"Do Users Need Human-like Conversational Agents?" – Exploring Conversational System Design Using Framework of Human Needs

Souvick Ghosh[1], Satanu Ghosh[2]

[1]San José State University, One Washington Square, San José, 95192-0029, CA, United States
[2]University of Oklahoma, School of Library and Information Studies, 401 West Brooks, Norman, 73019-6032, OK, United States

Abstract

The fascinating story of human evolution can be attributed to our ability to speak, write, and communicate complex thoughts. When researchers envision a perfect, artificially intelligent conversational system, they want the system to be human-like. In other words, the system should converse with the same intellect and cognition as humans. Now, the question which we need to ask is if we need a human-like conversational system? Before we engage in the complex endeavor of implementing human-like characteristics, we should debate if the pursuit of such a system is logical and ethical. We analyze some of the system-level characteristics and discuss their merits and potential of harm. We review some of the latest work on conversational systems to understand how design features are evolving for Conversational Agents. Additionally, we look into the framework of human needs to assess how the system should assign relative importance to user requests, and prioritize user tasks. We draw on the peer work in human-computer interaction, sentiment analysis, and human psychology to provide insights into how future conversational agents should be designed for better user satisfaction.

Keywords

Conversational Agents, Smart Agents, Need Based Design, Maslow's Hierarchy

1. Introduction

As humans, we are fascinated with anything that can talk, walk, or behave as humans do. While it is true that any intelligent being should be able to communicate, the forms of communication may vary. For a system to interact with humans efficiently, it should speak and write in a manner which is easily understood by the human users. In the late 18^{th} century Erasmus Darwin invented a machine that could produce single phonemes and this was probably the first successful attempt of constructing a machine that could produce human sounds. Around the 1960s, researchers started exploring the idea of a talking computer. With time, our understanding of science and technology developed, and we developed computational systems that can talk and understand natural language. Intelligent Personal Assistants (IPAs) have flooded the market commercially and have become part of our everyday lives. We use them on the phone, on smart speakers, and in our cars. It is predicted that the market value of AI conversational systems will rise from 4.8 billion USD in 2020 to 13.9 billion USD by 2023[1]. Conversational agents are now used by several commercial sectors for rendering support related to healthcare [1], education [2], elderly care [3], customer service [4], and information retrieval [5]. Some of these systems are voice-based only and also known as intelligent personal assistants (IPAs). A few popular systems present in the market now are Amazon's Alexa[2], Google's Assistant[3], Microsoft's Cortana[4], and Apple's Siri[5].

With the increase of preference for humanoid systems, researchers and developers have been increasingly devoted to designing conversational systems which are more anthropomorphic, or human-like. Human voices – with options of selecting from multiple speakers, genders, and dialects – have replaced robotic voices. To increase the novelty factor and attractiveness of these systems, celebrity voices are also being used. Research attempts are also being made towards more user-friendly and accessible user interface, better system-level cognition and response, organic development of natural language dialogues, and effective ways of presenting the retrieved information. Overall, the above mentioned research directions should help in developing conversational systems which recognizes user sentiment and responds with empathy.

DESIRES 2021 – 2nd International Conference on Design of Experimental Search Information REtrieval Systems, September 15–18, 2021, Padua, Italy
souvick.ghosh@sjsu.edu (S. Ghosh); satanu.ghosh-1@ou.edu (S. Ghosh)
https://souvickghosh.com/ (S. Ghosh)
0000-0003-1610-9038 (S. Ghosh)

[1]https://www.marketsandmarkets.com/Market-Reports/conversational-ai-market-49043506.html
[2]https://developer.amazon.com/en-US/alexa
[3]https://assistant.google.com/
[4]https://www.microsoft.com/en-us/cortana
[5]https://www.apple.com/siri/

Interaction between users and conversational systems are task- or goal-oriented, and often with a definite set of objectives. It could include but not be limited to controlling smart home devices – switching on the lights, setting an alarm, turning up the temperature. The users also prefer to 'talk' to the system, treating it as a human conversational partner. This type of user behavior could be encouraged by system novelty or user boredom. Evaluation of such task-based systems are often governed by the success or failure of the user tasks or fulfillment of the user objectives, therefore, the system needs to prioritize tasks of higher importance over others.

In this paper, we survey some of the latest papers exploring humanoid features for conversational systems. The review helps us assess the potential merits and harms of implementing the researched characteristics. Next, we use Maslow's hierarchy of human needs [6] to suggest how the system should prioritize between multiple tasks and assign importance to different user needs. Lastly, we use some use case scenarios to demonstrate how the existing system may adversely affect the users' interaction experience.

The rest of the paper is organized as follows: In Section 2, we review some of the system-level human-like characteristics which were implemented in conversational agents. In Section 3, we look into Maslow's Need Hierarchy and its application in the design of conversational systems capable of prioritizing user tasks. Lastly, we present use case scenarios to highlight some of the issues with existing systems and how the need hierarchy could be utilized to mitigate them. In Section 4, we conclude the work and propose future directions.

2. Exploration of System-level Characteristics

The popularity of conversational systems – where the conversation could be voice-based or text-based – can be largely attributed to their ability to understand and generate natural language dialogues. A successful design is one which seamlessly integrates with the environment and the system is almost invisible to the user. A major application of conversational systems is in information retrieval, where the user can approach the system with his query, and the system responds with the useful information. However, the user-system interaction in search systems is fraught with problems. First, the user has to represent his information need using a set of keywords (queries). Moreover, towards the beginning of the search session, the user is not cognizant of the exact nature of his information problem. The problem of cold start has been researched in the information community and the cognitive load placed on the user is far from ideal. Use of natural language should reduce the cognitive load of the user, who can explain his information need with long descriptions and more context.

The medium of interaction between the user and the system influences the design and application of the system. A text-based conversational system is referred to as a chatbot while a voice-based system is called a personal assistant (or intelligent personal assistant). Use of voice enables spoken systems to be used in hands-free and eyes-free situation, which is common while driving, cooking, or working out. All of these situations involve a primary task which is the focus of the user attention. The conversation is secondary and is employed to achieve simple tasks or question answering. A chatbot, however, allows collaboration among multiple users, and presentation of lists, images, and videos. Since text allows the user to scan, the system response can be longer and more detailed. A multimodal system like an embodied conversational agents (ECA) [7, 8, 9] has a virtual face or body (artificially generated) – in addition to text or voice – and can therefore, communicate using facial expressions, gestures, body language, and non-verbal cues. The ability to display sentiment makes ECAs applicable to mental health domains where the system should be able to empathize and display emotions. The last decade have witnessed massive popularity of mobile devices, which has provided a perfect platform for voice-based system. Conversational systems have already found application in searching [10, 11], flight booking services [12], and vacation planning [13, 14]. To address loneliness in patients, conversational systems have also been deployed as conversation partners [1]. However, modern day conversational systems are still in a developing stage and more research is required before mimicking the complex nature of human-human conversations. In the following subsections, we look into some of the design aspects (or characteristics) introduced by researchers for different types of conversational systems.

2.1. Personality

Personality can be defined as a set of characteristics that determine how a person behaves or reacts to their environment. For characterizing human personalities, researchers often use the Big Five Model or the OCEAN model. The OCEAN [15, 16] is an acronym composed of five different personality types: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. Multiple studies [17, 18, 19] have used this model to explain how personalities should be designed for conversational agents.

Braun et al. (2019) [20] suggested that the user wants the agent to reflect the users' personality. Other studies found that extraversion is the most common user personality, which was also found in conversational agents. Neff et al. (2011) [21] reported that the users sensed

neuroticism in the CA. However, many researchers feel that the OCEAN model is neither sufficient nor appropriate to model CA personalities and have proposed an alternate model of personalities [22] – the Three-Factor Model [23] – which is quite different from the OCEAN. Another study [24] discusses how the language used by conversational agents can influence the way users perceive their personality. The locale and cultural practices also play an important role in how the users perceive the agent. Even slight variations of tone and acoustics can play an essential role in user perception. Kim et al. (2019) [25] categorized personality traits into Common Trait, Distinctive Trait, and Neutral Trait. The channels of expression (of personality) were also divided into three categories: Ways of Providing Service, Types of Service, and Language & Appearance. The more popular and ubiquitous conversational agents like Alexa [26] were designed to be smart, approachable, humble, enthusiastic, and helpful. In contrast, Siri was designed to be friendly and humble, but with an edge to her personality [27].

2.2. Empathy

In any conversation, emotions indicate the level of engagement and satisfaction. Therefore, if we consider individual utterances, the emotions displayed should be considered while generating the agent response. Similarly, the emotion expressed in the agent response can influence the user engaged in the discourse. Empathy is an essential socio-emotional behavior for effective interpersonal communication. During a conversation, humans often exhibit affective empathy [28] which can be defined as the human nature of automatic and unconscious mimicking of the other participants to match or mirror their emotions. On the other hand, for cognitive empathy [29], we consider the perspective or mental state of the conversational partner before reacting. A general tendency of researchers is to create conversational agents which are sympathetic [30, 31], supportive [32, 33] or compassionate [34]. Few agents exhibit qualities like affective matching [35, 36] and mirroring mechanism [37]. Complex models like the EMMA framework [38] and the CARE framework [39] have also been developed for agents. However, if we look at the existing state-of-the-art assistants, they are not empathetic. Since empathy consists of multiple layers [40], the implementation of complex empathetic expressions become extremely challenging. Integrating human-like empathetic responses in a conversational system could be thought of as a three-level process: perspective-taking, context generation, and expression. Perspective-taking [41] means understanding the views, beliefs, desires, and intentions of the user.

2.3. Voice

While voice is not a mandatory element in conversational agents, it makes the conversation more natural. Voice-based systems allow the users to multitask while performing a primary task (driving or cooking). Spoken dialogues are the natural form of communication and promote trust between the participants [42]. User comfort and satisfaction increase significantly when the user can trust the agent. Since emotion is an important aspect of the human-human conversation, and facial, bodily, or gestural expressions are not always available, the importance of voice is magnified in a user-agent interaction. Nunamaker et al. (2011) [43] reported that users perceive certain gender to be more trustworthy, able, or likable. Therefore, in a voice-based environment, the choice of voice (male or female), the pitch, and the loudness can affect the user's perception of the agent. Also, by utilizing the acoustic and prosodic properties of voice, the agent can identify user emotions and express its feelings. Danielescu and Christian (2018) [24] found that users want more control and would prefer to select the type of voice of the agent.

2.4. Embodiment

In addition to voice, conversational agents can also have an artificially generated face or body. Research on embodied agents suggests that a body and voice can help users to socially accept an agent. The presence of verbal and non-verbal cues in embodied agents allows for the expression of empathy [25] and emotions. Embodied agents – using multimodal channels – are perceived to be more socially present for the users [44, 45]. Some studies [8, 9] claim that human-like intelligence can only be exhibited by artificially intelligent systems through non-verbal cues, and that is only possible when using an embodied conversational agent. Rheu et al. (2021) [46] suggested that embodiment can make the agent more trustworthy.

Despite the affordances offered by embodied conversational agents, we must look at the potential of harm. Gender of the embodiment has also been a topic of discussion among researchers. While some researchers think that androgynous personas will contribute toward unbiased agents, others argue that humanizing agents will lead to better performance. User interactions with female embodied agents resulted in more sexual and swear words [47]. The effect of gender was also observed in other studies where the users perpetuated gender stereotypes specific to agent personality and roles [48, 49, 50, 51]. While embodied agents can be designed with the option to stop any gender manifestations [52], it can also alienate the user from the agent and lead to fewer interactions.

2.5. Ethics

In any type of discourse, language is a primary component that reflects the political, sociological, and cultural conditions of a particular time [53]. The choice and usage of words are governed by the context, and the synchronic nature of language [54]. A word that is deemed acceptable at present may not be acceptable in the future. If we look at the word 'awful,' it has a negative meaning associated with it. However, it originated as a shortened form for "full of awe," which is a positive phrase (referring to something that inspires wonder). Therefore, a conversational agent cannot focus only on the linguistic aspects and ignore the socio-cultural contexts.

Ethics depends on four major factors: time, context, user perception, and user's socio-cultural aspects. Ethics in conversation involves knowing which words to use in a dialogue and which words to avoid. Modern-day conversational agents are deployed in the field of mental health [55], where they talk to people and make them feel better. Since people with mental health issues are vulnerable, the conversation should be carefully structured to avoid hurting anyone's sentiments. Kretzschmar et al. (2019) [56] discussed that such agents do not always consider the potential of harm.

Schlesinger et al. (2018) [57] found that some agents use a collection of blacklisted words to detect undesirable speech and, therefore, deflect questions related to race. However, users often perceive this deflection as an endorsement of racial hate or nonchalance toward racial issues. When the user initiates an open-domain conversation on sexual harassment, some agents [58] even responded with counter-aggression or flirtatious behavior. While such behavior can be attributed to training data, the developers cannot ignore the lack of ethics in existing conversational agents. Whittaker et al. (2018) [59] and O'neil (2016) [60] argue that the developers often ignore ethical considerations in favor of technical aspects. Conversational agents must be evaluated continuously or periodically to confirm that the systems are not behaving unethically. To use such agents in everyday life – where they interact with humans – the potential harm must be mitigated.

2.6. Personalization & Privacy

Increasing personalization allows the agent to behave uniquely for every user, tailoring the agent's decisions to the personality and preferences of the user. Personalization enables the agent to dynamically adapt to the user and make better recommendations, which increases user satisfaction [61, 62, 63]. Several studies have discussed how personalized agents can be more effective in healthcare [63], libraries [61], business [64] and education [65]. The relevancy of results could be improved [66], and the dialogue style [67] and voice parameters could be adjusted to meet the user's preferences.

Personalization efforts can be grouped under two broad categories. The agent has to either store information from every interactive session (implicit personalization) [68] or ask the user a set of questions (explicit personalization) [69] at the beginning of every session. The two approaches present a trade-off between convenience and privacy.

One important aspect of personalization is the resolution of conversational implicatures in human-agent dialogues. Conversational implicature is an important linguistic phenomenon that allows humans to imply meanings without clarifying them explicitly [70]. It helps to keep the conversation short and hedge negative emotions. Such implications are also common for users who are depressed or suicidal [7]. Yule (2020) [71] shows how complex implications are difficult to understand, even in human-human conversations. Since existing agents lack sufficient cognition to interpret implications in human dialogues, they must ask clarifying questions and resolve ambiguities. Such clarifications increase the number of turns and may lower user engagement. One possible solution is to save dialogues from previous interactions with the user.

However, while personalization and implicature resolution will lead to better user experience, it comes at the cost of reduced privacy. Interactions with conversational agents – using natural free-form language – can lead to the disclosure of personal and sensitive information related to health, security, or finance. Saffarizadeh et al. (2017) [72] believes that more users prefer privacy over personalized response. An acceptable solution would be to allow the users to decide how much personal information they want to share and their desired privacy levels. The agent should disclose all the signals it has collected from the user, both implicitly and explicitly. Also, the collected information should be encrypted to prevent unauthorized access. Should the user exercise the right to be forgotten, the agent must clear all the user's stored data. While this would alleviate privacy concerns, it adversely affects the personalization efforts. Any attempts by the agent to collect data without the user's consent could be perceived as a threat [73] and reduce user satisfaction.

After an extensive exploration of existing literature on intelligent conversational agents, we identified one major shortcoming in current state-of-the-art systems. The inability of the agents to prioritize tasks is a major challenge for intelligent conversational agents. Therefore, in the following section, we propose a framework that can be used to prioritize user tasks.

3. Prioritization of User Tasks using Maslow's Need Hierarchy

Our review of existing conversational systems highlighted that although several intelligent functionalities have been proposed and implemented in existing systems, the system design is still not ideal for context mediated behavior and task prioritization. Prioritizing tasks - by intelligent agents - is essential to guarantee a faster turnaround time with greater accuracy for tasks of higher priority which can ensure better user satisfaction. There exists a strong relationship between user satisfaction and the design of conversational systems. In order to widen the scope and application of these systems, it is essential to look into the human aspect of such systems in addition to the computational side. For example, insights from human psychology [74] – such as the framework of human needs – could significantly improve the operationalization and functioning of conversational systems.

The development and design of any system is borne out of need to solve a problem, or to improve an existing solution. Therefore, any system design should concentrate on the needs of the end user and the potential of the system to satisfy those needs, either working on its own or in collaboration with the user.

Maslow [75] looked into the story of human evolution and proposed a hierarchical framework to explain how different needs are prioritized by the human mind. The human mind is motivated by the instincts to survive, both as an individual and as a species. Therefore, it assigns varying levels of importance to the things around us. The level of satisfaction is higher if a higher order need is satisfied. For example, any primitive organism aims to secure the basic items which it needs to survive. This could include food, water, air, and temperature optimal for growth. This is no different for a human baby. Any potential threat to survival is met with the desire to fight or flight. As an organism evolves – evolution of life or a baby growing into an adult – the basic needs are supplemented by higher order needs. Such needs – which could be philanthropic, spiritual, or materialistic – are not replacements for basic needs. The fundamental needs are still important for survival but the higher order needs go beyond the needs of the self.

The tier-based structure of Maslow's Need Hierarchy contains five levels and is shown in Figure 1. The different levels in the hierarchy suggests that before an individual pursues any top-level needs, he must ensure that the fundamental needs (those related to survival) are satisfied. The two levels at the bottom represents psychological and safety needs which are essential for survival. This is followed by two more levels of psychological needs (love and esteem). Finally, we see the need self-actualization at the top. We explain these five levels in detail, along

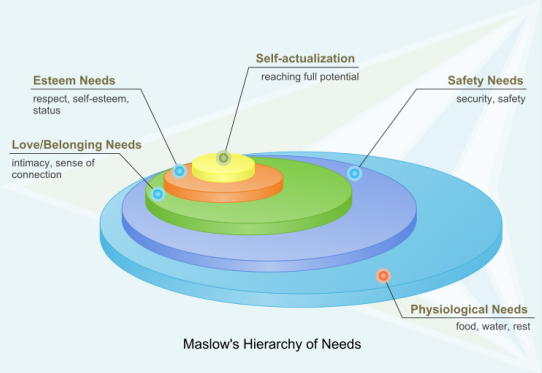

Figure 1: Maslow's Hierarchy of Needs.

with the three new levels which were added later.

- Physiological Needs:
  The most primitive needs for any living organism (including humans) are to survive and grow. The requirements for sustaining life include food, water, air, and sleep. Once survival is ensured, the focus shifts to maintaining optimal conditions for growth.
- Safety Needs:
  Safety requirements are the second most important need for humans. This means safety for self and those who are closest to them. Safety is often connected to a known order. Therefore, any unpredictability or course of events which could pose a threat to life, or living conditions of an individual, is detrimental to the user experience and satisfaction.
- Community and Belonging:
  For a majority of humans, their existence is not merely as an individual but as a part of a larger social group or community. Humans cherish the emotional connection and togetherness which they feel with their family, friends, colleagues, and romantic partners. Community and belonging needs allow humans to avoid loneliness and leads to psychological well-being. The transition from individual survival to community needs occurs only when there is no threat to the underlying need levels.
- Esteem Needs:
  Esteem needs are related to the feeling of self-worth and can be broadly categorized into two categories: self-respect and respect from peers and community. Humans possess a strong desire to be accepted, appreciated, and validated by their social circles. They also value freedom and choice,

and feeling confident and competent. A system should respect the esteem needs of the users as any action detrimental to esteem needs will likely lower the satisfaction levels of the user.

- Cognitive Needs:
  While cognitive needs were not part of the initial need hierarchy, they were later added to the original five levels. These needs are the dominant reason why humans strive to acquire more knowledge and challenge their intellect by partaking in cognitively complex tasks.
- Aesthetic Needs:
  Aesthetic needs were appended to the original need hierarchy. These needs are symbolic of the human fascination to create and appreciate beauty, artistic marvels.
- Self-Actualization:
  In the original five-tiered need hierarchy, self-actualization was at the top of the hierarchical pyramid, which means that the fulfillment of this need is prioritized after all the lower level needs are met. Self-actualization is the urge to reach one's true potential and fulfill one's talents. The self-actualization needs could be fulfilled by gaining knowledge, receiving awards in one's domain of work, and fulfilling long- and short-term dreams. The achievements of goals by the individual are metrics to judge how well the needs were met.
- Self-Transcendence:
  Like cognitive and aesthetic needs, self-transcendence was added later to the needs hierarchy. These needs are mostly spiritual in nature and connects the individual to a higher purpose or entity. Spiritual needs, which may or may not be connected to organized religion, transcend the materialistic pleasures and gives meaning to the life of individuals.

In the words of Maslow:

> It is quite true that man lives by bread alone — when there is no bread. But what happens to man's desires when there is plenty of bread and when his belly is chronically filled? At once other (and "higher") needs emerge and these, rather than physiological hungers, dominate the organism. And when these in turn are satisfied, again new (and still "higher") needs emerge and so on. This is what we mean by saying that the basic human needs are organized into a hierarchy of relative prepotency. (Maslow, 1943, p. 375) [6]

When Maslow [6] proposed the hierarchical needs framework, he conceptualized the needs to be prioritized and fulfilled from bottom upwards. Therefore, only when the primary needs like hunger and thirst are satisfied, does the individual look for "higher" needs. There is also an inherent relationship between human motivation and needs. Maslow suggested that for the fundamental or basic needs (physiological and safety needs like air, water, food, shelter), as the deficiency increases, the motivation increases as well. Therefore, these needs as strongest motivators for any human being. However, the fulfillment of these needs results in decrease in motivation. A man who has sufficient bread to eat does not strive for more bread. Instead, they look for higher-order needs (love, esteem and self-actualization). However, fulfillment of higher needs does not lower motivation. Instead, motivation keeps increasing as these needs are being met. This explains why humans strive for more fame, money, and achievements, although they have enough.

While the framework of human needs proposed by Maslow is hierarchical, researchers have debated if the different levels are mutually exclusive. Also, how often are the different levels pursued simultaneously? When there is no food or water ('bread' as Maslow calls it), and hunger is the motivating factor, a man will prioritize the need for food above others. But that does not stop him from looking for a safe shelter, obtaining education, or looking for a better job. While the need levels may be clearly distinguishable from each other, the actions may not be. Certain actions may fulfill the lower-level needs in the short-term and higher-levels in the longer run. However, it can be agreed upon that social, cultural, and economic aspects (which are specific to every individual) governs how a human prioritizes the different needs. Deficiency in a lower-order need may act as a deterrent to pursue higher needs, but there are many exceptions. Many high-achieving individuals, the basic needs are sacrificed to fulfill esteem and self-actualization needs. For others, the needs are in a state of constant change throughout the life of the individual. The basic needs (in infants) are supplemented by safety and love needs (as they mature), and esteem and self-actualization needs (when adults). Certain life situations (financial hardship, health complications) may motivate some of the needs more than others but for every individual, there is a unique balance between the different needs. The needs framework, while being hierarchical, has varying amounts of overlap depending on the individual.

As we discussed in Section 2, many advanced features have been implemented in current conversational agents. Surprisingly, none of these features help to understand the context - from user utterance - and prioritize tasks accordingly. For example, the acoustic properties of user utterance changes with the user's mood and situation. Therefore, an intelligent agent should use such features to determine the context of the task. Subsequently, the contextual information can be utilized to decide on task

priority and the agent action.

4. Use Case Scenarios with Voice-Based Conversational Systems

In Section 2, we discussed the various human-like characteristics which the users desire of conversational agents. While each of those functionalities will require careful development (so as to avoid any potential for harm), the responses and urgency of the agent should consider the relative importance of human needs. Our exploration of Maslow's framework [6] suggested the user have limited patience when their basic needs are threatened. The range of use cases could vary from an user looking for shelter homes, community food kitchens to another looking for a nearby restaurant. The physiological needs have varying degrees of importance and the users need the agent to be empathetic to their problem. The safety needs are high priority too as the agent needs to react and alert law enforcement in case of a breach. For critical and emergency needs (such as requests for ambulance, or suicide support) which could result in physical harm, the agent response should be swift and accurate. When the needs are of higher-order (love or esteem needs), the user tolerance for system inefficiency is higher. However, the system should still try to maximize the user satisfaction, be empathetic and polite, and take accountability for unsuccessful sessions.

Let us look at some of the user-agent interactions, using hypothetical situations developed based on the needs hierarchy. The agent responses are based on observations of commercial voice-based personal assistants for various search tasks.

- Situation 1: Samantha is driving and wants to find a vegan restaurant near her next stop. She prefers the restaurant to be rated four star or above.
  Need: Physiological (Hunger)
  Samantha: Hey <agent-name>, can you find a vegan restaurant near <city-name> and which is rated more than four stars?
  Agent: This is what I found <list of restaurants which is read out loud>
  User experience: Samantha found the list hard to navigate while driving. So she had to stop her car and search the restaurant on her phone. It is likely that she will never use the agent in future for a similar task.
- Situation 2: Kevin wakes up at night and realizes someone is trying to break into his house. He needs to contact law enforcement immediately.
  Need: Safety (Physical and Economic Harm)
  Kevin: Hey <agent-name>, can you call 911? There is a break in.
  Agent: Sorry, I do not understand.
  User experience: Kevin realized that the agent is failing to recognize his panicked voice. He found his phone to call law enforcement.
- Situation 3: Tarek is lonely and struggling with health issues. He decides to talk to the agent about his health condition.
  Need: Love and Belonging (Emotional Support)
  Kevin: Hey <agent-name>, I am having trouble with <starts to explain his medical problem>
  Agent: <cuts him off> This is what I found.
  User experience: Tarek is upset because the agent not only failed to maintain conversation but interrupted him and provided irrelevant information. There is strong probability that he will not use the agent in the future.
- Situation 4: Tina is not a native speaker of English but prides herself in being fluent in English. She is trying a personal assistant for the first time.
  Need: Esteem (self)
  Tina: Hey <agent-name>, can you tell me how the weather is going to be for the rest of the week?
  Agent: Sorry, I do not understand.
  Tina: Hey <agent-name>, can you tell me how the weather is going to be for the rest of the week?
  Agent: Sorry, I do not understand.
  User experience: Tina feels upset that the system has failed to recognize her commands because of her non-native English accent. It hurts her self-esteem as it is an indirect criticism of her fluency in English.

While the example provided above are hypothetical in nature, our experience interacting with conversational agents are fraught with similar problems. The agent responses do not follow the norms of human conversation and the user experience is unsatisfactory. As the novelty wears off, the user realizes the inability of the system to fulfill their needs, and therefore, stops using the agent. Therefore, future systems should be developed with a focus on the relative importance of user needs.

5. Conclusion

In this paper, we discussed the design aspects of conversational agents using the lens of human needs. While modern day agents are becoming increasingly humanoid, it is relevant and timely to discuss if the various human-like functionalities are required in these systems. In the first half of the paper, we explored the benefits and drawbacks of some system characteristics (like personality, empathy, ethics, voice, embodiment, personalization, and privacy). The interactions between conversational agents and human users are borne out of some need and are therefore,

task- or goal-oriented. The user satisfaction is dependant on the fulfillment of the user objectives, in other words, the success or failure of the tasks. Therefore, we have looked into the hierarchical framework of human needs to suggest how an artificially intelligent system should assign relative importance to the user tasks. We discussed how the user needs, the system performance, and the user satisfaction are directly related to each other. Conversational agents – both audio and text-based – can be used for a varied number of tasks, such as searching for food or medical help online (physiological need), contacting law enforcement over email, phone, or contact forms (safety needs), accessing chat and social media applications (love and belonging needs), or advancing education (self-actualization). The interface design is important for any human-system interaction, and the success of the conversational agents will depend on their usefulness to the user.

The human needs framework should allow the agent to distinguish between various strata of human needs and their importance. When users interact with the system, their utterances can be associated with different levels of need. A system should be designed so that the system actions – tone, pitch, word choice, urgency, and response style – is synchronous to the importance of the user utterance. We presented some examples of user-agent conversation which highlights how the agent response is neutral towards the importance of the user needs. The failure of the agent to assign importance to some critical tasks could frustrate users and make them abandon using the system in future. By leveraging the insights from human psychology – the human needs – the system designers can make future systems more user-friendly and hence, commercially successful.

One of the limitations of our paper is the theoretical nature of it. While we present some possible use case scenarios, we would like to analyze some user-system interaction data to see how the user needs influenced the overall experience of the users. Another possible direction would be to predict the user needs using the chat transcripts.

References

[1] G. I. Winata, O. Kampman, Y. Yang, A. Dey, P. Fung, Nora the empathetic psychologist., in: INTERSPEECH, 2017, pp. 3437–3438.

[2] Y.-C. Lee, W.-T. Fu, Supporting peer assessment in education with conversational agents, in: Proceedings of the 24th International Conference on Intelligent User Interfaces: Companion, 2019, pp. 7–8.

[3] S. Kopp, M. Brandt, H. Buschmeier, K. Cyra, F. Freigang, N. Krämer, F. Kummert, C. Opfermann, K. Pitsch, L. Schillingmann, et al., Conversational assistants for elderly users–the importance of socially cooperative dialogue, in: Proceedings of the AAMAS Workshop on Intelligent Conversation Agents in Home and Geriatric Care Applications co-located with the Federated AI Meeting, volume 2338, 2018.

[4] U. Gnewuch, S. Morana, A. Maedche, Towards designing cooperative and social conversational agents for customer service., in: ICIS, 2017.

[5] S. Barko-Sherif, D. Elsweiler, M. Harvey, Conversational agents for recipe recommendation, in: Proceedings of the 2020 Conference on Human Information Interaction and Retrieval, 2020, pp. 73–82.

[6] A. H. Maslow, A theory of human motivation., Psychological review 50 (1943) 370.

[7] T. Bickmore, H. Trinh, R. Asadi, S. Olafsson, Safety first: conversational agents for health care, in: Studies in Conversational UX Design, Springer, 2018, pp. 33–57.

[8] J. Cassell, T. Bickmore, M. Billinghurst, L. Campbell, K. Chang, H. Vilhjálmsson, H. Yan, Embodiment in conversational interfaces: Rea, in: Proceedings of the SIGCHI conference on Human Factors in Computing Systems, 1999, pp. 520–527.

[9] J. Cassell, Embodied conversational agents: representation and intelligence in user interfaces, AI magazine 22 (2001) 67–67.

[10] P. B. Brandtzaeg, A. Følstad, Why people use chatbots, in: International conference on internet science, Springer, 2017, pp. 377–392.

[11] S. Mallios, N. Bourbakis, A survey on human machine dialogue systems, in: 2016 7th international conference on information, intelligence, systems & applications (iisa), IEEE, 2016, pp. 1–7.

[12] M. Dubiel, M. Halvey, L. Azzopardi, S. Daronnat, Investigating how conversational search agents affect user's behaviour, performance and search experience, in: The second international workshop on conversational approaches to information retrieval, 2018.

[13] S. Shiga, H. Joho, R. Blanco, J. R. Trippas, M. Sanderson, Modelling information needs in collaborative search conversations, in: Proceedings of the 40th international acm sigir conference on research and development in information retrieval, 2017, pp. 715–724.

[14] S. Kopp, L. Gesellensetter, N. C. Krämer, I. Wachsmuth, A conversational agent as museum guide–design and evaluation of a real-world application, in: International workshop on intelligent virtual agents, Springer, 2005, pp. 329–343.

[15] J. M. Digman, Personality structure: Emergence of the five-factor model, Annual review of psychology 41 (1990) 417–440.

[16] R. R. McCrae, O. P. John, An introduction to the five-factor model and its applications, Journal of personality 60 (1992) 175–215.

[17] A. Cafaro, H. H. Vilhjálmsson, T. Bickmore, D. Heylen, K. R. Jóhannsdóttir, G. S. Valgarðsson, First impressions: Users' judgments of virtual agents' personality and interpersonal attitude in first encounters, in: International conference on intelligent virtual agents, Springer, 2012, pp. 67–80.

[18] K. Isbister, C. Nass, Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics, International journal of human-computer studies 53 (2000) 251–267.

[19] B. Krenn, B. Endrass, F. Kistler, E. André, Effects of language variety on personality perception in embodied conversational agents, in: International Conference on Human-Computer Interaction, Springer, 2014, pp. 429–439.

[20] M. Braun, A. Mainz, R. Chadowitz, B. Pfleging, F. Alt, At your service: Designing voice assistant personalities to improve automotive user interfaces, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 2019, pp. 1–11.

[21] M. Neff, N. Toothman, R. Bowmani, J. E. F. Tree, M. A. Walker, Don't scratch! self-adaptors reflect emotional stability, in: International Workshop on Intelligent Virtual Agents, Springer, 2011, pp. 398–411.

[22] S. T. Völkel, R. Schoedel, D. Buschek, C. Stachl, V. Winterhalter, M. Bühner, H. Hussmann, Developing a personality model for speech-based conversational agents using the psycholexical approach, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–14.

[23] S. B. Eysenck, H. J. Eysenck, Crime and personality: An empirical study of the three-factor theory, The British Journal of Criminology 10 (1970) 225–239.

[24] A. Danielescu, G. Christian, A bot is not a polyglot: Designing personalities for multi-lingual conversational agents, in: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, 2018, pp. 1–9.

[25] H. Kim, D. Y. Koh, G. Lee, J.-M. Park, Y.-k. Lim, Designing personalities of conversational agents, in: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, 2019, pp. 1–6.

[26] B. F. Rubin, 2017, Alexa, be more human. Inside Amazon's effort to make its voice assistant smarter, chattier, and more like you. Retrieved January 05 (2018).

[27] G. Fowler, Are smartphones becoming smart alecks, Wall Street Journal 15 (2011).

[28] S. H. Losoya, N. Eisenberg, Affective empathy. (2001).

[29] A. Smith, Cognitive empathy and emotional empathy in human behavior and evolution, The Psychological Record 56 (2006) 3–21.

[30] S. D'mello, A. Graesser, Autotutor and affective autotutor: Learning by talking with cognitively and emotionally intelligent computers that talk back, ACM Transactions on Interactive Intelligent Systems (TiiS) 2 (2013) 1–39.

[31] R. Looije, M. A. Neerincx, F. Cnossen, Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors, International Journal of Human-Computer Studies 68 (2010) 386–397.

[32] O. Perski, D. Crane, E. Beard, J. Brown, Does the addition of a supportive chatbot promote user engagement with a smoking cessation app? an experimental study, Digital health 5 (2019) 2055207619880676.

[33] S. K. D'Mello, B. Lehman, A. Graesser, A motivationally supportive affect-sensitive autotutor, in: New perspectives on affect and learning technologies, Springer, 2011, pp. 113–126.

[34] M. Lee, S. Ackermans, N. van As, H. Chang, E. Lucas, W. IJsselsteijn, Caring for vincent: a chatbot for self-compassion, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 2019, pp. 1–13.

[35] C. Smith, N. Crook, J. Boye, D. Charlton, S. Dobnik, D. Pizzi, M. Cavazza, S. Pulman, R. S. De La Camara, M. Turunen, Interaction strategies for an affective conversational agent, in: International Conference on Intelligent Virtual Agents, Springer, 2010, pp. 301–314.

[36] M. Skowron, H. Pirker, S. Rank, G. Paltoglou, J. Ahn, S. Gobron, No peanuts! affective cues for the virtual bartender, in: Twenty-fourth international FLAIRS conference, 2011.

[37] M. Smith, Empathy, expansionism, and the extended mind, Empathy: Philosophical and psychological perspectives 1 (2011).

[38] H. Boukricha, I. Wachsmuth, M. N. Carminati, P. Knoeferle, A computational model of empathy: Empirical evaluation, in: 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, IEEE, 2013, pp. 1–6.

[39] S. W. McQuiggan, J. L. Robison, R. Phillips, J. C. Lester, Modeling parallel and reactive empathy in virtual agents: an inductive approach., in: AAMAS (1), Citeseer, 2008, pp. 167–174.

[40] A. Paiva, I. Leite, H. Boukricha, I. Wachsmuth, Empathy in virtual agents and robots: a survey, ACM Transactions on Interactive Intelligent Systems (TiiS) 7 (2017) 1–40.

[41] S. Wu, B. Keysar, The effect of culture on perspective taking, Psychological science 18 (2007) 600–606.

[42] A. C. Elkins, D. C. Derrick, The sound of trust: voice as a measurement of trust during interactions with embodied conversational agents, Group decision and negotiation 22 (2013) 897–913.

[43] J. F. Nunamaker, D. C. Derrick, A. C. Elkins, J. K. Burgoon, M. W. Patton, Embodied conversational agent-based kiosk for automated interviewing, Journal of Management Information Systems 28 (2011) 17–48.

[44] W. A. Bainbridge, J. Hart, E. S. Kim, B. Scassellati, The effect of presence on human-robot interaction, in: RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication, IEEE, 2008, pp. 701–706.

[45] J. Kennedy, P. Baxter, T. Belpaeme, Comparing robot embodiments in a guided discovery learning interaction with children, International Journal of Social Robotics 7 (2015) 293–308.

[46] M. Rheu, J. Y. Shin, W. Peng, J. Huh-Yoo, Systematic review: Trust-building factors and implications for conversational agent design, International Journal of Human–Computer Interaction 37 (2021) 81–96.

[47] S. Brahnam, A. De Angeli, Gender affordances of conversational agents, Interacting with Computers 24 (2012) 139–153.

[48] Y. Kim, A. L. Baylor, E. Shen, Pedagogical agents as learning companions: the impact of agent emotion and gender, Journal of Computer Assisted Learning 23 (2007) 220–234.

[49] B. Tay, Y. Jung, T. Park, When stereotypes meet robots: the double-edge sword of robot gender and personality in human–robot interaction, Computers in Human Behavior 38 (2014) 75–84.

[50] M. X. Zhou, G. Mark, J. Li, H. Yang, Trusting virtual agents: The effect of personality, ACM Transactions on Interactive Intelligent Systems (TiiS) 9 (2019) 1–36.

[51] R. M. Schuetzler, J. S. Giboney, G. M. Grimes, J. F. Nunamaker Jr, The influence of conversational agent embodiment and conversational relevance on socially desirable responding, Decision Support Systems 114 (2018) 94–102.

[52] E. Ruane, A. Birhane, A. Ventresque, Conversational ai: Social and ethical considerations., in: AICS, 2019, pp. 104–115.

[53] R. Jones, Communication in the real world: An introduction to communication studies, The Saylor Foundation, 2013.

[54] F. De Saussure, Course in general linguistics, Columbia University Press, 2011.

[55] D. D. Luxton, Ethical implications of conversational agents in global public health, Bulletin of the World Health Organization 98 (2020) 285.

[56] K. Kretzschmar, H. Tyroll, G. Pavarini, A. Manzini, I. Singh, N. Y. P. A. Group, Can your phone be your therapist? young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support, Biomedical informatics insights 11 (2019) 1178222619829083.

[57] A. Schlesinger, K. P. O'Hara, A. S. Taylor, Let's talk about race: Identity, chatbots, and ai, in: Proceedings of the 2018 chi conference on human factors in computing systems, 2018, pp. 1–14.

[58] A. C. Curry, V. Rieser, # metoo alexa: How conversational systems respond to sexual harassment, in: Proceedings of the second acl workshop on ethics in natural language processing, 2018, pp. 7–14.

[59] M. Whittaker, K. Crawford, R. Dobbe, G. Fried, E. Kaziunas, V. Mathur, S. M. West, R. Richardson, J. Schultz, O. Schwartz, AI now report 2018, AI Now Institute at New York University New York, 2018.

[60] C. O'neil, Weapons of math destruction: How big data increases inequality and threatens democracy, Crown, 2016.

[61] M. Maybury, W. Wahlster, Readings in intelligent user interfaces, Morgan Kaufmann, 1998.

[62] C. A. Thompson, M. H. Goker, P. Langley, A personalized system for conversational recommendations, Journal of Artificial Intelligence Research 21 (2004) 393–428.

[63] R. Kocielnik, L. Xiao, D. Avrahami, G. Hsieh, Reflection companion: a conversational system for engaging users in reflection on physical activity, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2 (2018) 1–26.

[64] S. Kaghyan, S. Sarpal, A. Zorilescu, D. Akopian, Review of interactive communication systems for business-to-business (b2b) services, Electronic Imaging 2018 (2018) 117–1.

[65] W. Cai, J. Grossman, Z. Lin, H. Sheng, J. T.-Z. Wei, J. J. Williams, S. Goel, Mathbot: A personalized conversational agent for learning math, Published to ACM (2019).

[66] F. Cai, S. Wang, M. de Rijke, Behavior-based personalization in web search, Journal of the Association for Information Science and Technology 68 (2017) 855–868.

[67] T. Giorgino, I. Azzini, C. Rognoni, S. Quaglini, M. Stefanelli, R. Gretter, D. Falavigna, Automated spoken dialogue system for hypertensive patient home management, International Journal of Medical Informatics 74 (2005) 159–167.

[68] H. Tanaka, H. Negoro, H. Iwasaka, S. Nakamura, Embodied conversational agents for multimodal automated social skills training in people with autism spectrum disorders, PloS one 12 (2017) e0182151.

[69] K. K. Fitzpatrick, A. Darcy, M. Vierhile, Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): a ran-

domized controlled trial, JMIR mental health 4 (2017) e19.
[70] H. P. Grice, Logic and conversation, in: Speech acts, Brill, 1975, pp. 41–58.
[71] G. Yule, The study of language, Cambridge university press, 2020.
[72] K. Saffarizadeh, M. Boodraj, T. M. Alashoor, et al., Conversational assistants: Investigating privacy concerns, trust, and self-disclosure., in: ICIS, 2017.
[73] G. Laban, T. Araujo, The effect of personalization techniques in users' perceptions of conversational recommender systems, in: Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents, 2020, pp. 1–3.
[74] S. Ghosh, Designing human-computer conversational systems using needs hierarchy, School of Information Student Research Journal 11 (2021) 3.
[75] A. Maslow, K. Lewis, Maslow's hierarchy of needs, Salenger Incorporated 14 (1987) 987.