Hey Google, Do You Have a Personality? Designing Personality and Personas for Conversational Agents

Alisha Pradhan
University of Maryland, College Park
alishapr@umd.edu

Amanda Lazar
University of Maryland, College Park
lazar@umd.edu

ABSTRACT

Conversational agents designed to interact through natural language are often imbued with human-like personalities. At times, the agent might also have a distinct persona with traits such as gender, age, or a backstory. Designing such personality or persona for conversational agents has become a common design practice. In this work, we review the emerging literature on designing agent persona or personality, and reflect on these approaches, along with the personas that are created for common conversational agents. We discuss open questions with regards to three aspects: meeting user needs, the ethics of deception, and reinforcing social stereotypes through conversational agents. We hope this work can provoke researchers and practitioners to critically reflect on their approach for designing personality or persona of conversational agents.

CCS CONCEPTS

• **Human-centered computing**; • **Human computer interaction (HCI)**; • **HCI design and evaluation methods**;

KEYWORDS

Conversational agents, voice assistants, persona, personality

ACM Reference Format:
Alisha Pradhan and Amanda Lazar. 2021. Hey Google, Do You Have a Personality? Designing Personality and Personas for Conversational Agents. In *CUI 2021 - 3rd Conference on Conversational User Interfaces (CUI '21), July 27–29, 2021, Bilbao (online), Spain.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3469595.3469607

1 INTRODUCTION

Conversational agents (such as chatbots [18] or intelligent personal assistants [19]), are software systems that can interact with users using natural language similar to human-to-human conversation [6]. Conversational agents have long been used in healthcare [16], e-learning [14] or for social companionship purposes [29]. More recently, a growing number of individuals are using commercial conversational voice agents such as Alexa, Siri, Cortana, or Google Assistant, which are embodied in different devices such as smartphones or speakers [35].

One practice that is becoming widespread with researchers in industry and academia is to design conversational agents with particular personalities in mind [5, 18, 25, 36]. Personality is defined as "a set of traits that is stable across situations and time and acts as a guiding influence on agent behavior and interactions" [17]. For example, some sources mention that Alexa is designed to be "smart, approachable, humble, enthusiastic, helpful and friendly" [5, 15, 25] or Siri is designed to be "friendly and humble – but also with an edge" [1, 24, 34]. Another common practice with the goal of imbuing human-like traits in conversational agents is to create personas[1] for the agent. A persona for a conversation agent is a fictional character and can have a name, age, education or job, or even a defined backstory and personalities [11, 15]. Below are two examples of persona for two conversational agents: XiaoIce (Microsoft's China-based conversational agent [10]) and Google Assistant.

> XiaoIce has a persona of as an 18-year-old girl who is always reliable, sympathetic, affectionate, and has a wonderful sense of humor. Despite being extremely knowledgeable due to her access to large amounts of data and knowledge, XiaoIce never comes across as egotistical and only demonstrates her wit and creativity when appropriate [33].

> Google assistant is a young woman from Colorado, the youngest daughter of a research librarian and physics professors who has a B.A. in history from Northwestern, an elite research university in the United States; and as a child, won US$100,000 on Jeopardy Kids Edition, a televised trivia game. Google Assistant used to work as a personal assistant to a very popular late night TV satirical pundit and enjoys kayaking [12, 30].

Some research argues that having a distinct persona or personality can contribute a cohesive and consistence presence of the conversational agent for users [17] and increase trust and the intention to use the technology [21]. As such, researchers seem to be beginning to adopt personas in their study of conversational agents as well. For example, Liao and He [18] created personas for conversational agents that had distinct gender and race to understand user preferences [18]. Below, we review current approaches to designing personality and persona of conversational agents. We then discuss some open questions and implications of these approaches.

CUI '21, July 27–29, 2021, Bilbao (online), Spain

ACM ISBN 978-1-4503-8998-3/21/07…$15.00
https://doi.org/10.1145/3469595.3469607

[1] In some works, persona and personality are used interchangeably in discussing conversational agents (e.g., [17]), possibly because these concepts are relatively new, though personas appear to be more descriptive and elaborate. Both concepts impart human-like traits to the agent and serve the same purpose, i.e., to create a cohesive and consistent presence of the agent.

2 DESIGNING PERSONALITY AND PERSONA OF CONVERSATIONAL AGENTS

Although imparting personality or persona to conversational agents appears to be a common practice, there is scarce research on the design decisions involved in the actual practice of designing personalities and persona. Below we review the approaches discussed in the literature.

Lessio and Morris [17] propose using the psychological personality OCEAN model (also referred as the Big Five model: Extroversion, Agreeableness, Conscientiousness, Neuroticism, and Openness) to model the personality of conversational agents. Based on this OCEAN model, they map human personality traits to conversational agents. For example, the trait of "agreeableness" translates to "empathy" for the agent (i.e., "is the agent empathetic? can it remember the user and emotionally support them?"). They also consider the persona of the user (a fictional identity construct of the user representing some characteristics of the user along with their needs and goals [2]) and map the personality of the agent to the user persona. This approach is recommended in other research as well [9]. Based on this mapping, if the user persona is "engaging" then the agent would be "engagement oriented" and interact with the person via emotions and psychology.

However, some emerging research indicates that classifying the personality of the conversational agent based on the OCEAN model might not be as straightforward as it appears. Völkel et al. argue that commonly used Big Five model for human personality does not adequately describe agent personality and they propose a personality model with alternative dimensions (confrontational, dysfunctional, serviceable, unstable, approachable, social-entertaining, social-inclined, social-assisting, self-conscious, and artificial) [31]. Kim et al. conducted design workshops to understand how practitioners design personality for conversational agents (specifically healthcare agents) [15]. Here, most participants designed the agent to be "empathizing, trustworthy, submissive, and smart-yet-modest" with some even describing "healthcare agent is a supporter... so it would be very docile". Some individuals also assigned distinct traits (that varied from designer to designer) such as "kind-hearted", "witty". However, when it came to traits such as race, gender, economic level or politics, their participants determined that these should remain abstract, and did not specify them.

Few works mention designing agent personas guided by user data. As one example of a project that is guided by user data, persona XiaoIce was designed based on a large scale analysis of human conversations [33]. In doing so, the designers found that the majority of "desired" users are young and female. Hence, they designed XiaoIce's persona around an "18-year-old girl" [33]. As another example, Danielescu and Christian [5] designed personas for a conversational coaching system where they involved customers by interviewing them and brainstorming with them, finding that their preferences may vary based on their culture and region.

3 REFLECTING ON CURRENT PRACTICES AND THEIR IMPLICATIONS

Above we described approaches for designing agent personality/persona. Here, we reflect on these current approaches as well as their outcomes – the personas and personalities that are imparted to existing conversational agents. We discuss possible implications of these approaches, along with open questions and directions for future research.

3.1 Are we really meeting user needs?

Some researchers (discussed above), suggest modeling agent persona/personality based on user persona [9, 17]. This could be successful if the user persona is accurate. However, HCI researchers are pointing out how user personas are not always accurate as they are often based on designers' assumptions [2]. Another assumption made in modeling agents after user personas is that the user *wants* the agent to be modelled around their personality. This assumption may not always result in the kinds of interactions that users find meaningful. Additionally, mapping roles for agents to traits may involve incorrect assumptions by designers. Kim et al.'s study [15] provides a glimpse of such a phenomenon, when practitioners approached designing a healthcare agent. In this study, designers mapped the agent's role to the designed personality, i.e., since the healthcare agent is a "supporter" it should be "docile" and "submissive". But do users want a supporter to be submissive and docile? Also, what design decisions are involved in translating these traits into behaviors? How is a submissive and docile agent perceived by users?

Possibly because of such assumptions and limited user involvement in design, some research suggests that users might not always perceive the resulting conversations as meaningful. The literature includes reactions to programmed personality in conversational agents as weird [3] or fake, dry, boring, unfunny, and artificial [7]. This suggests that we need to think of ways to appropriately capture actual user preferences to design interactions that would be more meaningful to users. One approach here could be to include users directly in the process of creating personas through co-design approaches. Parallel to how Kim et al. conducted design workshops with *practitioners* to understand how they designed agent personality [15], researchers could conduct similar workshops with diverse *users* of conversational agents (diversity with respect to age, gender, education, living condition, etc.) to understand users' preferences for agent personality.

3.2 The ethics of deception

Creating personality and personas for conversational agents makes human-agent interaction more naturalistic. But this could be concerning if this naturalistic interaction results in deceiving the user into feeling and acting differently than they might intend. For example, one study found that the human-like qualities of the agent affect how users might act towards the agent, where some users did not want to hurt the agent's feelings [3]. Users might personify these agents (i.e., ascribe human-like properties, characteristics, or mental states to nonhuman agents [8]) and perceive them as human-like to even associate social roles of friend or a partner with the agent [19, 22, 23]. Some researchers are even questioning, could these personas designed for having more human-like conversational agents be tricking people to share more data with the agent [32].

This body of research impose new questions: can conversational agents, that are programmed to emulate human-like characteristics,

really change the way we perceive and interact with our technologies? What does it mean for humans to think about the feelings of an agent? By designing more human-like agents are we tricking users in some way or persuading them to do something they would not otherwise? Researchers working on Human-Robot Interaction have already cautioned us of the underlying concerns of such deception (including misplaced trust, replacing human care) [27, 28], and is something we should not be sidelining as we design the next generation of conversational agents. In doing so, we also need to critically question the power relation between the agents and humans, and how creating human-like agents can impact users' autonomy [26].

3.3 Reinforcing stereotypes in our society

Looking closely at the personas described above (e.g., Google Assistant, XiaoIce), we see a common thread that these agents are modeled after young women, who are typically portrayed as having an assistant or "butler" like subservient roles [4, 17]. Kim et al.'s participants also designed their healthcare agents to be submissive and docile. But what happens when this agent is female and is portrayed as submissive and docile? Will this reinforce negative gender stereotypes and also marginalize non-binary individuals, as some researchers begin to argue [4, 20, 30, 32]? We argue that examining the persona that designers create – which play a key role in determining the voice of the agent and programmed conversations [11]– is key to understanding alternatives to existing, sometimes problematic, approaches.

Gender, as described above, has been recognized as a trait embodied in many personas in ways that may reinforce negative stereotypes. There is much work to be done to understand how personas and personalities might inadvertently be perpetuating other harmful stereotypes that are yet to be unveiled. For example, most of the personas we reviewed had no explicit mention of race or region. While not having these traits makes sense in a way, since it is not a living thing to have such qualities, but what does it mean to design a fictional human-like character that does not have a race or belong to a particular region? Looking at this through the lens of racial blindness perpetuated through technology (which enables unintentional racism through claims of ignoring race) [13], one could argue that the default normal might be a *White Woman from the West*. In other words, not explicitly mentioning race for a human-like conversational agent might still be reinforcing certain dominant stereotypes. As we create such human-like personas for agents, we should ask if our approach is in the interest of a more equitable society. Would the interactions resulting from the dominant persona be perceived as meaningful by users who are racially diverse and from diverse geographical locations? Will it reinforce the default for all users? Even though some emerging work shows racial preferences might not disappear in human-agent interaction [18], we still continue to follow a race blind approach to designing conversational agents. Future work needs to further investigate approaches to create (or perhaps not create) agent personality/persona in ways that do not reinforce such stereotypes in our society.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation award IIS-1816145.

REFERENCES

[1] Arnold Kim. 2011. How Apple Approached Developing Siri's Personality. *MacRumors*. Retrieved April 6, 2021 from https://www.macrumors.com/2011/10/15/how-apple-approached-developing-siris-personality/

[2] Yen-ning Chang, Youn-kyung Lim, and Erik Stolterman. 2008. Personas: from theory to practices. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges* (NordiCHI '08), 439–442. https://doi.org/10.1145/1463160.1463214

[3] Benjamin R. Cowan, Nadia Pantidi, David Coyle, Kellie Morrissey, Peter Clarke, Sara Al-Shehri, David Earley, and Natasha Bandeira. 2017. "What can i help you with?": infrequent users' experiences of intelligent personal assistants. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (MobileHCI '17), 1–12. https://doi.org/10.1145/3098279.3098539

[4] Andreea Danielescu. 2020. Eschewing Gender Stereotypes in Voice Assistants to Promote Inclusion. In *Proceedings of the 2nd Conference on Conversational User Interfaces* (CUI '20), 1–3. https://doi.org/10.1145/3405755.3406151

[5] Andreea Danielescu and Gwen Christian. 2018. 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* (CHI EA '18), 1–9. https://doi.org/10.1145/3170427.3174366

[6] Stephan Diederich, Alfred Brendel, and Lutz Kolbe. 2019. Towards a Taxonomy of Platforms for Conversational Agent Design. *Wirtschaftsinformatik 2019 Proceedings*. Retrieved from https://aisel.aisnet.org/wi2019/track10/papers/1

[7] Philip R. Doyle, Justin Edwards, Odile Dumbleton, Leigh Clark, and Benjamin R. Cowan. 2019. Mapping Perceptions of Humanness in Intelligent Personal Assistant Interaction. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services* (MobileHCI '19), 1–12. https://doi.org/10.1145/3338286.3340116

[8] Nicholas Epley, Adam Waytz, and John T. Cacioppo. 2007. On seeing human: a three-factor theory of anthropomorphism. *Psychological Review* 114, 4: 864–886. https://doi.org/10.1037/0033-295X.114.4.864

[9] Catia M Ferreira and Sviatlana Hoehn. Crafting Conversational Agents' Personality in a User-centric Context. In *Proceedings of the 31st Benelux Conference on Artificial Intelligence (BNAIC 2019) and the 28th Belgian Dutch Conference on Machine Learning (Benelearn 2019)*, 2.

[10] Geoff Spencer. 2018. Much more than a chatbot: China's Xiaoice mixes AI with emotions and wins over millions of fans. *Microsoft Stories Asia*. Retrieved April 6, 2021 from https://news.microsoft.com/apac/features/much-more-than-a-chatbot-chinas-xiaoice-mixes-ai-with-emotions-and-wins-over-millions-of-fans/

[11] Google. Create a persona - Conversation design process - Conversation design. Retrieved April 2, 2019 from https://designguidelines.withgoogle.com/conversation/conversation-design-process/create-a-persona.html#create-a-persona-how-do-i-create-one

[12] Isobel Asher Hamilton. 2018. Google gave Google Assistant a weirdly specific backstory to get her voice just right. *Business Insider*. Retrieved September 14, 2020 from https://www.businessinsider.com/google-gave-google-assistant-a-weirdly-specific-backstory-2018-10

[13] Ihudiya Finda Ogbonnaya-Ogburu, Angela D.R. Smith, Alexandra To, and Kentaro Toyama. 2020. Critical Race Theory for HCI. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (CHI '20), 1–16. https://doi.org/10.1145/3313831.3376392

[14] Alice Kerry, Richard Ellis, and Susan Bull. 2009. Conversational Agents in E-Learning. In *Applications and Innovations in Intelligent Systems XVI*, 169–182. https://doi.org/10.1007/978-1-84882-215-3_13

[15] Hankyung Kim, Dong Yoon Koh, Gaeun Lee, Jung-Mi Park, and Youn-kyung Lim. 2019. Designing Personalities of Conversational Agents. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3290607.3312887

[16] Liliana Laranjo, Adam G Dunn, Huong Ly Tong, Ahmet Baki Kocaballi, Jessica Chen, Rabia Bashir, Didi Surian, Blanca Gallego, Farah Magrabi, Annie Y S Lau, and Enrico Coiera. 2018. Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association* 25, 9: 1248–1258. https://doi.org/10.1093/jamia/ocy072

[17] N. Lessio and A. Morris. 2020. Toward Design Archetypes for Conversational Agent Personality. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 3221–3228. https://doi.org/10.1109/SMC42975.2020.9283254

[18] Yuting Liao and Jiangen He. 2020. Racial mirroring effects on human-agent interaction in psychotherapeutic conversations. In *Proceedings of the 25th International Conference on Intelligent User Interfaces*, 430–442. https://doi.org/10.1145/3377325.3377488

[19] Irene Lopatovska and Harriet Williams. 2018. Personification of the Amazon Alexa: BFF or a Mindless Companion. In *Proceedings of the 2018 Conference on Human Information Interaction&Retrieval - CHIIR '18*, 265–268. https://doi.org/10.1145/3176349.3176868

[20] Marie Louise Juul Søndergaard and Lone Koefoed Hansen. 2018. Intimate Futures: Staying with the Trouble of Digital Personal Assistants through Design Fiction. In *Proceedings of the 2018 Designing Interactive Systems Conference* (DIS '18), 869–880. https://doi.org/10.1145/3196709.3196766

[21] Sara Moussawi and Raquel Benbunan-Fich. 2020. The effect of voice and humour on users' perceptions of personal intelligent agents. *Behaviour & Information Technology* 0, 0: 1–24. https://doi.org/10.1080/0144929X.2020.1772368

[22] Alisha Pradhan, Leah Findlater, and Amanda Lazar. 2019. Phantom Friend or Just a Box with Information: Personification and Ontological Categorization of Smart Speaker-based Voice Assistants by Older Adults. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW: 214. https://doi.org/10.1145/3359316

[23] Amanda Purington, Jessie G. Taft, Shruti Sannon, Natalya N. Bazarova, and Samuel Hardman Taylor. 2017. "Alexa is My New BFF": Social Roles, User Satisfaction, and Personification of the Amazon Echo. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (CHI EA '17), 2853–2859. https://doi.org/10.1145/3027063.3053246

[24] Rhodri Marsden. 2015. When Siri says some-thing amusing, is it a man or a machine? *The Independent*. Retrieved April 6, 2021 from https://www.independent.co.uk/life-style/gadgets-and-tech/features/when-siri-says-something-amusing-who-makes-it-happen-man-or-machine-10394914.html

[25] Janko Roettgers. 2019. How Alexa Got Her Personality. *Variety*. Retrieved July 28, 2020 from https://variety.com/2019/digital/news/alexa-personality-amazon-echo-1203236019/

[26] Supraja Sankaran, Chao Zhang, Mathias Funk, Henk Aarts, and Panos Markopoulos. 2020. Do I have a say? Using conversational agents to re-imagine human-machine autonomy. In *Proceedings of the 2nd Conference on Conversational User Interfaces* (CUI '20), 1–3. https://doi.org/10.1145/3405755.3406135

[27] Amanda Sharkey and Noel Sharkey. 2012. Granny and the robots: ethical issues in robot care for the elderly. *Ethics and Information Technology* 14, 1: 27–40. https://doi.org/10.1007/s10676-010-9234-6

[28] Amanda Sharkey and Noel Sharkey. 2020. We need to talk about deception in social robotics! *Ethics and Information Technology*. https://doi.org/10.1007/s10676-020-09573-9

[29] Cameron Smith, Nigel Crook, Simon Dobnik, Daniel Charlton, Johan Boye, Stephen Pulman, Raul Santos de la Camara, Markku Turunen, David Benyon, Jay Bradley, Björn Gambäck, Preben Hansen, Oli Mival, Nick Webb, and Marc Cavazza. 2011. Interaction Strategies for an Affective Conversational Agent. *Presence: Teleoperators and Virtual Environments* 20, 5: 395–411. https://doi.org/10.1162/PRES_a_00063

[30] UNESCO. 2019. I'd blush if I could: closing gender divides in digital skills through education. Retrieved June 18, 2019 from https://unesdoc.unesco.org/ark:/48223/pf0000367416.page=1

[31] Sarah Theres Völkel, Ramona Schödel, Daniel Buschek, Clemens Stachl, Verena Winterhalter, Markus Bühner, and Heinrich Hussmann. 2020. 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* (CHI '20), 1–14. https://doi.org/10.1145/3313831.3376210

[32] Heather Suzanne Woods. 2018. Asking more of Siri and Alexa: feminine persona in service of surveillance capitalism. *Critical Studies in Media Communication* 35, 4: 334–349. https://doi.org/10.1080/15295036.2018.1488082

[33] Li Zhou, Jianfeng Gao, Di Li, and Heung-Yeung Shum. 2020. The Design and Implementation of XiaoIce, an Empathetic Social Chatbot. *Computational Linguistics* 46, 1: 53–93. https://doi.org/10.1162/coli_a_00368

[34] 2011. Apple gave Siri AI personal assistant its edge. *MacDailyNews*. Retrieved April 6, 2021 from https://macdailynews.com/2011/10/16/apple-gave-siri-ai-personal-assistant-its-edge/

[35] Voice assistants used by 46% of Americans, mostly on smartphones. *Pew Research Center*. Retrieved April 1, 2021 from https://www.pewresearch.org/fact-tank/2017/12/12/nearly-half-of-americans-use-digital-voice-assistants-mostly-on-their-smartphones/

[36] Her name is Cortana. Her attitude is almost human. | Engadget. Retrieved August 7, 2020 from https://www.engadget.com/2014-06-04-cortana-microsoft-windows-phone.html