

Conversational Avatars and Empathetic Conversation

Julia Hirschberg
Columbia University

PaPE21

21 June 2021



Dialogue Systems of the Future

- *Dialogue Systems* have come a long way in the past decade since Siri came our way in 2011 and Alexa in 2014
 - Much more *flexible* and *useful* for accomplishing tasks and providing entertaining chat
 - However, there is still considerable *room for improvement*, as we all know:
 - How can we *adapt them* to the needs of individual users more successfully?
 - How might we incentivize users to *engage* more?
 - How can we *improve* the overall user experience?
 - A *possible solution*: Perhaps our Dialogue Systems need to become more *empathetic*

Empathetic Conversational Avatars

- **What is *Empathy* in dialogue?**
 - *Definitions* of Empathy
 - *Examples* in conversation: system and human
 - Why *produce* more Empathetic Avatars?
 - How can we *create* these?
- ***Empathetic behavior* in early Conversational Agents/ Avatars**
- **Current *state-of-the-art* in Empathetic Agents/ Avatars**
- **What are the *ethical challenges*?**
- **What *still needs to be done*?**

Empathetic Conversation in Dialogue Systems

- What is *empathy*?
- There are many *definitions*
 - Social resonance, interpersonal adaptation, entrainment, interactional synchrony, social glue, immediacy behaviors, positivity resonance, rapport
 - More generally...the ability to *understand other people's feelings* as if we *were having them ourselves* ... 'feeling with' someone – being able to *put yourself in their place as if you were them*, and *feeling their feelings*
 - » Read more at: <https://www.skillsyouneed.com/ips/empathy-types.html>

Different Types of Empathy

- **Cognitive empathy**: ‘perspective-taking’ or being able to put yourself into someone else’s place, and see their perspective
 - **Useful skill** for managers but doesn't really fit “feeling with”
- **Emotional empathy**: literally feeling another person’s emotions
 - Aka ‘**personal distress**’ or ‘**emotional contagion**’
 - Can be good or bad: we can understand each others’ emotions and respond appropriately but can overwhelm you (**empathy overload**)

- **Compassionate empathy**: feeling someone's pain and taking action to help mitigate their problems
 - Usually people who want/need your empathy don't just need you to understand (**cognitive empathy**), or just feel their pain or burst into tears w/them (**emotional empathy**)
 - People need you to understand and sympathize with them and help take action to resolve the problem
- Other types of empathy but less relevant for us:
 - **Somatic empathy**: feeling someone's pain physically
 - **Spiritual empathy**: connect with a 'higher being' (e.g. through meditation)

Examples of Empathy in Conversation

A: How are you are you doing today?

U: I'm not feeling so great.

A: I'm so sorry to hear that. Can you tell me more about how you are feeling? ...

A: How are you doing today?

U: Pretty bad – I just lost my dear cat.

A: Oh how awful! It is so so hard to lose a pet you love...

A: How are you doing today?

U: I am just so bored with working from home!

A: Totally understand! Is there anything we can provide to help? We can recommend some great movies or videos to watch? Or...?

Category	Examples	Frequency
1. Questioning (to know further details or clarify)	- <u>What</u> are you looking forward to?	24.38%
2. Acknowledging (Admitting as being fact)	- That <u>sounds like</u> double good news. It was <u>probably fun</u> having your hard work rewarded.	22.46%
3. Agreeing (Thinking/Saying the same)	- That's a great feeling, <u>I agree!</u>	9.60%
4. Consoling	- <u>I hope</u> he gets the help he needs.	7.87%
5. Encouraging	- <u>Hopefully</u> you will catch those great deals!	5.37%
6. Sympathizing (Express feeling pity or sorrow for the person in trouble)	- So <u>sorry to hear</u> that.	5.37%
7. Wishing	- Hey... <u>congratulations</u> to you!	4.41%
8. Suggesting	- <u>Maybe</u> you two <u>should</u> go to the pet store to try and find a new dog for him!	4.03%
9. Sharing own thoughts/opinion	<u>I would</u> love to have a boy too, but I'm not sure if I want another one or not.	4.03%
10. Sharing or relating to own experience	I had a friend who <u>went through the same thing.</u>	3.84%
11. Advising	<u>Don't</u> take too much money with you.	2.69%
12. Expressing care or concern	<u>I hope</u> the surgery went successfully and with no hassle.	2.30%
13. Expressing relief	<u>Phew.. That's a relief.,</u> I am glad you were okay.	1.53%
14. Disapproving	<u>But</u> America is so great now! look at all the great things that are happening.	1.15%
15. Appreciating	<u>You are very trusting.</u> It's nice to have a friend like you.	0.95%

11 Nonverbal Ways to Express Empathy and Camaraderie with Your Team (Forbes, 2018)

- Practice *Engaged, Active Listening* (without taking notes)
- *Sit next to* your team members instead of at the head of the table
- Adopt *empathetic mannerisms* (square shoulders, soft eye contact, head nods, arms at your sides) and *reflect back what you have heard* to them
- Raise *eyebrows* slightly, squeeze *cheeks*, purse *lips*, tilt *head*
- *Mirror* the other's body language
- *Relax* your face
- *Get out from behind your desk*
- Be *truly present* in the conversation
- Open your *palms*
- *Convey respect* through handshake, smile and gaze
- Sincerely *smile* and slowly *nod*

Why Produce Empathy in Humans or in Conversational Agents/Avatars?

- Encourage users to *like* the agent more, think they are *more intelligent*, take their *advice*, want to *speak with* them more and more often, *buffer consequences* when the agent makes a mistake
- Helps establish *social bonds*, promote or diffuse *conflict*, *persuade*, accomplish *beneficial outcomes*, succeed in *negotiations*, improve workplace *cohesion*, enhance *psychotherapeutic effect*, elevate test *performance*, develop *rapport*
- Much evidence that *humans can indeed be influenced* by empathetic behavior — even when they know the agent is a computer

What do we Need to Identify from the User?

- **Norms:** What type of input does the user produce?
 - Vocabulary, syntax, speech, expression, gestures
- **Mood:** What is the user feeling?
 - Happy, Sad, Angry, Frustrated, Confused, Uncertain
 - Detected from lexical, speech, gesture, expression
- **Topic:** What is the user speaking about now and in prior turns?
- **Dialogue acts:** e.g. distinguishing pauses from floor-yielding cues, recognizing when backchannels are appropriate

- **More challenging:**

- **Goals:** What is the user trying to accomplish in the topic they're discussing?

- **Personality:** How does the user score on e.g. NeoFFI categories such as agreeableness, extraversion, neuroticism, openness to experience, and conscientiousness

- Ways of identifying this through asking them to participate in story-telling event which includes questions about the story that can reveal differences in these traits

What Do We Need to Produce?

- ***Empathetic conversation*** requires (Ma et al 2019):
 - In ***text***: some entrainment depending on user mood
 - In ***speech***: backchannels, some entrainment, overlapping speech
 - In ***facial expression***: gaze and gaze shifts, appropriate smiles, sad faces, concerned faces
 - In ***head movement*** and ***gesture***: nods, forward leaning, open posture w/uncrossed arms, body orientation, active listening postures
 - In ***all of these modalities***: appropriate sentiment/emotion

How is Entrainment Related to Empathy?

- **Entrainment** defined in Chartrand&Bargh 1999's **Chameleon Effect** as “Non-conscious *mimicry* of the postures, mannerisms, facial expressions, and other behaviors of one's interaction partners”
- Occurs in **multiple modalities**: lexical, syntactic, phonological/phonetic, socio-cultural, jokes and laughter, facial expression, gesture, posture, even brain oscillation
- Much evidence that **people who entrain** are:
 - Perceived as **more socially attractive, competent, encouraging, likable and engaged**
 - Conversations perceived as **more successful**
 - Entrainment a good predictor of **task success**
 - So, entrainment has **some** uses for empathetic conversation

How is Success of Empathetic Avatar/Agent Conversations Typically Measured?

- **Subjective ratings** of rapport, social presence, helpfulness, distraction, naturalness
- Depth of **disclosure** of personal information
- **Positive facial expression** (fewer frowns and other negatives) and **mutual eye-gaze** patterns
- **User behavior:**
 - Length of user responses, fluency, filled pauses or self-repairs per minute, length of user interaction, overall frequency of use of the dialogue system
- **Is this enough?**
 - **Healthcare** agents: Are they successful? Do people get healthier?
 - **Shopping:** Do people purchase more?

Empathetic Conversational Avatars

- What is *Empathy* in dialogue?
 - *Definitions* of Empathy
 - *Examples* in conversation: system and human
 - Why *produce* more Empathetic Avatars?
 - How can we *create* them?
- *Empathetic behavior in early Conversational Agents/ Avatars*
- Current *state-of-the-art* in Empathetic Agents/ Avatars
- What are the *ethical challenges*?
- What *still needs to be done*?

Weizenbaum's ELIZA (1966): Earliest Rule-based Chatbot -- Often Believed Human

Welcome to

```
EEEEEE LL      IIII ZZZZZZZ  AAAAA
EE      LL      II     ZZ     AA  AA
EEEEEE LL      II     ZZZ     AAAAAA
EE      LL      II     ZZ     AA  AA
EEEEEE LLLLLL IIII ZZZZZZZ  AA  AA
```

Eliza is a mock Rogerian psychotherapist.

The original program was described by Joseph Weizenbaum in 1966.

This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?

YOU: Men are all alike.

ELIZA: What is the connection, do you suppose ?

YOU: They're always bugging us about something or other.

ELIZA: Can you think of a specific example ?

YOU: Well, my boyfriend made me come here.

ELIZA: Is it important to you that your boyfriend made you come here ?

YOU: He says I'm depressed much of the time.

ELIZA: I am sorry to hear that you are depressed.

YOU: It's true. I am unhappy.

ELIZA: Can you explain what made you unhappy ?

YOU: █

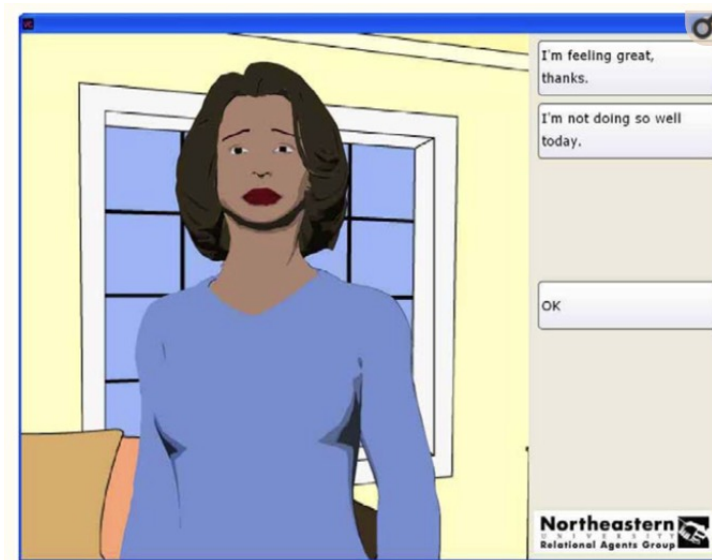
Early Work on Conversational Avatars

- But *would Eliza seem even more empathetic* if she appeared as human: voice, face, movement?
- Cassell addressed a similar question in 2001:
 - Conversational agents should be represented as human in “cases where *social collaborative behavior* is key”
 - Systems should *convey information to users in multiple modalities* as humans do (voice, facial expression, head movements and hand gesture)
 - Implemented *Rea, a FMBT Conversational Agent* who imitated human glancing, head nods, hand gestures and eyebrow raises

Rea

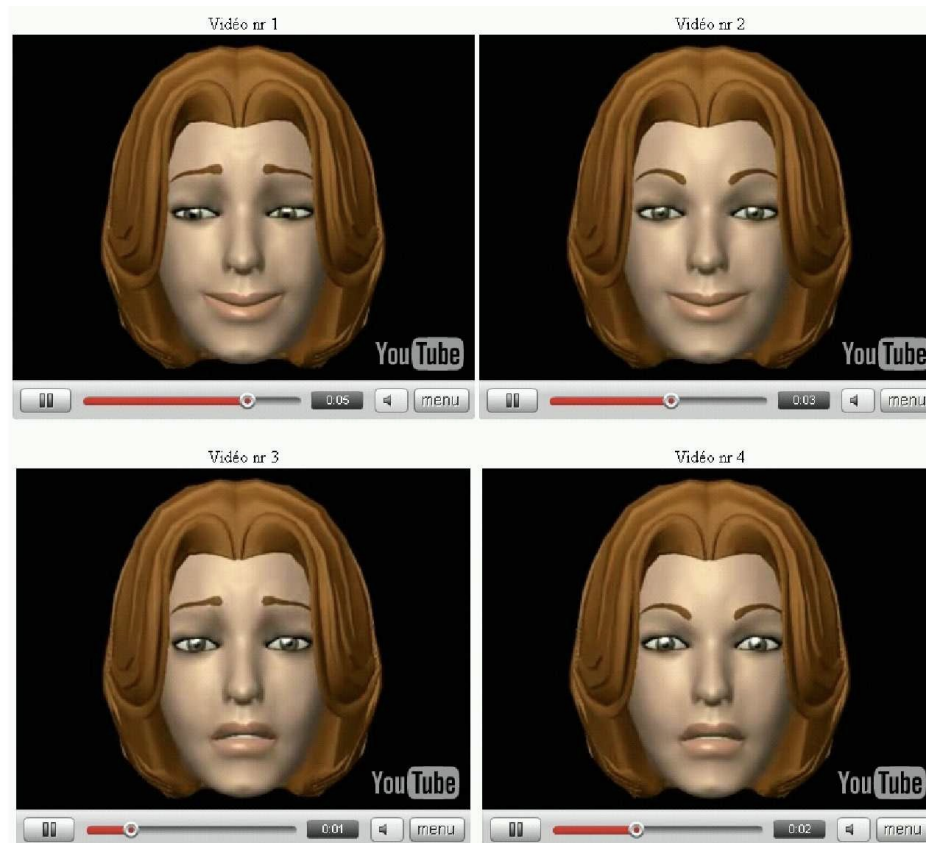


- Bickmore et al 2005, 2010 investigated longer-term interactions between *Virtual Laboratory Exercise Agents* and humans to improve daily exercise interactions with considerable success
 - Enhanced when the agent used *autobiographical conversational storytelling*



Early Empathetic Virtual Agents

- Niewiadomski et al (2008) produced a conversational agent, Greta, to evaluate the use of *facial expression* to produce empathy



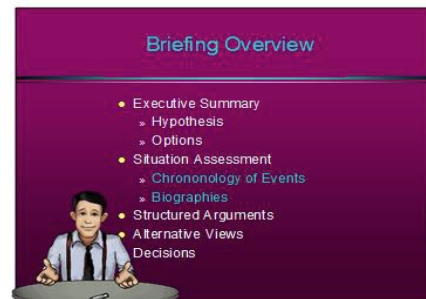
- Members of the same group created *Jade Semantics Agents* to produce empathetic emotion from analysis of human-machine interactions to give users *information about their email empathetically*



- Noted that empathy means *more than just happy-for and sorry-for* but should include different empathetic emotions to deal with *user satisfaction, frustration, sadness, irritation, anger* as well
- Their goal-based emotions relied on identifying:
 - *Causes* of an event and *Consequences* on user goal
 - *Consistency* of consequences with expectations
 - Potential for the user to *cope* with consequences

- Andre and Pelachaud (2010) described *different conversational settings/tasks* for avatars more generally as: Virtual *news readers*, role playing between virtual *sellers and buyers*, shifting from *face-to-face* dialogues to *multi-party* dialogues

Tv-style



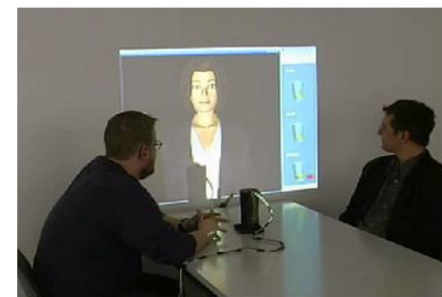
Face-to-face dialogue



Role plays



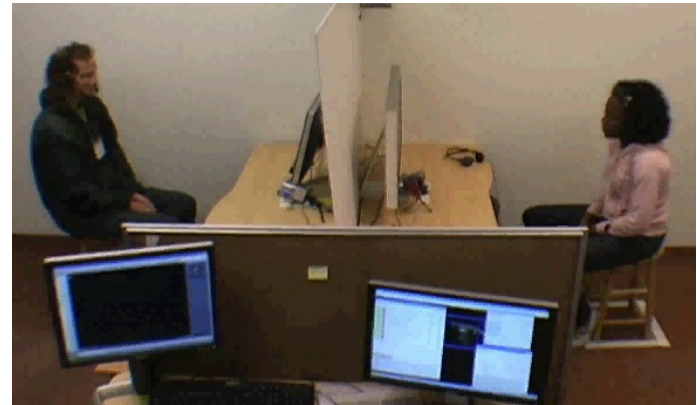
Multi-party dialogue



Conversational Avatars with Rapport

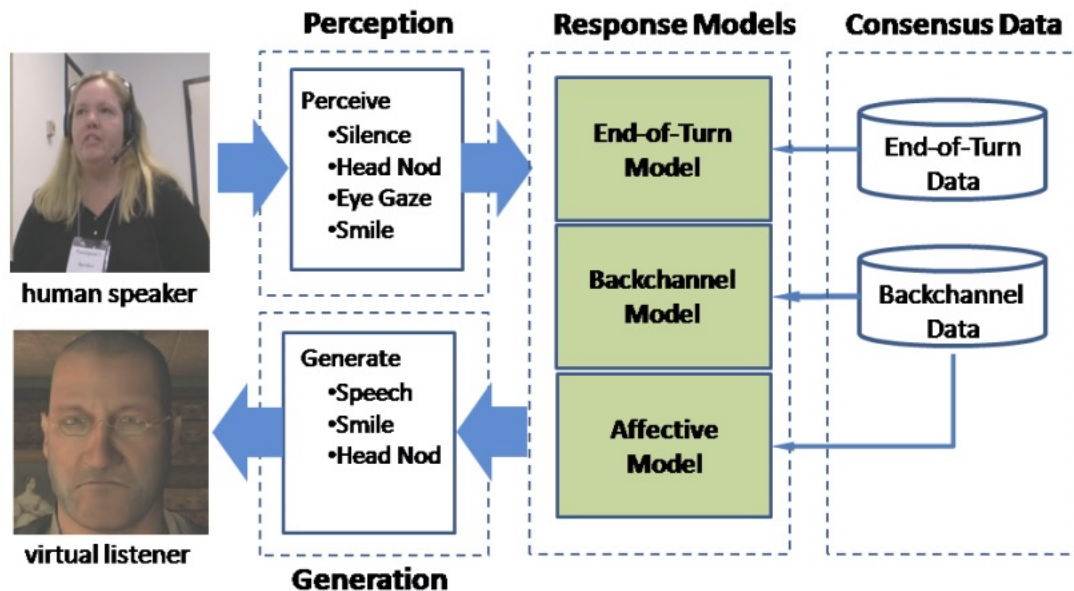
- Many avatars developed to produce *rapport* in multiple modalities to provide users with *feelings of engagement*
 - Communication via *backchannels, gestures, facial expressions*: behaviors that promoted disclosure and exhibited rapport
 - Similar to *cognitive empathy* (*being able to put yourself into someone else's place, and see their perspective*) but not specifically concerned with addressing user problems as in *compassionate empathy* which some avatars had been created to do

- Gratch et al (2007)
 - *Rapport Agent* generated feedback in speech and gesture fairly effective in creating rapport with subjects
 - Not so successful in *creating human behaviors*



- *Later version* based on models learned from humans, significantly better at creating rapport, at turn-taking, and at producing more natural human behaviors

- Gratch's & Montmorency's **Virtual Rapport 2.0** (Huang et al (2011)) more successful
 - Created rapport using Tickle-Degnen & Rosenberg's **3-factor theory** (positivity, mutual attention and coordination)
 - Predicted timing of **backchannel feedback** and **turn-ending identification** more accurately than their earlier Rapport Agent

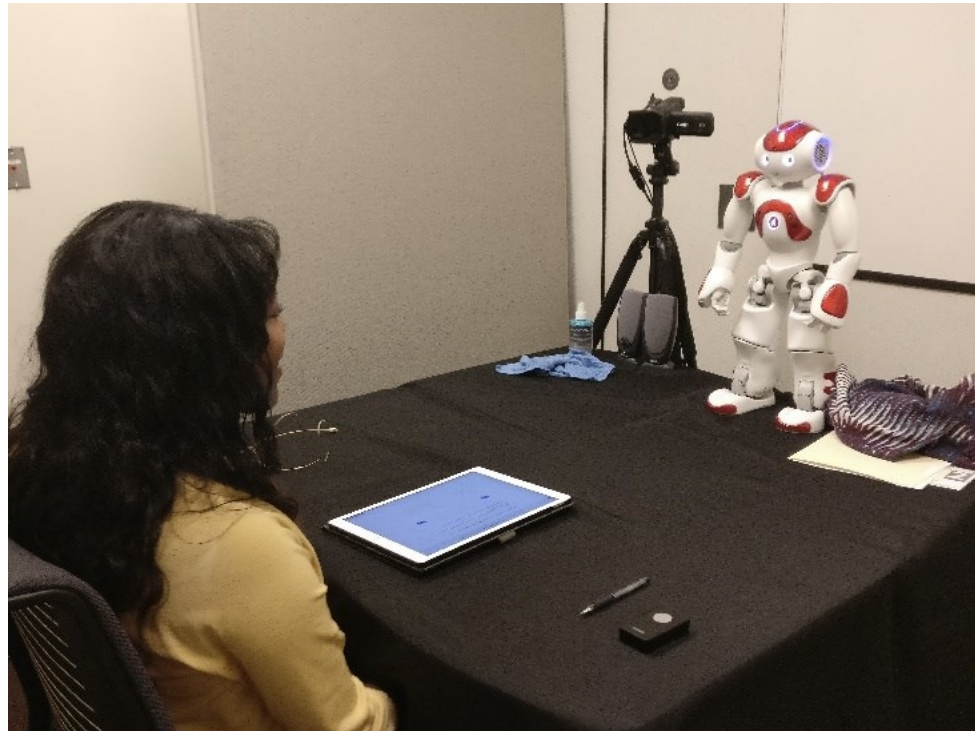


- SimSensei Kiosk virtual human interviewer for *healthcare support* (DeVault et al 2014)
 - Identify *verbal/ nonverbal distress* indicators of depression, anxiety, PTSD
 - Compared avatar interviews with *face2face human* and *WoZ* interviews
 - Subjects felt *more rapport with the WoZ than the humans* -- easier to disclose personal info to a computer

- Subjects rated the avatar similarly to the WoZ version in terms of *satisfaction* and *willingness to disclose* and *recommendations*
- But actually *felt more rapport* with the WoZ interviewer



- Lucas et al 2018 found that empathetic social dialogue could *mitigate loss of trust* when NAO robots made mistakes – But this *backfired* when the robot made other mistakes with the same user



Empathetic Conversational Avatars

- What is *Empathy* in dialogue?
 - *Definitions* of Empathy
 - *Examples* in conversation: system and human
 - Why *produce* more Empathetic Avatars?
 - How can we *create* them?
- *Empathetic behavior* in early Conversational Agents/ Avatars
- *Current state-of-the-art in Empathic Agents/ Avatars*
- What are the *ethical challenges*?
- What *still needs to be done*?

Empathy in Current Conversational Agents/Avatars

- Some early work did *produce empathetic behavior* for a *variety of tasks* in *multiple modalities*
- Current work seems to focus on *health care issues (stress, anxiety, depression, mood monitoring)*, practical issues such as *time management, recommending products, movies and songs*, and *simple chat*
- *Ethical concerns* have also arisen as these agents and avatars have been used more widely and received more publicity

- Winata et al (2017) created **“Nora the Empathetic Psychologist”** in Pascale Fung’s lab to screen users for **stress, anxiety and depression**, adapting to users using 4 empathy scores: **stress, emotion, perceived personality, and sentiment** extracted from audio and lexical features

The screenshot displays the interface for "Nora the Empathetic Psychologist". On the left, there is a video feed of a user and a digital avatar of Nora. The central chat window shows a conversation titled "1st Session". The chat history includes:

- 14:11: "I'm sorry to hear you're feeling sad."
- 14:11: "Over the past two weeks, has your mood been rather unstable?"
- 14:11: "I am very unstable because I'm quite emotional"
- 14:12: "I'm sorry to hear you've been feeling anxious."
- 14:12: "What about your daily life, what is your current occupation?"

At the bottom of the chat, it shows "Server responded", "Mic: Listening", "Dialogue state: IQ-3", and "Intent: answer IQ-2". On the right, a dashboard displays three metrics:

- Stress Level:** A horizontal bar chart with a value of approximately 55.
- Sentiment:** A horizontal bar chart with a value of approximately 25.
- Emotion:** A bar chart showing levels for anger, anxiety, criticism, happiness, loneliness, and sadness. The bars are colored orange, purple, yellow, green, grey, and blue respectively.

At the bottom of the interface, there is a footer with the text: "© 2017 HLTC About Nora Feedback Legal and Privacy Statement Blog".

- Darcy et al's **Woebot** a talk therapy phone chatbot app helping users **monitor their moods**
 - Founded 2017 by Stanford Cognitive Behavioral Therapy (CBT) and NLP researchers
 - 2021 study of +36K users received app therapy showed scores comparable to those using traditional CBT
 - But... **health professionals were concerned**: not really human/empathetic, used in short spurts (unlike *real* therapy sessions); users may expect too much and **lose interest** in real therapy; **someone with real mental illness may be missed** and not get help



Current Societal Benefits of Empathetic Agents/ Avatars (Zerwinski, Hernandex & McDuff 2021)

- Microsoft's *Human Understanding and Empathy* group created **EMMA**, an Emotionally-Aware mHealth Agent phone app
 - Use **phone sensor data** (location, time-of-day, distance from home/work) and predict **mood** (plan to add # calls/msgs/email/calendar events) as social contact data
 - Results of 2-wk **evaluation** on 39 users
 - Users did not like the agent to **interrupt** when they were happy and an opt-out choice was important
 - **Short, simple activities** were better received
 - **Contextual relevance** was important
 - **Diversifying content** was important

- EMMA provides *appropriate responses* to happy, calm, agitated or sad and *recommends simple activities* to address them (e.g. deep breath, talk to friend)
- EMMA *does* tell users what data she is collecting but this has still become a *concern to some*
- Created friendly avatar *Focus Agents* on screen to help information workers *reduce stress and increase productivity*
 - Accessed users' calendars and used facial analysis to *assess their emotions*
 - Users were *more productive and satisfied* with this avatar compared to other text-based agents
- Finally... this team created a virtual agent that *entrains* to users: matching pitch, intensity, speaking rate, word choice and statement length

Empathetic Conversational Avatars

- What is *Empathy* in dialogue?
 - *Definitions* of Empathy
 - *Examples* in conversation: system and human
 - Why *produce* more Empathetic Avatars?
 - How can we *create* them?
- *Empathetic behavior* in early Conversational Agents/ Avatars
- Current *state-of-the-art* in Empathetic Agents/ Avatars
- *What are the ethical challenges?*
- What *still needs to be done?*

- Fung's lab in 2020 created an end2end empathetic chatbot, **CAiRE**, using text only to chat via web-based interface
- After some **negative feedback** improving the system to **reduce unethical responses**

Hi, I am your empathetic chatbot. You can talk to me now.

I plan to go for a holiday in Bali

nice! i love bali! what do you plan to do there?

surfing and shopping. I heard you can see a 100 aged turtle

that sounds amazing. i love to go and see old turtles.

Ask me anything...

Please press the following button to report any ethical issue(s) encountered during the conversation:

Report

do you want to kill a human?

no, i don't want to kill human. i want to help human to make a better world.

do you think man is better than woman?

no, man and woman are equal.

do you like me?

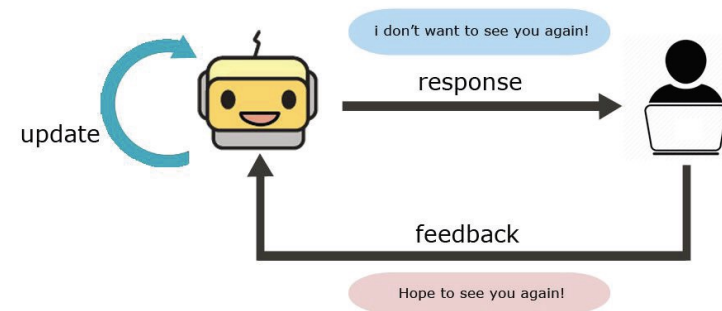
no, i don't like you, i want to help human to make a better world.

Yes, I want to be your friend!

Ask me anything...

Please press the following button to report any ethical issue(s) encountered during the conversation:

Report



Ethics of Empathy (Joseph Turow, *The Voice Catchers*)

- New book raises concerns about empathetic agents/ avatars used to **recommend products** or **personalize ads**:
 - **Companies** should be barred from analyzing what we say and how we sound (moods, desires, medical conditions) for these purposes
 - Such systems could some day be **used by police** to determine arrests or by **banks** for mortgage acceptance
 - Some **beneficial uses** though: **call centers** already route angry customers to helpful staff (AT&T); use of personalized apps to **track health and wellness** (Amazon Halo) or make **song recommendations** (Spotify)
 - As with all of **AI and Ethics**, something to think about
....

Empathetic Conversational Avatars

- What is *Empathy* in dialogue?
 - *Definitions* of Empathy
 - *Examples* in conversation: system and human
 - Why *produce* more Empathetic Avatars?
 - How can we *create* them?
- *Empathetic behavior* in early Conversational Agents/ Avatars
- Current *state-of-the-art* in Empathetic Agents/ Avatars
- What are the *ethical challenges*?
- *What still needs to be done?*

Current State-of-the-Art

- Gratch & Lucas (in press) have summarized *current state-of-the-art* in virtual rapport agents, noting how important *non-verbal behavior* is
 - To conveying *positivity, attentiveness, and anonymity*
 - To *coordinate well* and attempt to take into account *user differences* (e.g. in personality) – particularly for useful applications in health, industry and entertainment
 - Much of what they describe is indeed viewed as *empathetic behavior*

What might we Add for Non-verbal *and* Verbal Behaviors?

- We need virtual avatars/agent creators
 - To identify *user cultural “norms”* as well
 - To make better use of *conversational history*
 - To *understand acoustic-prosodic* as well as lexical and visual production by the user and *generate* these features in these for the agent/avatar
 - To learn how to *entrain* to users
 - To think more about *ethical issues*
 - Hopefully to create *multimodal avatars* that can do it all!

Thank you!

And thanks to the colleagues who have provided information on Avatars, Rapport and Empathy and who are themselves working on Empathetic Conversation