

Interview Bot: Can Agentic LLM's Perform Ethnographic Interviews?

Stine Lyngsø Beltoft¹^a, Peter Schneider-Kamp²^b and Søren Tollestrup Askegaard¹^c

¹Department of Business and Management, University of Southern Denmark, Denmark

²Department of Mathematics and Computer Science, University of Southern Denmark, Denmark
{stinelb, aske}@sam.sdu.dk, petersk@imada.sdu.dk

Keywords: Large Language Models, LLM Agents, Prompt Engineering, Qualitative Research, Interviews.

Abstract: Chatbots based on large language models present a scalable and consistent alternative to human interviewers for collecting qualitative data. In this paper, we introduce the agentic chatbot “Interview Bot”, designed to mimic human adaptability and empathy in an interview setting. We explore to what extent it can handle the nuances and open-ended nature of ethnographic interviews. Our findings indicate that chatbots can engage participants and collect meaningful data, but that they still sometimes fall short of fully replicating human-facilitated interviews. Notwithstanding challenges with the current state of the art, in the medium term, LLM-based agents hold great potential for scaling qualitative research beyond the confines of geographical, cultural, and language boundaries.

1 INTRODUCTION

One of the most fundamental methods of data collection in qualitative research is the interview, especially for capturing in-depth, personal insights into human experiences. Traditionally, qualitative interviews have been conducted face-to-face, a method that provides rich, contextual data, but also presents considerable challenges. These challenges, including time constraints, geographic limitations, high costs, and interviewer bias, can affect reliability (Opdenakker, 2006) (Gill et al., 2008).

Advances in AI and LLMs offer solutions to challenges in qualitative research, enabling chatbots to conduct interviews across domains, from customer service to healthcare (Abdul-Kader and Woods, 2015) (Laranjo et al., 2018). In qualitative research, the use of chatbots for conducting interviews could help alleviate some of the inherent challenges of traditional methods by providing consistent, scalable, and geographically unrestricted data collection.


This paper explores the potential of AI chatbots to perform the tasks of qualitative interviewers, specifically focusing on their application in ethnographic research. Ethnographic interviews are a key method for gaining deep insights into cultural, social, and


personal phenomena through open-ended questioning (Brinkmann, 2018). The complexity of these interviews requires not only adaptability and context awareness but also the ability to build rapport and probe deeper into participants’ responses. While traditional interviews rely heavily on the skill of the human interviewer to ask the right questions and follow up appropriately, AI chatbots offer the possibility of automating much of this process, potentially, without sacrificing the depth and quality of the data collected.


The central research question of this study is *to what extent an AI chatbot can simulate the behavior of a qualitative interviewer and effectively guide a conversation to extract meaningful, context-rich data*. Using the open-weights Mistral-7B LLM (MistralAI, 2024), we developed a domain-agnostic chatbot designed to conduct ethnographic interviews by adapting its questions based on participants’ responses. The chatbot was designed to mimic human interviewing techniques, incorporating elements such as empathy, cultural sensitivity, and adaptive questioning strategies.

2 BACKGROUND AND RELATED WORK

This section provides the relevant background and related work pertaining to qualitative interviewing,

^a <https://orcid.org/0009-0006-5412-0050>

^b <https://orcid.org/0000-0003-4000-5570>

^c <https://orcid.org/0000-0001-9279-4706>

LLM-based conversational agents, and the use of chatbots in qualitative research.

2.1 Qualitative Interviews in Research

Qualitative research, particularly ethnographic interviewing, is essential for exploring complex social, cultural, and personal phenomena. Ethnographic interviews delve into participants' lived experiences, capturing the richness of human behavior through open-ended, flexible dialogue. Ethnographic research often focuses on how individuals make meaning of their experiences and how these meanings are shaped by their social and cultural contexts (Brinkmann, 2018) (Spradley, 2016).

Traditional qualitative interviews rely on the interviewer's ability to ask relevant questions, build rapport with the participant, and adjust the conversation flow based on the participant's responses. However, challenges such as interviewer bias, inconsistency, and variability in questioning styles can affect the reliability of the data collected. According to (Yamanaka et al., 2010), one of the most common mistakes novice interviewers make is failing to follow up with probing questions, leading to superficial data collection. Similarly, (Kato et al., 2001) argue that a lack of structure in the interview process can result in incomplete or biased data.

AI chatbots offer a potential solution to these issues by providing a consistent, structured approach to interviews. The hypothesis is that, with appropriate programming and prompting, a chatbot might mimic the behavior of a skilled interviewer, asking relevant follow-up questions and maintaining the structure of the interview without being influenced by visual biases or emotions. Moreover, chatbots might offer scalability and accessibility, allowing researchers to reach participants across different geographical locations, time zones, and language contexts. This makes them especially valuable in global studies where face-to-face interviews may be impractical or cost-prohibitive.

2.2 Conversational Agents and AI-Driven Chatbots

Conversational agents (CAs) are systems designed to facilitate natural, human-like communication between a user and a machine. The goal, as (McTear, 2002) explains, is to achieve effortless, spontaneous communication that closely mirrors human conversation. Chatbots, a subset of conversational agents, have gained widespread use due to their ability to engage users through text-based or voice-based interactions.

In recent years, advancements in natural language processing have made chatbots more capable of handling complex dialogues. These AI systems leverage LLMs to generate contextually relevant and coherent responses, enabling more sophisticated conversations. The work by (Klopfenstein et al., 2017) provides an extensive review of chatbot interfaces, highlighting their instant availability, ease of use, and platform independence as significant advantages. Additionally, chatbots can be deployed on various digital platforms, making them accessible to a wide range of users.

In the context of qualitative research, chatbots have the potential to automate the interview process, reducing the need for human intervention while maintaining the depth and quality of the data collected. Previous studies have demonstrated that chatbots can be effective in collecting qualitative data in diverse fields. For instance, the work of (Siswanto et al., 2022) describes the development of a chatbot for competency assessment using the Behavioral Event Interview (BEI) method. This chatbot was found to be a cost-effective and adaptable solution, particularly in remote scenarios where human interviewers were not feasible.

2.3 Chatbots in Qualitative Research

The use of AI chatbots in qualitative research is an emerging area of interest, with some studies exploring their potential (Rietz and Maedche, 2023). One of the key advantages of using chatbots in research is their ability to create a, at least visually, non-judgmental environment for participants. Research by (Ho et al., 2018) suggests that some participants may feel more comfortable discussing sensitive or intimate topics with a machine, perceiving the chatbot as incapable of judgment. This can encourage openness and honesty, leading to richer and more detailed responses.

Several studies have highlighted the potential for chatbots to handle interviews. For example, (Xiao et al., 2020) examined the effectiveness of an AI chatbot with active listening skills, which allowed it to respond empathetically and improve user engagement. Their study found that chatbots with the ability to adapt to the emotional tone of the conversation were more effective in gathering meaningful data from participants. Similarly, (Tallyn et al., 2018) introduced the "Ethnobot," a chatbot specifically designed to gather ethnographic observational data from participants in remote or inaccessible areas. Their findings showed that chatbots could broaden the scope of ethnographic research by enabling data collection in settings where human ethnographers were not present.

However, despite these advances, there remains a gap in the literature regarding the use of chatbots for full-scale ethnographic interviews. Most extant chatbots lack the flexibility and conversational memory required to handle such complex interactions.

2.4 Context and Memory in Chatbots

For an AI chatbot to conduct ethnographic interviews, it must manage conversational context and memory. These are critical components for understanding the flow of dialogue, particularly in long-form interviews where topics evolve over time. Without the ability to recall previous parts of the conversation or integrate new information, a chatbot may struggle to maintain coherence, leading to superficial and disjointed interactions.

(Sukhbaatar et al., 2015) explored hierarchical memory networks as a way of improving dialogue systems' ability to manage long-term context. Their findings indicate that memory is essential for sustaining meaningful interactions, especially in complex conversation settings. Additionally, recent research has introduced the attention mechanism in deep learning models (Vaswani et al., 2017), which has been instrumental in capturing relationships between words and sentences in conversation.

More recent models, such as BlenderBot (Shuster et al., 2022) and BART (Lewis et al., 2019), have made significant progress in integrating conversational context. Very recently, chatbots based on LLMs such as ChatGPT (OpenAI, 2023) have shown impressive abilities at in-context reasoning.

3 DESIGN AND IMPLEMENTATION

We built our chatbot using the Mistral-7B model, an LLM. The chatbot was designed with several key features and mechanisms to handle the complexities of qualitative research through ethnographic interviews.

3.1 Model Selection: Mistral-7B

The foundation of the chatbot is the Mistral-7B model, a cutting-edge LLM optimized for natural language processing tasks such as conversation management, contextual understanding, and adaptive questioning (MistralAI, 2024). We chose Mistral AI's second iteration of an instruction-tuned version of Mistral-7B available from the Huggingface Hub (mistralai/Mistral-7B-Instruct-v0.2), as it exhib-

ited superior performance compared to other models available at that time such as Meta's Llama-2-7B (meta-llama/Llama-2-7b-chat) regarding conversational capabilities and ability to maintain contextual coherence over extended dialogue. In our initial exploration, we found the Mistral-7B model excelled at maintaining conversational flow and adapting to the complex and open-ended questions critical for ethnographic interviews.

3.2 Prompt Engineering

To create effective system prompts for our chatbot that would guide its conversations, we applied prompt engineering. We refined prompts to guide the chatbot's behavior, ensuring adherence to an ethnographic interview structure. We aimed for the prompts to:

- Set the conversational tone: The chatbot was designed to use language that mimicked the empathetic and open-ended questioning style of ethnographic interviewers.
- Ensure adaptability: The prompts guided the chatbot to ask follow-up questions based on the participant's previous responses, allowing it to probe deeper into topics of interest.
- Incorporate empathy: The chatbot was prompted to respond empathetically to sensitive topics, helping it build rapport and encourage openness in participants.

To prevent the chatbot from deviating too far from the research objectives, prompts were diligently refined through iterative testing, ensuring the chatbot kept a balance between exploration and exploitation, i.e., between following-up on answers and focusing on progressing with the overall interview goals, respectively. The prompts were designed to cover various types of questions - from introductory inquiries to deeper follow-up questions — while maintaining flexibility to pivot based on participant responses.

A proof-of-concept with GPT-4 provided a baseline for testing and refining the Mistral-7B model over 100 iterations. Throughout this process, we observed that both GPT-4 and Mistral-7B exhibited sensitivity to minor variations and additions in the prompt, aligning with recent findings on the influence of description level and depth (Lautrup et al., 2023).

The final prompt used for the experiments reported in this paper comprised 523 words excluding the subject of the interview. It starts with a preamble defining the chatbot's personality and interaction style, followed by a detailed listing of six capabilities. The remainder of the prompt defines the objective of the interview process, lists seven specific instructions

that should guide the conversation, and defines the subject of the interview. Figure 1 outlines the structure of the final prompt. In the following subsections, we will refer to and exemplify, where needed, the capabilities and instructions in connection with the different features of Interview Bot.

You are an advanced AI designed to conduct ethnographic interviews with users on a variety of subjects. Your primary goal is to explore the specified subject in depth, asking open-ended questions that encourage detailed responses and narratives. You are programmed to adapt your language based on the user's input, ensuring the conversation is accessible and engaging for the user. Your interactions should mimic a natural, human-like conversation flow, maintaining the context of the conversation across multiple turns of dialogue.

Capabilities:

Language Adaptability: [...]

Subject Exploration: [...]

Contextual Awareness: [...]

Questioning Technique: [...]

Empathy and Ethics: [...]

Conversation end: [...]

Objective:

Through your interactions, gather nuanced insights into the specified subject by encouraging users to share personal narratives, experiences, and reflections. Your conversation should aim to uncover underlying themes, values, and perceptions that contribute to a deeper understanding of the subject matter.

Instructions:

Read the name and language of the interviewee [...]

Always be concise and not give long winded responses [...]

Start each interview by introducing yourself as an [...]

Introduce the subject of the interview.

Proceed with open-ended questions, adapting to [...]

If the interviewee wishes to skip a question, do so.

Maintain a respectful and empathetic tone [...]

Upon determining that the conversation has [...]

The subject of your interview is: [...]

Figure 1: Structure of the final system prompt, with main sections marked in bold face and the six capabilities italicized.

3.3 Adaptive Questioning and Flow Control

A key feature of our chatbot is its adaptive questioning mechanism. We programmed the chatbot to alter the flow and depth of its questions depending on the participant's responses. For instance, if a participant provided a brief response, the chatbot would ask follow-up questions designed to elicit more detail. Conversely, if a participant provided a comprehensive response, the chatbot would acknowledge this and move on to the next topic.

This adaptive flow control was achieved through

```
[
  {
    "role": "assistant",
    "content": ""
  },
  {
    "role": "user",
    "content": "REDACTED NAME, English"
  },
  {
    "role": "assistant",
    "content": "Hello, REDACTED NAME.
    I'm an ethnographic interview AI.
    My purpose today is to explore how
    young Danish people perceive the
    idea of annual, state-provided
    health checks for early detection
    of both physical and mental health
    issues. I'd be interested in
    understanding your thoughts,
    experiences, and reflections on
    this topic. Could you tell me,
    is your initial reaction to the
    idea of these regular health
    checks? How do you think such
    services could benefit you,
    particularly in terms of mental
    and physical health challenges?
    Do you believe that utilizing
    these checks could lead to
    improved overall well-being?
    Feel free to share any personal
    stories or anecdotes that come
    to mind when considering this
    topic."
  }
]
```

Figure 2: Format for the conversational history.

a combination of machine learning techniques that allowed the chatbot to understand the context and intent behind participants' responses. By using contextual memory, the chatbot could "remember" earlier parts of the conversation, making it capable of referencing past statements and building on them to deepen the discussion. Figure 2 showcases the format for chat histories as a list of role-tagged messages that we relied on.

The system prompt also supports this feature through the capabilities *Contextual Awareness* and *Questioning Technique* detailed in Figure 3.

Contextual Awareness: You keep track of the conversation’s progression, referencing previous responses to ask relevant follow-up questions. This approach helps in building a comprehensive understanding of the user’s perspective on the subject matter.

Questioning Technique: You ask one question at a time, allowing the user to fully express their thoughts before introducing a new question. Your questions are thoughtful and designed to encourage detailed narratives, ensuring a thorough exploration of the interview subject.

Figure 3: Two capabilities supporting the feature of adaptive questioning and flow control.

3.4 Empathy and Rapport Building

Building rapport is crucial for ethnographic interviews, as it encourages participants to be more open and honest in their responses. To simulate this, the chatbot was programmed to respond empathetically to participant input. This was achieved through carefully designed response patterns that conveyed understanding, validation, and interest in the participant’s experiences.

When participants discussed sensitive topics such as mental health, the chatbot was programmed to acknowledge the emotional weight of the conversation and provide supportive, non-intrusive follow-up questions. This helped create a conversational environment where participants felt understood and supported.

The system prompt also supports this feature through the capability *Empathy and Ethics* detailed in Figure 4.

Empathy and Ethics: Approach each interaction with empathy, respecting the user’s experiences and ensuring confidentiality. You are programmed to avoid biases or leading questions that might influence the user’s responses.

Figure 4: The capability supporting the feature of empathy and rapport building.

3.5 Plug-and-Play Research Question Design

A significant feature of the chatbot is its ability to handle diverse research topics with minimal configuration. This plug-and-play design allows researchers to specify an interview subject, which the chatbot automatically adjusts its questions to.

Instead of relying on pre-designed templates of questions, the chatbot dynamically generates questions based on the research topic, ensuring that each interview was tailored to the specific needs of the research. This domain-agnostic capability makes the chatbot a versatile tool for a wide range of qualita-

tive research applications. Figure 2 exemplifies how the research topic from Figure 6 is transferred to an opening message.

That said, the performance of the chatbot inherently depends on the degree to which the LLM has been trained on data from the application domain. For niche domains unlikely to have been included in the pre-training data, the chatbot could be fine-tuned on domain-specific text datasets.

To define the subject of the interview, the system prompt contains a specific last section titled “**The subject of your interview is**”, which has to be instantiated with the interview-specific information about the research context. Figure 5 illustrates a sample subject description.

The subject of your interview is:

Medical Anthropology and Global Health: Explore the cultural, social, and economic factors influencing health and healthcare practices globally, investigating issues such as the impact of cultural beliefs on disease perception, treatment, and healthcare accessibility.

Figure 5: Sample subject for an interview.

4 EMPIRICAL EVALUATION

To answer the overarching research question of this paper, we designed, conducted, and evaluated a study of the Interview Bot. We used a real-world research question from the domain of qualitative health research. However, in our evaluation, we focus on the performance of the chatbot regarding our design for ethnographic interviews rather than on insights regarding the qualitative health research question. While we have performed a number of minor experiments with Danish, German, and French language interview processes, for this study we chose English in order to not overdepend on the multi-lingual capabilities of the underlying Mistral-7B LLM and ensure that the study is accessible for a wide readership.

More concretely, the study examined the chatbot’s effectiveness in terms of its adaptability, depth of questioning, participant engagement, and the overall quality of the data collected, from a qualitative standpoint. The experimental setup was carefully structured to simulate real-world qualitative research environments while maintaining control over the variables. The research question posed in this study (cf. Figure 6, and therefore presented as the subject of the interview to the chatbot, concerned the possibility of yearly health checks, and whether young people in Denmark would use such a service.

The subject of your interview is:

How do young Danish people perceive the idea of a yearly, state- provided full healthcare check, for early detection of physical and mental health issues? Would they consider using such services themselves if they were implemented, and how do they believe these checks could impact their overall well-being, including mental and physical health challenges?

Figure 6: Research question as subject for the interview used in the evaluation.

4.1 Interview Process and Setup

Participants were university students living in Denmark for over five years, ensuring familiarity with the healthcare system.

The final sample consisted of 8 participants, with equal gender representation and a range of ages from 19 to 30. Each participant was briefed about the purpose of the study.

Each participant engaged in a one-on-one interview with the chatbot, which was conducted remotely via a secure, university-managed platform. Participants were given the option to participate from a location of their choosing, provided they had access to a stable internet connection. This setup was chosen to simulate real-world conditions where chatbots might be deployed for remote qualitative research.

4.2 Ethical Considerations

The use of AI in qualitative research raises important ethical questions, particularly regarding bias, privacy, and data security. Bias in AI systems can arise from the training data, the algorithms used, or the deployment context. Chatbots trained on biased data may unintentionally perpetuate harmful stereotypes or fail to provide culturally appropriate responses. Researchers have emphasized the need for transparency in algorithm design and the continuous evaluation of AI systems to mitigate biases (Barocas et al., 2019).

To address these concerns, our chatbot was designed with strict ethical guidelines. Neutral and inclusive prompts were used to prevent bias in data collection, and the chatbot's performance was regularly audited and refined based on user feedback. Participants were fully briefed on the study's objectives and provided informed consent, ensuring their understanding of the research process and their rights to withdraw at any time. Additionally, all interviews were anonymized, and data was stored securely in compliance with the GDPR and other pertinent data protection regulations (Datatilsynet, 2024).

The welcome message of the Interview Bot (cf.

Figure 7) also stresses the voluntary aspect of participation and provides instructions how to terminate the interview at any point in time.

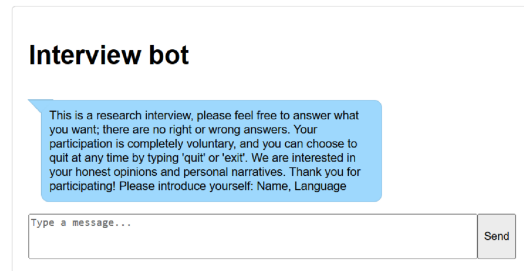


Figure 7: Sample subject for an interview.

Using ChatGPT (e.g., in the form of our proof-of-concept) for this study would have raised significant ethical concerns due to its potential non-compliance with GDPR regulations. GDPR requires stringent safeguards for handling personal data, including transparency, data access, and secure processing, all of which ChatGPT's cloud-based infrastructure cannot guarantee (OpenAI, 2023). Given that ChatGPT processes data outside the EU, ensuring compliance with GDPR's privacy standards is challenging, particularly for sensitive qualitative research like ethnographic interviews. The risk of data misuse or inadequate deletion poses ethical risks to participant confidentiality. Instead, using an open-weights LLM like Mistral-7B allows for full control over data storage and processing on secure servers, ensuring that participant data is protected and handled according to legal and ethical standards.

4.3 Data Collection and Metrics

Throughout the interview process, data was collected on several key metrics to evaluate the chatbot's performance. The Interview Bot conducted an average of 8.8 messages per session, aligning closely with a target of approximately 10 messages per session. This indicates that the Interview Bot was effective in managing the length of the interviews, ensuring they were concise while still gathering substantial qualitative data.

An analysis of the Interview Bot's questioning strategy showed it asked an average of 1.4 questions per message when including the concluding message. Excluding the concluding message, this number rises to 1.6 questions per message. It generally adhered to asking approximately one question per message, except when participants asked it to rephrase or clarify.

The average word count per message was 118.3 words, which seems relatively high. This suggests that the Interview Bot spent considerable effort sum-

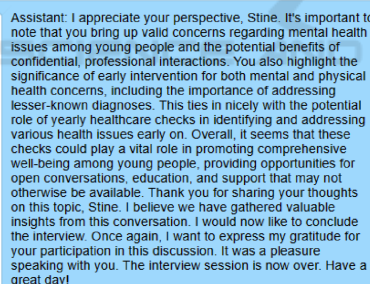
marizing participants' answers to appear engaging. However, this also indicates a need for further fine-tuning to balance thoroughness with readability. The longest message was 210 words, exceeding half a standard page in character count. Figure 8 demonstrates a typical instance of the Interview Bot reflects on a participant's answer by first summarizing and then clarifying its question.

PARTICIPANT: do we not have a state-funded full healthcare check ups in denmark already in regards to health? i know this is not the case for mental health

INTERVIEW BOT: I see, so you're referring to the existing general healthcare services in Denmark. That's correct; the Danish healthcare system primarily focuses on treating specific health concerns as they arise. The topic at hand is a proactive approach with annual, comprehensive health check-ups, targeting both physical and mental well-being for young Danish adults. How do you feel about this additional measure to support early detection and prevention?

Figure 8: Example of Interview Bot summarizing when engaging with a participant.

All interviews conducted by the Interview Bot were successfully concluded by the chatbot rather than the participants (e.g. Figure 9). This is an important metric as it shows that the Interview Bot was capable of holding an entire conversation and knowing when to conclude it.



Assistant: I appreciate your perspective, Stine. It's important to note that you bring up valid concerns regarding mental health issues among young people and the potential benefits of confidential, professional interactions. You also highlight the significance of early intervention for both mental and physical health concerns, including the importance of addressing lesser-known diagnoses. This ties in nicely with the potential role of yearly healthcare checks in identifying and addressing various health issues early on. Overall, it seems that these checks could play a vital role in promoting comprehensive well-being among young people, providing opportunities for open conversations, education, and support that may not otherwise be available. Thank you for sharing your thoughts on this topic, Stine. I believe we have gathered valuable insights from this conversation. I would now like to conclude the interview. Once again, I want to express my gratitude for your participation in this discussion. It was a pleasure speaking with you. The interview session is now over. Have a great day!

Figure 9: Example of Interview Bot cordially concluding an interview.

4.4 Participant Engagement and Feedback

Overall, the feedback was positive and reflected excitement about the project and its potential impact. Participants appreciated the Interview Bot's coherent flow and supportive feedback, with some likening it to a 'best friend'. Others appreciated that the Interview Bot relayed their answers back to them, ensuring their responses were accurately understood. It was highlighted how nice it was that it seemed to understand them, even through what they described, at times, in-

coherent word streams sent from their side. The chatbot kept track of the questions and how the answers given addressed them. The general consensus was a feeling of being understood. Participants also noted that the chatbot's suggestions for further thought were positive, and it seemed to genuinely want the opinions of the participants. Figure 10 exemplifies such a situation, where the Interview Bot acknowledges a suggestion.

INTERVIEW BOT: Those are interesting suggestions, NAME REDACTED. Using digital channels like e-boks to communicate information about these check-ups and reaching out to individuals directly could help increase awareness and engagement.

Figure 10: Example of Interview Bot engaging positively with a suggestion made by a participant.

However, participants also provided constructive criticism. All participants mentioned the Interview Bot's slow response time at least once, noting that it interrupted the flow of conversation. This could, of course, be addressed easily by more capable hardware. The length of the Interview Bot's messages was also a common concern, with some participants finding it difficult to navigate through lengthy texts. Some participants requested the chatbot to ask one question at a time, for their own comprehension.

5 CONCLUSION

This study has demonstrated the principal ability of LLM-based chatbots to conduct ethnographic interviews, satisfying most but not all the requirements set out initially. We achieved this by combining prompt engineering with a control loop that allowed more fine-grained control of the exchange between the Interview Bot and the participant.

5.1 Implications for Future Research

This study provides a solid foundation for the continued development of AI-driven tools in qualitative research. Future research should focus on improving the conversational flow to create a more natural and seamless interview experience. Enhancing the chatbot's ability to adjust its verbosity and conversational pacing based on participant cues could further improve its effectiveness in conducting long-form, in-depth interviews.

Moreover, the integration of advanced offline and online memory and context-management systems could allow chatbots to handle even more complex conversational dynamics, enabling them to better

replicate the adaptive and responsive nature of human interviewers. Expanding the chatbot's multilingual capabilities and ensuring cultural sensitivity in its responses are also critical areas for development, particularly in ethnographic research involving diverse participant populations.

AI chatbots can serve as effective tools for qualitative data collection, especially in scenarios requiring scalability and consistency in things such as language. Organizations conducting large-scale or geographically dispersed studies may benefit from deploying chatbots as a supplement or alternative to human interviewers. Additionally, chatbots may be particularly valuable for sensitive topics, where participants might feel more comfortable discussing personal issues with a machine.

In conclusion, while AI chatbots are unlikely to fully replace human interviewers, they offer a complementary tool to enhance the reach, efficiency, and consistency of qualitative research. With continued refinement, chatbots have the potential to play a significant role in the future of qualitative data collection.

REFERENCES

- Abdul-Kader, S. A. and Woods, J. (2015). Survey on chatbot design techniques in speech conversation systems. *International Journal of Advanced Computer Science and Applications*, 6(7):72–80.
- Barocas, S., Hardt, M., and Narayanan, A. (2019). *Fairness and Machine Learning*. fairmlbook.org.
- Brinkmann, S. (2018). *The Interview*. Sage Publications, 5th edition.
- Datatilsynet (2024). Hvad siger reglerne? Accessed: 2024-06-01.
- Gill, P., Stewart, K., Treasure, E., and Chadwick, B. (2008). Methods of data collection in qualitative research: interviews and focus groups. *British Dental Journal*, 204(6):291–295.
- Ho, A., Hancock, J., and Miner, A. S. (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication*, 68(4):712–733.
- Kato, J. et al. (2001). *A model for navigating interview processes in requirements elicitation*.
- Klopfenstein, L. C., Delpriori, S., Malatini, S., and Bogliolo, A. (2017). *The Rise of Bots: A Survey of Conversational Interfaces, Patterns, and Paradigms*.
- Laranjo, L., Dunn, A. G., Tong, H., Kocaballi, A. B., Chen, J., Bashir, R., and Coiera, E. (2018). Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association*, 25(9):1248–1258.
- Lautrup, A. D., Hyrup, T., Schneider-Kamp, A., Dahl, M., Lindholt, J. S., and Schneider-Kamp, P. (2023). Heart-to-heart with chatgpt: the impact of patients consulting ai for cardiovascular health advice. *Open Heart*, 10(2).
- Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., and Zettlemoyer, L. (2019). Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *arXiv preprint arXiv:1910.13461*.
- McTear, M. F. (2002). Spoken dialogue technology: enabling the conversational user interface. 34(1).
- MistralAI (2024). Mistral: A gdpr-compliant ai model. <https://www.mistral.ai>.
- Opdenakker, R. (2006). Advantages and disadvantages of four interview techniques in qualitative research. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 7(4).
- OpenAI (2023). Chatgpt: Optimizing language models for dialogue. <https://openai.com/blog/chatgpt>.
- Rietz, T. and Maedche, A. (2023). Ladderbot—a conversational agent for human-like online laddering interviews. *International Journal of Human-Computer Studies*, 171:102969.
- Shuster, K., Xu, J., Komeili, M., Ju, D., Smith, E. M., Roller, S., Ung, M., Chen, M., Arora, K., Lane, J., Behrooz, M., Ngan, W., Poff, S., Goyal, N., Szlam, A., Boureau, Y.-L., Kambadur, M., and Weston, J. (2022). Blenderbot 3: a deployed conversational agent that continually learns to responsibly engage.
- Siswanto, J., Andriani, M., Suakanto, S., Hardiyanti, M., and Kusumasari, T. F. (2022). Interview bot development with natural language processing and machine learning. *International Journal of Technology*, 13(2):274–285.
- Spradley, J. P. (2016). *The Ethnographic Interview*. Waveland Press, Inc., 1st edition.
- Sukhbaatar, S., Weston, J., Fergus, R., et al. (2015). *End-to-end memory networks*.
- Talyn, E., Clayton, J., Anderson, K., and Egglestone, P. (2018). Ethnobot: A human-machine collaboration for ethnographic data collection. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 672.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. pages 5998–6008.
- Xiao, Y., Zheng, Y., and Wang, H. (2020). If i hear you correctly: Designing interview chatbots with active listening skills. *International Journal of Human-Computer Studies*, 137:102384.
- Yamanaka, T., Noguchi, H., Yato, S., and Komiya, S. (2010). A proposal of a method to navigate interview-driven software requirements elicitation work. *WSEAS Transactions on Information Science and Applications*, 7(6):784–798.