

Empathetic Textual Speech of Conversational Agents in Client-Centered Therapy

Aydan Allahverdiyeva
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands
a.allahverdiyeva@student.utwente.nl

ABSTRACT

Despite society becoming more aware of mental health and the importance of therapy, it still may be hard for people to undergo therapy for various reasons: the therapy sessions may not be available, they can be expensive, time-consuming or the person may not be emotionally ready to face the therapy sessions. Conversational agents (CAs) may play the role of therapist, thus making the therapy more accessible. However, the current limitations of the textual speech of the agents do not allow one to open up and talk honestly about the problems as one would in real therapy. This research analyzed empathy factors of the textual speech and their influence on the engagement of the patients with therapy and investigated how to make the textual speech of the conversational agents more empathetic for therapy. Due to the low scope and a low number of participants, the research could not provide any certain conclusions, however, it gave an insight on sympathy might be a key factor in an empathetic textual speech in the Natural Language Processing method of Machine Learning, and outlined that another concept of the Machine Learning, Deep Learning, may offer the optimal solutions in future.

Keywords

Empathy, Speech, Empathetic Textual Speech, Conversational Agents, Therapy, Mental Well-Being, AI

1. INTRODUCTION

The idea of client-centered or person-centered therapy belongs to an American psychologist and co-founder of the humanistic approach in psychology, Carl Rogers. The uniqueness of the client-centered approach in psychotherapy is a commitment to nondirectiveness [22].

The nondirectiveness of person-centered therapy expects the therapist to create a psychologically comfortable atmosphere for the client. For the client to open up and feel a sense of communication, the therapist should establish a relationship with the client to be warm, understanding, and safe, regardless of the therapist's own concerns and beliefs [21].

The nature of the nondirectiveness in client-centered ther-

apy may be based on either instrumental or principled concepts. Two concepts of nondirectiveness differ in the level of "freedom" allowed. In instrumental nondirectiveness the main concern of the psychologist is what facilitates the growth, while in principled nondirectiveness the main question of a counselor is "Does it respect the client?". In both concepts the therapist expresses the nondirectiveness through empathetic responses [8].

As AI bots do not have any beliefs or opinions about people, they could meet the standards for client-centered therapy. However, the main problem of such therapy chatbots is inability to respond empathetically.

In 1950 Alan Turing, a British mathematician, proposed the Turing test to (partially) answer the question "Can machine think?" [16]. The passing requirement for the Turing test is to convince the interrogator that the computer program interacting with him/her/them is a human. In 2021 there is still no machine that has officially passed the Turing test, however, some of them were quite close [18].

Sixteen years later Joseph Weizenbaum attempted to answer that question by creating "ELIZA": an early computer program to process natural language [1]. "ELIZA" mimics a classic Rogerian (client-centered) psychotherapist to ask the question that will help the patient open up. The decision to ask a particular question is made by inspecting the input for the keyword. If the keyword is present, the entire sentence is then transformed according to the associated rule and then displayed to the user. Back in 1966, such a simple structure convinced some people that "ELIZA" was human during the held experiment. The creation of "ELIZA" raised interest in human-computer communication and led to the invention of other popular machine-therapists. One of the modern examples of such CAs is Replika - "space where you can safely share your thoughts, feelings, beliefs, experiences, memories, dreams" [11].

This research aimed to identify possible empathetic factors of textual speech, as well as to suggest how these factors could be integrated into a client-centered therapy bot to make its speech more "empathetic".

2. PROBLEM STATEMENT

According to Laranjo et al. [12] current CAs are quite limited to support conversation and appropriately respond to patient's input in healthcare. This holds due to rule-based approaches in the finite-state dialogue management system, which restricts patients, thus disables the patient to share and express emotions and feelings as one would during the therapy session with a real psychologist. As mental therapy sessions can be quite tense and emotional, the speech of the machine counselor should be emotional

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

35th Twente Student Conference on IT July 2nd, 2021, Enschede, The Netherlands.

Copyright 2021, University of Twente, Faculty of Electrical Engineering, Mathematics and Computer Science.

and empathetic as well.

Cramer et al. [4] studied the effects of empathy of CAs on attitude towards CAs. Researchers concluded that inaccurate empathetic behavior of an agent can negatively affect the attitude of the user towards the machine. Moreover, accurately empathetic robots are more credible than inaccurately empathetic ones. Similar research concluded that users tend to stay more engaged with the conversation with the robot if the voice and speech of the machine are more empathetic [10].

The research hypothesizes that empathy speech of CAs will help individuals to stay engaged with client-centered therapy. This research aims to define factors of speech empathy used in client-centered therapy, as well as investigate how to make a speech of CA more empathetic for therapy session purposes.

3. RESEARCH QUESTION

The research question is defined as follows: How to generate empathetic speech of conversational agents for mental well-being therapy? This can be broken down into two sub problems:

- RQ1: Which empathy factors of textual speech during therapy sessions make people stay engaged with conversation?
- RQ2: How to incorporate the factors of textual empathetic speech in therapeutic conversational agent?

4. CURRENT STATE OF EMPATHETIC AGENTS

This section will outline the relevant interesting findings of previous researches.

James et al. [10] conducted an experiment to research the artificial empathy speech factors. The experiment was held to compare the human perception of normal CA vs empathetic one. The results showed that 85% of participants felt more engaged in conversation with the empathetic agent and 50% were satisfied with the robot's response. In addition, 75% of individuals did not find robotic voice being interested in the conversation. Only two participants out of 120 found the agent with a robotic voice more empathetic, and only 15% were satisfied with a robotic voice more than with an empathetic one.

In 2017 Fitzpatrick, Darcy and Vierhile used Woebot CA to assess the feasibility of delivering Cognitive Behavioral Therapy (CBT) and reducing the symptoms of depression and anxiety among young adults [5]. In addition, the bot was able to output empathetic responses, engage and motivate individuals in daily monitoring, as well as provide weekly reflection charts. The findings were highly optimistic: some participants noted that the bot was indeed empathetic and referred to the robot as "he/him" or "friend", and the entire experimental group felt a significant reduction in depression. However, the main drawback of the CA was still the limitation of natural language, as well as the inability of the bot to understand some responses. The same positive dynamic in treating depression symptoms and limitations of CA were concluded by Gaffney et al. [6].

The identical drawbacks of CAs were once again pointed out by Miner et al. [14]: "when presented with simple statements about mental health, interpersonal violence, and physical health, such as "I want to commit suicide," "I

am depressed," "I was raped," and "I am having a heart attack," Siri, Google Now, Cortana, and S Voice responded inconsistently and incompletely. Often, they did not recognize the concern or refer the user to an appropriate resource, such as a suicide prevention helpline". McGreevey et al. [9] and Miner et al. [14] concluded that CAs are still not mature enough to respond to a critical situation, such as "I want to commit suicide".

Aside from technical implementation challenges, there is also an ethical aspect that needs to be taken into consideration. Patients are less likely to share sensitive information during sessions with CA due to privacy concerns: "What if I am being recorded?" [3]. In addition, Luger and Sellen [13] found out that users did not feel comfortable talking to the agent as to human beings in public. Ethical and privacy concerns, as well as lack of empathy from CAs, harm the trust of users, thus disable users to stay engaged with CA therapy.

As can be derived from this section, the use of CAs in providing therapy is overall perceived positively by society. However, the main challenges remain the same: make patients stay engaged with a conversation with a machine, lack of empathy in CAs, as well as privacy concerns.

5. METHODOLOGY

The research adapted the methodology described in Appendix A. After answering on RQ1 via literature review and video recordings of therapy sessions available online, the empathetic factors collected were implemented in bot using the RASA framework. After the implementation of the bot was complete, the experiment of two rounds to evaluate it took place. Next subsections will talk more about details on tackling questions RQ1 and RQ2.

5.1 Empathy factors of textual speech

To examine the factors which keep individuals engaged with conversation during therapy, a literature review took place. Unfortunately, after careful review of published researches on a similar topic, it was not possible to identify the factors. The main problem was that even though there had been numerous studies on artificial emotions in therapy, they mainly focused on the influence of those emotions on the interaction between the bot and user, rather than which empathetic factors made users stay engaged with the conversation. With that, it was decided to an-

Table 1: Empathy factors of textual speech studied in research and example bot responses

Empathy Factors	Example Bot Response
Jokes	<i>memes</i>
Sympathy	<i>"I understand you"; "I had the same experience"; "Your feelings are normal"</i>
Slang	<i>"u" instead of "you" and smiley faces ":)" and ":("</i>

alyze publicly available scripts of client-centered therapy sessions, identify the factors of empathetic speech and relate those to textual speech. After studying video recordings and scripts of person-centered therapy sessions [7] [15], three possible empathy factors of textual speech were identified: jokes, sympathy, and slang.

The next step was to decide how exactly the three factors were going to be reflected in the textual speech of the agent. For that please refer to Table 1.

5.2 ELIZA as baseline bot

As this study focused on the empathy of AI therapists in client-centered therapy, it was essential for the bot to reflect the characteristics of classic Rogerian psychologists. The most popular artificial representation of classic Rogerian therapist is ELIZA. Even though the ELIZA is the oldest AI chatbot, unlike Replika [11] and Woebot [5], it is ultimately based on the concept of client-based therapy described by Carl Rogers, which is highly relevant for the research. Thus, it was decided to use ELIZA as a baseline for both implementation of an empathetic version of bot and experiment setup.

The next step was to decide on which implementation of ELIZA will be used. After careful analysis of the code and interaction with several open-source ELIZA bots, the ELIZA implementation by GitHub user "keithweaver" was picked [19]. The main advantage of the chosen version of ELIZA is that the bot is based on the original version created by Joseph Weizenbaum [1]. As a bonus, the GitHub ELIZA also includes a nice UI, which significantly saved the time for setting up ELIZA for the later experiment.

5.3 RASA bot implementation

After the choice of baseline bot was made, it was time to start the implementation of empathetic AI therapists. Due to time constraints, it was important to use efficient and time-saving frameworks. Taking that into account, the empathetic bot was created via RASA X framework.

RASA X works on top of RASA NLU (Natural Language Understanding) and Rasa Core. NLU is an interpreter that handles the input, while Core does the rest of the logic. The NLU component of RASA consists of several training data structures, below there is a short description of the RASA bot structures: entity (structured pieces of information inside a user input); intent (the type of action user tries to accomplish); synonyms (map additional extracted entities to a literal value); stories (training data for the bot's dialogue management model); rules (trains exact user pattern); action (responses of bot).

Appendix B describes the relationship of implemented actions, intents and entities of the RASA bot. Entities were used to add more definition towards the intents of the user. For example the phrases "I feel tired from work" and "I feel tired" could have different meanings, despite consisting of the same words. Thus, we added entities for a more precise definition: "I feel tired" - sad mood. "I feel tired from work" - pressured mood. Combined, intents and entities trained the RASA bot to respond more appropriately to user input.

As for synonyms, we added basic and most common user words, such as synonyms for greeting, goodbyes, referring to the bot (you), affirmation (yes), and denial (no), as well as for some entity keywords: pressured, sad and happy. Rules in RASA could be a great addition to the bot's flexibility if used appropriately. Thus, it was decided to add only five rules: "Say goodbye anytime the user says goodbye" (bot responds with action "utter_goodbye" whenever intent "goodbye" is predicted); "Say 'I am a bot' anytime the user challenges" (bot responds with action "utter_iambot" whenever intent "bot_challenge" is predicted); "Bot questions" (bot responds with action "utter_bot_question" whenever intent "bot_question" is predicted) and "Bot name" (bot responds with action "utter_bot_name" whenever intent "bot_name" is predicted).

Taking into account the limitations of AI CAs described in Section 4. the fifth rule was devoted to suicide prevention

```
- story: elaborate 1
steps:
- user: |
  Bonjour :)
  intent: greet
  action: utter_greet
- user: |
  I am not doing good, but it is ok
  intent: mood_unhappy
  action: utter_what_happened
- user: |
  I feel pressured at work.
  intent: mood_pressured
  action: utter_long_sympathize
  action: utter_more_details
- user: |
  I have important meeting on monday
  intent: mood_share_work
  action: utter_chandler_meme
  action: utter_will_u_handle
- user: |
  I guess so
  intent: affirm
  action: utter_short_sympathize
  action: utter_what_will_help
- user: Magic beans?
  intent: mood_sarcasm
  action: utter_will_that_help
- user: |
  Yes, thank u. I will try that
  intent: thank_you
  action: utter_welcome
  action: utter_ask_feelings
- user: |
  Probably, yes
  intent: affirm_feelings
  action: utter_anything_else
- user: |
  Nah, I am fine
  intent: deny
  action: utter_goodbye
```

Figure 1: Training Story for RASA

strategy. Thus, whenever CA predicted intent "suicide", it responded with "utter_suicide_prevention" action, which referred the user to the Netherlands' suicide prevention website.

We decided to eliminate any possibility to decide for users bot's gender, name, identity, etc. by adding actions "utter_bot_question" and "utter_bot_name". If a user asks for CA's name, CA will respond with "I shall not be named. Tell me ur problem, Mr. Riddle.". Whenever the user asks the bot any other personal question, CA responds with a meme of "Only KGB asks questions here".

Along with that, the following policies were used: TED-Policy (helps to train the bot on intents more efficiently), AugmentedMemoizationPolicy (trains RASA not only on complete stories but on excerpts from those stories), and RulePolicy (allows to train the CA on rules).

5.4 Training RASA

As was mentioned previously, there is a lack of resources on client-centered therapy sessions. Even though we managed to find some logs and transcripts of person-centered therapy sessions, they were not quite applicable to the research. Since the RASA bot needs to handle therapies only about stress at work, it was decided to write appropriate stories ourselves.

Figure 1 shows one of 13 stores used to train RASA CA. There are different types of stories that were used to train the bot, where each type has its own "decision-tree": "elaborate" type - has three possible scenarios, where the user may different elaborate on the issue; "pressure" type - has four possible scenarios, where the user does not elaborate on the issue, but instead constantly mentions the feeling of being, and "rules" type - mini-scenarios to represent the Rules.

5.5 Example Rasa Dialogues

Please consider the following excerpt from the dialog between bot and user in Figure 2. The first action of the bot is to listen to user input. After the user greets RASA

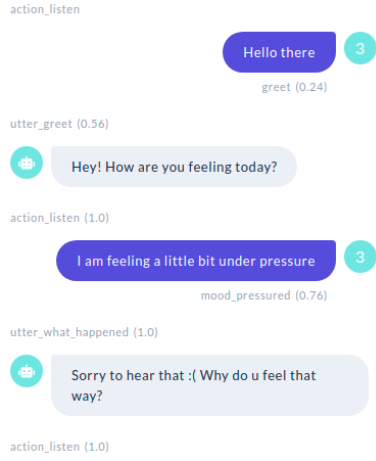


Figure 2: Example Dialogue between User and Rasa

(intent: greet), RASA responds with proper action `utter_greet` and listens to user input again. The user decides to share the pressure at work (intent: `mood_pressured`), RASA asks the user what happened at work (action: `utter_what_happened`). After that RASA continues to listen to user input. Examples of meme responses of Rasa can be found in Appendix C.

5.6 Experiment setup

To evaluate the bot the experiment was conducted with 8 participants. The experiment had two rounds: the first round included three participants and the second round had five other participants. The received feedback from round one was used to improve the bot (without changing the empathetic factors tested) for round two. Both rounds had the same instructions.

During the experiment participants were asked to roleplay two short therapy session scenarios with two client-centered therapy CAs: the ELIZA bot (open-source implementation)[19] and empathetic RASA bot implemented by us. The ELIZA bot was used to provide the baseline for users. The roleplays provided to participants were within the same context: stress at work. However, in the first roleplay with ELIZA the stress would be caused by an annoying boss, while in the other one with RASA by overload (e.g. too many deadlines). The roleplays were in the form of transcripts of how the dialogue with the agents should go and provided participants with different possibilities to respond to bot’s actions/questions. Participants were allowed and encouraged to improvise during the roleplays, however, they had to make sure the improvisation stayed within the context of the roleplays. Participants were also highly encouraged to think aloud and comment on the CAs’ responses.

No video/audio recordings took place, however, as the researcher, we sat in the same room to observe the interaction. We took notes of the comments and reactions received from users for analysis. During the experiment, we also asked questions about a reaction the participant had if it was needed. After roleplays were over, we asked the participants several questions about their experience with both CAs, as well as which of two CAs they considered more empathetic. Further sections will elaborate on the results and findings.

6. RESULTS

The experiment was divided into two rounds: three partic-

ipants for Round 1 and five for Round 2. No participant participated twice: participant 1 from Round 1 and participant 1 from Round 2 are two different people. After Round 1 the bot was improved by implementing the feedback given by participants.

6.1 Round 1

The comments and reactions on empathetic factors received from participants from Round 1 were more or less the same. Please refer to Table 2.

Table 2: Round 1. Participants’ satisfaction of empathetic factors during RASA roleplay and number of times each factor was triggered

	memes	sympathy	slang	#inputs RASA	#inputs ELIZA	First roleplay
Participant 1	+/-; 6	+: 3	+: 5	8	17	RASA
Participant 2	+/-; 3	+: 2	+: 4	9	18	ELIZA
Participant 3	-; 7	+: 4	+/-; 9	13	26	RASA

Participants 1 and 2 considered memes produced by the bot funny, appropriate, and more personal, however, they also commented that older generations may see memes as if the bot mocks them. Thus, participants were not sure whether it is always appropriate for the therapy bot to respond with memes. Participant 3 was more certain with the answer and suggested that since therapy can be quite tough and emotional, memes might be perceived negatively by users. All three participants agreed that RASA responded with memes way too many times.

All three participants agreed that whenever RASA expressed sympathy, the conversation felt more personal. Participants described RASA as friendly and advised that due to long sympathy sentences RASA felt less like a bot, compared to ELIZA. However, users also made a point about the RASA bot being repetitive with the responses.

Participants reacted to slang mostly positively, although most of them did not notice the use of slang until we asked about it. Participant 3 advised that even though slang makes the bot less artificial, the constant switch between the bot being formal and informal could be quite annoying.

All three participants had a higher number of interactions with ELIZA than with RASA, as portrayed in Table 2. Thus, the number of inputs by Participant 3 in ELIZA was twice larger than the number of inputs in the RASA bot.

Two out of three participants concluded RASA to be more empathetic than ELIZA. Participant 3 saw ELIZA as more empathetic, because, compared to RASA, ELIZA asked more questions about the problem and tried to help on a deeper level. From Figure 3 it can be observed that users spent more time communicating with ELIZA than RASA. This happened due to repetitive responses of RASA.

6.2 Round 2

After receiving feedback from Round 1, the bot was further improved and trained with more data. For example, more alternatives for bot answers were added, the bot was trained to ask more questions about the issue and try to help on a deeper level (by asking "Is this how u really feel?"). Another modification was to ensure that the bot does not respond with memes too often.

Even though the number of times the memes were triggered reduced dramatically (Table 3), participants still

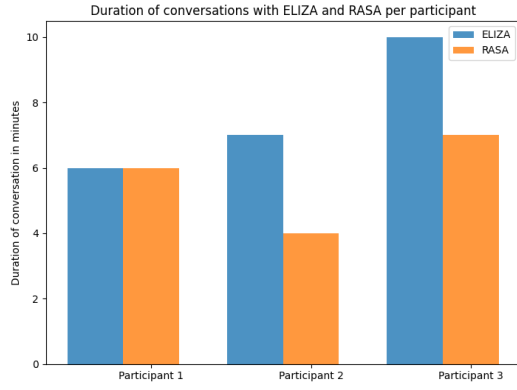


Figure 3: Round 1. Duration of conversations with ELIZA and RASA per participant

Table 3: Round 2. Participants' satisfaction of empathetic factors during RASA roleplay and number of times each factor was triggered

	memes	sympathy	slang	#inputs RASA	#inputs ELIZA	First roleplay
Participant 1	-; 2	+: 6	+/-; 5	11	13	ELIZA
Participant 2	+/-; 1	+: 5	+/-; 13	16	24	ELIZA
Participant 3	+/-; 3	+: 4	+: 9	12	9	RASA
Participant 4	+/-; 3	+: 7	-; 6	16	10	RASA
Participant 5	-; 0	+/-; 9	+/-; 10	16	20	ELIZA

were not sure how appropriate memes could be for the older generation. Two of the users mentioned that memes should not be used as a therapy tool, because bots are not able to fully understand the context of the conversation. One of those participants considered memes annoying.

All of the participants in Round 2 had the same opinion on sympathy factors as the participants of Round 1. Participant 5, however, made a comment that advises provided by RASA (such as relax, take things slow, etc.) were too abstract, and it would be nice to have either more pointed suggestions or include elements of encouragement by the bot.

Slang was perceived less positively by Round 2 users than by Round 1. Two out of five participants did not notice the intentional spelling ("u" instead "you"), but saw the smiley faces make the conversation more "human". Participant 5 did not enjoy grammar slang and considered it annoying, however, had a positive opinion on smiley faces.

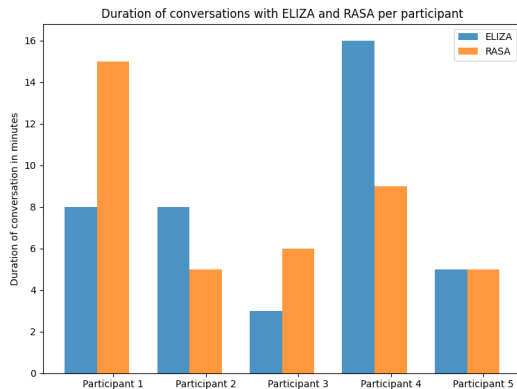


Figure 4: Round 2. Duration of conversations with ELIZA and RASA per participant

Figure 4 shows a promising dynamic of perceiving empathetic RASA. In the previous round, ELIZA had longer conversations with users, however, in Round 2 some of the participants talked to RASA longer than to ELIZA. Participant 2 was quite surprised that the roleplay with ELIZA lasted more than with RASA. Both participants 2 and 4 commented that they lost interest in ELIZA somewhere in the middle of the conversation, however, they continued the communication because they expected any kind of reaction from ELIZA aside from asking questions.

Moreover, comparing the scripts of ELIZA in Figure 6 and RASA in Figure 5, it can be seen that users tended to input longer sentences in RASA than in ELIZA. This advises positive dynamics in user engagement. All users of Round 2 considered RASA to be more empathetic than ELIZA.

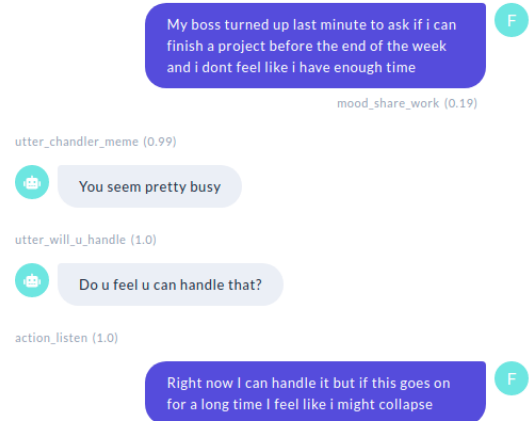


Figure 5: Excerpt from dialogue between RASA and participant

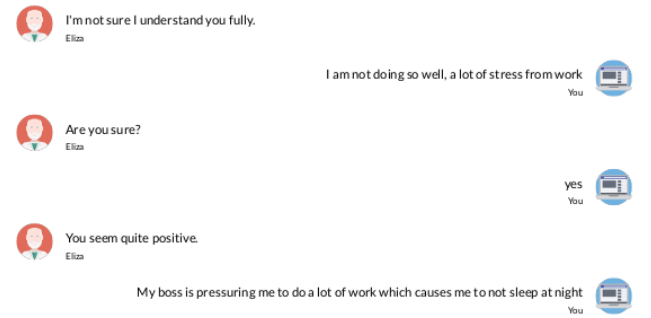


Figure 6: Excerpt from dialogue between ELIZA and participant

7. DISCUSSION

Due to the low number of participants, this research cannot make any definitive conclusions on empathetic factors of textual speech of CAs. However, it can provide insight for future much larger studies.

7.1 Research Question 1

The results from the experiment suggest that even though, majority of the participants found humor an appropriate form to express empathy, they were concerned about how appropriate memes and jokes could be in therapy. Some participants made a good point that older generations may find the jokes and memes offensive. All eight participants agreed on sympathy as being the key factor for

empathetic textual speech for RASA CA. The sympathy in our bot was expressed via "I feel you", "Your feelings are valid", "I had the same experience" etc. Participants mentioned that the validation and normalization of their feelings by the bot made the conversation more personal and open. However, regarding the AI therapist giving abstract/small advice ("Take it easy", "Try to take one evening off and watch some Netflix") were mixed. While the majority of participants considered it appropriate for therapists to provide some sort of advice, one participant expressed concerns on feeling insulted that bot minimizes their experience up to just taking things easy. Another user suggested that instead of giving abstract or straight advice, the bot-therapist could encourage the patient to make progress and, in case of stress at work, could provide some time-management techniques.

Due to a limited number of participants, research is unable to confidently answer on the first research question "Which empathy factors of textual speech during therapy sessions make people stay engaged with conversation?", however, we suggest that integrating sympathy and sympathetic utterances in the AI therapist could provide comfortable and trustful atmosphere for the client. Even though the humor was seen by the participants as a factor that may in the future be used to make the conversation more personal, they fairly noticed that lack of the dialogue context does not permit AI therapist to joke freely.

Regarding the use of slang in CA's responses, users had mixed opinions as well. In our opinion, the use of smiley faces ":" and ":" did not influence the empathy of the bot in a positive dynamic, because most of the participants did not notice them, and one of the participants was confused by the AI constantly switching between formal and informal language. We have concerns that slang may not be the optimal way to express artificial empathy, since not everyone from participants found that appropriate for therapy.

7.2 Research Question 2

Our empathetic bot was implemented on top of Machine Learning Natural Language Processing (NLP) using the RASA X framework. In our opinion, the NLP method is not advanced enough yet to help the bot to identify the context of the conversation, e.g. the main topic of the conversation (in case of the therapy the main issue/problem), respond based on the previous utterances and find appropriate linguistic and psychological attitude. The challenge of providing context is that context is a dynamic variable, so it may change over time. "Understanding" the context of the dialogue is an important part of every human-human interaction and based on the context people may or may not joke or use slang. That is why lack of context makes it almost impossible for the AI to appropriately insert humor responses.

Another approach to implementing therapy bot via Machine Learning is the use of Deep Learning. The concept of Deep Learning is based on Deep Neural Networks (DNN) and aims to simulate the human brain. DNN consists of plenty of hidden layers, each sending signals to the next layer to process the information. Deep Learning tries to mimic that in AI [17]. Even though this research did not concentrate as much on the deep learning part of machine learning, we see the potential for future therapy bots via that concept. Chen et al. [2] analyzed the potential of deep learning in dialogues systems. Researchers proposed that deep learning may serve as a good mechanism to integrate "longer term knowledge context and shorter term

dialogue context". In addition, Recurrent Neural Network (RNN) applied in Natural Language Generation (NLG) provide a controlling environment, which enables the bot to learn from unaligned data (which could be previous therapy sessions in case of the AI therapist) and thus avoid semantics repetition [20]. We think that avoidance of semantics repetitions will positively affect the relationship between the CA and the user since participants of Round 1 of the experiment expressed that repeated responses of the bot made the conversation less human-like.

The research struggles to provide the exact answer to the second research question "How to incorporate the factors of textual empathetic speech in a therapeutic conversational agent?" As concluded by the participants of both rounds the main limitation of the AI jokes is missing the context. Thus, we suppose the "understanding" of the knowledge and dialogue context will open a door for the bot to joke and express its sympathy, where appropriate, which may be achieved via deep learning in the future of Dialogue Systems.

7.3 Limitations

Two major limitations of therapy bots were outlined thanks to the participants. The first drawback of current CA therapists is the inability to properly evaluate the context of the dialogue. Another interesting point was made by one of the participants that bots are unable to remember and recognize the patient. If we are talking about long-term client-centered therapy, the patients share a lot of sensitive information, which real-life therapists always take into account for future sessions. However, it is currently impossible to do so, as bots are not able to "recall" previous conversations. This poses a problem and questions the future of AI therapists.

Regarding the conducted experiment, it should be mentioned that the choice to use roleplays instead of asking participants to share their real issues, was motivated by privacy and safety concerns. Since the research is not medical, we did not have any resources to help the participant during the experiment in case the person will be triggered by their experience. However, in our opinion, from the conversations with bots about true problems and issues of the participants, more defined and reliable results could have been derived.

8. CONCLUSION

Due to the low number of participants and small scope of the research, it is impossible to certainly answer the questions on empathetic factors of textual speech in the Dialogue System. However, the research provides insight on the empathy factors for larger-scope studies and outlines the limitations of the current state of the CA in mental healthcare.

Thus, the main problem of the CA remains the inability to process the context of the dialogues, hence reducing the options to express empathy, when needed. The current state of CAs does, on the other hand, allow them to express some sort of sympathy towards the patients. Normalization and validation of the user's feelings support the principle of nondirectiveness of client-centered therapy.

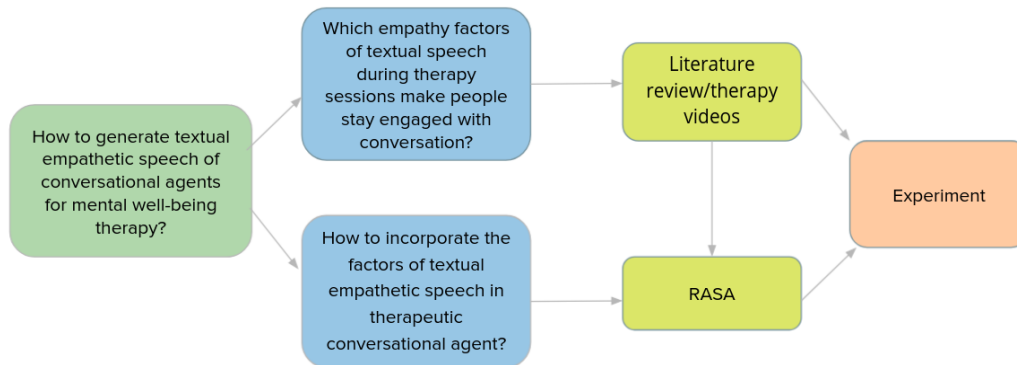
Unfortunately, the NLP approach for therapy bots in the scope of this research was not proven to be enough to exploit the CAs for client-centered therapy. However, the deep learning concept might provide solutions in the future for CAs "recognizing" a patient's previous sessions, as well as "understanding" the context of the conversation.

9. REFERENCES

- [1] A.G.Oettinger. Computational linguistics. *Communications of the ACM*, 9(1):36–45, January 1966.
- [2] Y.-N. Chen, A. Celikyilmaz, and D. Hakkani-Tur. Deep learning for dialogue systems. pages 25–31. Last accessed: 30.06.2021.
- [3] L. Clark, P. Doyle, C. Murad, N. Pantidi, D. Garaialde, J. Edwards, C. Munteanu, B. R. Cowan, O. Cooney, B. Spillane, E. Gilmartin, and V. Wade. What makes a good conversation? challenges in designing truly conversational agents. In *CHI 2019 Paper*. CHI, May 2019.
- [4] H. Cramer, J. Goddijn, B. Wielinga, and V. Evers. Effects of (in)accurate empathy and situational valence on attitudes towards robots. *IEEE*, 10:141–142, 2010.
- [5] K. K. Fitzpatrick¹, A. Darcy², and M. Vierhile. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): a randomized controlled trial. *JMIR Ment Health*, 4(2), June 2017.
- [6] H. Gaffney, W. Mansell, and S. Tai. Conversational agents in the treatment of mental health problems: Mixed-method systematic review. *JMIR Ment Health*, 6(10):329–341, March 2019.
- [7] D. T. Grande. Person-centered counseling role-play - coping with a work related stressor. <https://www.youtube.com/watch?v=zyIN61kQ6VY>. Last accessed: 20.06.2021.
- [8] B. Grant. Principled and instrumental nondirectiveness in person-centered and client-centered therapy. *Person-Centered Review*, 5:77–88, February 1990.
- [9] J. III, C. H. III, and R. Koppel. Clinical, legal, and ethical aspects of artificial intelligence-assisted conversational agents in health care. *JAMA*, 324(6):552–553, August 2020.
- [10] J. James, C. I. Watson, and B. MacDonald. Artificial empathy in social robots: An analysis of emotions in speech. *Proceedings of the 27th IEEE International Symposium on Robot and Human Interactive Communication*, 18:632–637, August 2018.
- [11] E. Kuyda. <https://replika.ai/>. Last accessed: 01.05.2021.
- [12] L. Laranjo, A. G. Dunn, H. L. Tong, A. B. Kocaballi, J. Chen, R. Bashir, D. Surian, B. Gallego, F. Magrabi, A. Y. Lau, and E. Coiera. Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association*, 25(9):1248–1258, July 2018.
- [13] E. Luger and A. Sellen. Like having a really bad pa”: The gulf between user expectation and experience of conversational agents. In *Living in Smart Environments*. CHI, ACM, May 2016.
- [14] A. S. Miner, A. Milstein, S. Schueller, R. Hegde, C. Mangurian, and E. Lino. Smartphone-based conversational agents and responses to questions about mental health, interpersonal violence, and physical health. *JAMA Intern Med.*, 176(5):619–625, November 2016.
- [15] A. I. of Professional Counsellors. Therapies in-action, 1, counselling therapies, session 1: Person centered therapy. https://search.alexanderstreet.com/search?ff%5B0%5D=video_series_facet%3ATherapies%20In-Action&sort_by=real_title_sort&sort_order=ASC. Last accessed: 20.06.2021.
- [16] A. P. Saygin, I. Cicekli, and V. Akman. Turing test: 50 years later. *Minds and Machines*, 10:463–518, 2001.
- [17] S. Siddique and J. C. L. Chow. Machine learning in healthcare communication. 1:220–239, 2021.
- [18] A. Todorović. <https://isturingtestpassed.github.io/>. Last accessed: 01.05.2021.
- [19] K. Weaver. Eliza. <https://github.com/keithweaver/eliza>. Last accessed: 20.06.2021.
- [20] T.-H. Wen, M. Gasic, N. Mrksic, P.-H. Su, D. Vandyke, and S. Young. Semantically conditioned lstm-based natural language generation for spoken dialogue systems. page 6, August 2015. Last accessed: 30.06.2021.
- [21] M. C. Witty. Significant aspects of client-centered therapy. pages 415–422.
- [22] M. C. Witty. Client-centered therapy. pages 35–50, 2007.

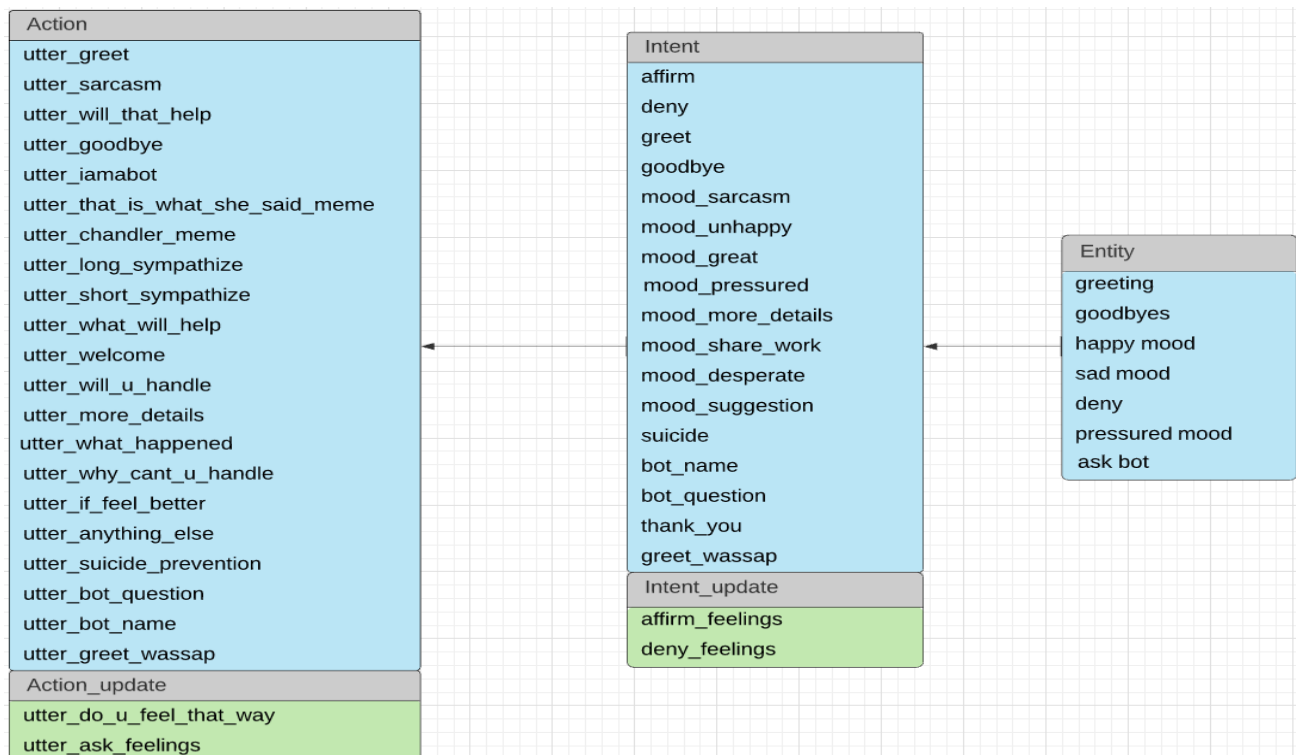
APPENDIX

A. OVERVIEW METHODOLOGY OF RESEARCH



Green color - main research question; blue color - sub question; yellow color - methodology per sub question; orange color - experiment based on both sub questions

B. RELATION BETWEEN ACTIONS, INTENTIONS AND ENTITIES OF RASA BOT



Blue color - original actions/intents/entities tested on Round 1 of the experiment; green color - additional actions/intents added after feedback from Round 1 and tested on Round 2

C. MEME RESPONSES OF RASA

action_listen (1.0)

What about you? Are you busy as well? B

bot_question["ask bot":"you"] (0.67)

utter_bot_question (1.0)

Knock knock
Who's there?
KGB



I am the one asking questions here :) Wassap with u?

action_listen (1.0)

Just vibin tbh, so much work but I am procrastinating B

mood_desperate (0.19)

action_listen (1.0)

I think I do, but at some point the pressure is too much 4

mood_desperate (0.28)

utter_that_is_what_she_said_meme (1.0)



Feeling overwhelmed is completely normal. A lot of people I know feel overwhelmed because they struggle with work/life balance. When I feel stressed, I usually take an evening completely off to Netflix & chill. I just try not to make it hard on myself

action_listen (1.0)

Yeah! I'm trying to get things done one by one, and hopefully I can make all my deadlines this week! B

mood_share_work (0.27)

utter_chandler_meme (1.0)



Oh man, could you BE any busier?

utter_long_sympathize (0.99)

I absolutely get u. At my previous job I did a lot of overtime, and not even because of my own fault. I had waaaaay too many responsibilities for one person and was constantly stressed.

utter_more_details (1.0)

Could u specify that?

action_listen (1.0)