

How do conversational interfaces become gendered?

Katarzyna Cieślak¹[0000-0002-5539-4313]

¹ Faculty of Humanities, AGH University of Krakow, Kraków, Poland

kcieslak@agh.edu.pl

Abstract

The category of gender has been extensively studied in relation to communication with bots (Curry & Rieser, 2018), cooperation in IT teams (Bear & Woolley, 2011), and the use of technology (Obinali, 2019). However, it has not been adequately analyzed from a sociological perspective as a category embedded in the design process and co-generates technology. Thus, the objective of this project is to identify and analyze gender patterns that are produced and reproduced in the process of designing conversational interfaces. The definition of gender used in this project is influenced by integrative gender theories (Connell, 2009; Risman & Davis, 2013; Leszczyńska, 2016). Drawing inspiration from Raewyn Connell's concept of gender regimes (2009), the focus primarily analyzes gender rules and practices identified in the design process regarding four gender structures: power order, labor order, symbolic dimension, and body-emotional. The empirical part of the project is based on qualitative research techniques, such as individual in-depth interviews with individuals involved in the implementation of chatbots or voicebots, and participatory observation in Polish organizations that have deployed conversational interfaces.

Keywords: conversational interfaces, gender, feminist technology, design process, gender structures.

1 Introduction

1.1 The aim of the project

The aim of my research is to identify and analyse gender patterns produced and reproduced (further: (re)produced) in the design processes of artificial intelligence (AI) systems based on human-machine communication. These machines are systems that I understand in the project as computer programs divided into text-based (chatbots) and voice-based (voicebots) interfaces that, through proper programming, are designed to sustain a conversation with a human and imitate human speech (McTear et al. 2016). I define gender patterns taking inspiration from gender integration theory as a recursive, complex relationship between social norms and reproducible gendered social practices (Connell 2009; Risman & Davis 2013; Leszczynska 2016). I want to analyse how conversation interfaces are developed and what models of gender, understood as social practices and rules, are produced and reproduced in the process of designing conversation interfaces. In my project, I intend to focus on analysing selected design processes, in which different specialists who are involved in the process of creating Polish-language bots participate in organisations in the high-tech industry which specialise in commercial bot implementation. The research problem stated in the project, therefore, concerns the question: Which gender patterns are reproduced in the process of designing conversation interfaces? The underlying assumption for the project is that genderisation of interfaces also takes place in the design process, before the final outcome of the process is used by the final user. The development process of interface technologies is a unique and complex undertaking, which requires involvement (to its limited-in-time development) of qualified persons from different fields, who work in line with the adopted method of project management.

1.2 Gender and technology – main research areas

Research pertaining to the relationship between gender and technology has been conducted from sociological perspective mainly as part of Science Technology and Studies (i.a. Adam 2006; Lagesen 2012) or Feminist Technology Studies (i.a. Faulkner 2001;) also defined as technofeminism (Wajcman 2004). That research concerns different fields, including how technology is gendered in terms of use and design, and it explores gender “in and of” technological artifacts (Faulkner 2001), or how women acquire knowledge about computers as part of their ways of doing gender (Lagesen 2012). Overall, it mainly concerns the idea that gender and technology are co-constructed, as technology itself is not gender-neutral, and several dimensions of gender are reproduced in the design and usage of technology (Zou, Schiebinger 2018). Other important topics taken up in research on gender and technology present multi-gender ways and skills of using technology (Ono & Zavodny, 2002; Cai et al., 2017), gendered design assumptions for end users of computer programs and domestic devices or robots for home usage (Bath 2014; Søråa 2017). On the other hand, conversation interfaces have thus far often been

analysed in terms of their audience. There has been research on the processes of the anthropomorphising of voice assistants by their users and the consequences of consolidating stereotypical gender beliefs this entails (Obinali 2019; Wagner & Schramm-Klein, 2019). In research on genderization of technology, gender is often treated as a default and intuitive category, rarely as a complex and multidimensional structure. Gender models that I would like to identify and explain are inspired by integrative gender theories, which place gendered practices at the center of analysis. My research examines these practices through an analysis of the activities involved in the process of designing conversational interfaces. Integrative theories combine different perspectives to explain gender practices, which may be conditioned by a certain system, structure or social institution. Therefore in this project by examining practices in technology development we can gain a univocal perspective for understanding the process of stereotype production. Furthermore, artificial intelligence is not only studied in the context of technology development but also in terms of social sciences (Liao & Sundar 2022) but rarely associated with sociological lenses that incorporated concepts from gender studies which my project is undertaken. In terms of social sciences there are several domain that covers analysis of tensions between AI development and society including philosophy (Mohamed et al., 2020), more specifically ethics (Paraman & Anamalah, 2023), cognitive science (Zhao et al., 2022), or psychology (Ho et al., 2018) particularly in scope of human machine interaction studies (Liao et al., 2020; Xie & Pentina, 2022). Moreover, due to growing numbers of implementations of AI methods in decision-making systems in industries such as finance (Bredt, 2019), healthcare (Rajpurkar, 2022) or hiring (Kodiyar, 2019) the need for transparency and explainability becomes notable. Therefore, concept of explainable AI which addressed those needs is gaining more attention especially with the growing complexity and numbers of parameters in the generative AI based systems (Zini & Awad, 2022).

2 Perspective of designers as social actors

It is worth to note that it is the author (designer) who makes the decisions on the ways a bot is to be programmed: how it is to communicate with its audience, what content it is to convey, and what gendered characteristics it is to present. Development of conversation interface technology is, therefore, not only a technological process, but also a social one, where creators of a technology involved in the process of its development, make decisions regarding which training data and which language (corpora) models are to be uploaded. Even though AI, along with subfields such as machine learning (using various training data for remembering communication patterns and models), artificial neural networks used for speech recognition and necessary to process information, and natural language processing itself are the underlying technologies behind conversation interfaces (Masche, Le 2017), embodied women and men, who take different actions and take different positions in the process, are behind these technologies (Wellner, Rothman 2020).

Different specialists in the AI industry are responsible for the process of interface development, including business analysts, content writers, programmers, and project managers. Bearing these insights in mind, in the project I analyse how conversation interfaces (chatbots and voicebots) are developed and what models of gender, understood as social practices and rules, are produced and reproduced in the process of designing conversation interfaces. In that process, there are people who make decisions regarding which anthropomorphic features given bot will have, how the bot will communicate (formally or informally, e.g. by using slang words or joking), what the order of such messages will be, and finally, what the content of such messages will be. Therefore, it requires communication between persons from different stages of the same endeavour within a design team, as well as development of methods and tools to share expertise which is used to reconcile the various needs of users, business requirements, and technical capabilities.

3 Concept of gender and methodology

In the project I use the gender definition based on integrative theories treated gender as the relationship between social practices and social rules which emphasize gender differences that define femininity and masculinity (Connell 2009). These relations contest gender differences and organise inequalities based on that difference. The relationships between rules and practices are framed in structures that I analyse in the different stages of the design process. Inspired by Raewyn Connell's concept of gender regimes (2009), I will focus primarily on analysing the gender rules and practices that shape the four basic gender structures: (1) Gendered power: What positions do women and men have in the design process? Who makes final decisions on approval of a bot version? What is the structure and order of communication in a design team?; (2) Labour order by studying the division of responsibilities in a team and which tasks are assigned to women, and which tasks are assigned to men when designing the content of a conversation for a bot; (3) Symbolic structures, considering, for example, how femininity and masculinity are interpreted at the stage of bot personality modelling (e.g. metaphorized, depicted linguistically). How are the tone of voice and modulation of played messages in bots that have female or male names selected? How are bots given names?; (4) Body-emotional structures (so-called cathexis), concerning questions about the emotional practices to which a bot imagined as a woman or a man refers to. The empirical part of the project is take the form of a multi-sited analysis, for which I used qualitative research techniques such as individual in-depth interviews and participant observation. As far as in-depth interviews are concerned, the research sample was selected on purpose, initially based on the snowball method. The research sample includes both women and men who participate at different stages of conversation interface design processes and belong to the group of so-called creators: designers, business analysts, developers, and project managers in Polish organisations that implement AI-based solutions (mostly customer services). Currently, the size of research sample is 23, but I will conduct 7 interviews by the end

of the year. Open participatory observation took place in one organisation that implements projects related to conversation interfaces for Polish users (customer service voicebot). The observation time was two months due to the cyclical nature and duration of design tasks. I observed activities related to the implementation of individual stages of designing conversation interfaces, which include a.o. ways of obtaining training data, the project tools and techniques used, preparation of documentation for model audiences and a bot image, team meetings concerning, e.g. concept decisions, the process of creating content of messages played by bots, and how the designed solutions are evaluated.

4 Preliminary findings and further research

At present, the project is in the stage of analyzing the collected research material to define the situation in which respondents find themselves when designing and implementing chatbots and voicebots. In my project, the concept of situation refers to a category of sociological analysis that is inspired by Barbara Risman's (2018) theory of gendered practices. Practices is understood as a sequence of gender-constructing actions that are conditioned by the structures and contexts in which they occur. In my project I identify at which stages of designing the bot, designers perform actions that give gendered meaning to the interface. By incorporating Raewyn Connell's theory on gendered regimes, I highlight the significance of labour divisions in design process. Apart from understanding how respondents perceive their experiences of different expressions of power, I also want to explore their usage of facilities such as tools, design techniques, and project management methods that are enhancing inclusive design strategies. Preliminary findings suggest that there are recurring business and technical constraints related to the decision of chatbot or voicebot gender. The study reveals patterns of choosing female voices, particularly in industries such as healthcare and banking customer service, which were initially established by stakeholders. Respondents also point to "technical reasons," as female voices tend to sound less robotic. It appears that time pressure and high development costs are the main reasons why individuals choose not to change client decisions if they prefer only female voices or female chatbot avatars. On the other hand, some individuals mentioned that workshops with clients can be a valid opportunity to discuss and identify potential biased content in outputs and/or graphic visualizations of bot personality. In such cases, designers can reflect on their resources to change the situation (context), attempt to increase awareness of stereotypical assumptions, and implement some inclusive design strategies. The resources mentioned by the respondents, conceptualized in my study as opportunities and constraints, provide a broader perspective on the boundary conditions of design decisions and define designer situation. The interviewees highlighted a conflict that arises in deciding on the image of bots. They prefer gender-neutral images, communication style and assistant's voice, but clients often suggest gendered characteristics.

The next stage of the research will analyze power structures and the designers' sense of agency in the context of the decision to give gendered characteristics to bots. In conclusion, a systematic and professional analysis of the material is crucial to comprehend the social relations created during the design process. This approach provides insight into the behavior of designers, the potential for reconstructing structures, and the use of facilities to build inclusive design strategies.

Additional information

Funding

The research on which this project presentation is based was supported by the Polish National Science Centre (decision number UMO-2022/45/N/HS6/02725).

References

1. Adam, A. (2006). *Artificial knowing: Gender and the thinking machine*. Routledge.
2. Adams, R. (2021). Can artificial intelligence be decolonized?. *Interdisciplinary Science Reviews*, 46(1-2), 176-197.
3. Bath, C. (2014). Searching for methodology. *Feminist Technology Design in Computer Science*. P
4. Bredt, S. (2019). Artificial Intelligence (AI) in the financial sector—Potential and public strategies. *Frontiers in Artificial Intelligence*, 2, 16.
5. Cai, Z., Fan, X., & Du, J. (2017). Gender and attitudes toward technology use: A meta-analysis. *Computers & Education*, 105, 1-13.
6. Connell, R. (2009). *Gender*. Short Introduction, Cornwall: Polity Press.
7. Curry, A. C., & Rieser, V. (2018). # MeToo Alexa: How conversational systems respond to sexual harassment. In *Proceedings of the second acl workshop on ethics in natural language processing* (pp. 7-14).
8. Faulkner, W. (2001). The technology question in feminism: A view from feminist technology studies. In *Women's studies international forum* (Vol. 24, No. 1, pp. 79-95). Pergamon
9. Ho, A., Hancock, J., & Miner, A. S. (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication*, 68(4), 712-733.
10. Kodyan, A. A. (2019). An overview of ethical issues in using AI systems in hiring with a case study of Amazon's AI based hiring tool. *Researchgate Preprint*, 1-19.
11. Liao, Q. V., Wang, Y. C., Bickmore, T., Fung, P., Grudin, J., Yu, Z., & Zhou, M. (2019). Human-agent communication: Connecting research and development in HCI and AI. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing* (pp. 122-126).

12. Liao, Q. V., & Sundar, S. S. (2022). Designing for responsible trust in AI systems: A communication perspective. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1257-1268).
13. Lagesen, V. A. (2012). Reassembling gender: Actor-network theory (ANT) and the making of the technology in gender. *Social Studies of Science*, 42(3), 442-448.
14. Leszczyńska, K. (2016). *Płeć w instytucje uwikłana*. Wydawnictwo Naukowe Scholar.
15. Masche, J., & Le, N. T. (2017). A review of technologies for conversational systems. In *International conference on computer science, applied mathematics and applications* (pp. 212-225). Springer.
16. McTear, M. F., Callejas, Z., & Griol, D. (2016). The conversational interface (Vol. 6, No. 94, p. 102). Springer.
17. Mohamed, S., Png, M. T., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology*, 33, 659-684.
18. Obinali, C. (2019). The perception of gender in voice assistants.
19. Ono, H., & Zavodny, M. (2005). Gender differences in information technology usage: A US-Japan comparison. *Sociological Perspectives*, 48(1), 105-133.
20. Paraman, P., & Anamalah, S. (2023). Ethical artificial intelligence framework for a good AI society: principles, opportunities and perils. *AI & SOCIETY*, 38(2), 595-611.
21. Poushneh, A. (2021). Humanizing voice assistant: The impact of voice assistant personality on consumers' attitudes and behaviors. *Journal of Retailing and Consumer Services*, 58, 102283.
22. Risman, B. J., & Davis, G. (2013). From sex roles to gender structure. *Current Sociology*, 61(5-6), 733-755.
23. Risman, B. J. (2018). *Where the millennials will take us: A new generation wrestles with the gender structure*. Oxford University Press.
24. Søraa, R. A. (2017). Mechanical genders: how do humans gender robots?. *Gender, Technology and Development*, 21(1-2), 99-115.
25. Terzopoulos, G., & Satratzemi, M. (2020). Voice assistants and smart speakers in everyday life and in education. *Informatics in Education*, 19(3), 473-490.
26. Wagner, K., & Schramm-Klein, H. (2019). Alexa, Are You Human? Investigating Anthropomorphism of Digital Voice Assistants-A Qualitative Approach. In ICIS.
27. Wajcman, J. (2004). *TechnoFeminism*. Cambridge, UK: Polity Press.
28. Wellner, G., & Rothman, T. (2020). Feminist AI: Can we expect our AI systems to become feminist?. *Philosophy & Technology*, 33(2), 191-205
29. Xie, T., & Pentina, I. (2022). Attachment theory as a framework to understand relationships with social chatbots: a case study of Replika.
30. Zhao, G., Li, Y., & Xu, Q. (2022). From emotion AI to cognitive AI. *International Journal of Network Dynamics and Intelligence*, 65-72.
31. Zini, J. E., & Awad, M. (2022). On the explainability of natural language processing deep models. *ACM Computing Surveys*, 55(5), 1-31.
32. Zou, J., & Schiebinger, L. (2018). AI can be sexist and racist—it's time to make it fair. *Nature* 559, 324-326 (2018)