Wonder document

Midterm assignment

[Group 3]

Niels Gabriel Theissen - nielsgt@uio.no
Stian Grimsrud - stiangri@uio.no
Gwendolyn Borchsenius - gwendob@uio.no
Henriette Elde - henreld@uio.no
Harald Maartmann-Moe - haralmaa@uio.no

About us 1

Area of interest 1

Background 2

Methods 3

- Overall approach 3
- Data collection method 3
- Further planned research 4

Findings 5

Appendix 1 - Chabot design 6

Appendix 2 - Machine learning, MovieChatbot 7

Bibliography 9

About us

Our group consists of five master students and one Ph.D. candidate, four of whom have backgrounds from the bachelor program Design, use and interaction at IFI and one with a bachelor in both Media Studies and Media- and Interaction Design from UiB.

Area of interest

Over the past decades, interaction via Graphical User Interfaces (GUI) has become a common type of interaction. In fact, most technologies we use in our daily life enable interaction via GUIs. However, for some people with cognitive or sensory impairments, interaction via GUIs may be challenging, decreasing the user experience, excluding particular user groups. Therefore, people living with cognitive, developmental, intellectual, mental, physical, or sensory impairments have repeatedly ended up as an afterthought, excluded by technology. About 15 percent of the world's population lives with cognitive, social, or physical impairments (WHO, 2020). As these groups make up a smaller part of the total population, they are frequently less represented, or less prioritized during technology development and design. This lack of inclusion in design and development can lead to social exclusion (Foley & Ferri, 2012, p.192). Inclusive design requires additional arrangements or a different type of recruitment, which many technology developers and designers tend to evade.

Fortunately, some technology has also been developed with these groups in mind, benefiting them. Audiobooks, video captioning, remote controls, video conferencing, eye tracking/detection, and many more technologies have increased the quality of life for people living with impairments. Recently, AI has contributed with improved speech recognition, speech synthesis, sign language translation, simplifying content for people living with cognitive impairments, and visual aids describing whatever the user shows their camera. These developments open up technological possibilities and can contribute to the autonomy of previously excluded groups.

Through our groups' previous experience, we have found a possible beneficial application of AI in assistive technology/UD. In Maartmann-Moe's (2019) work with older adults, many participants suggested that sending messages would improve their communication with family and friends whom they value immensely, and often talk about as one of their greatest sources of joy. Message sending has also become an increasingly large part of using services and systems in phones designed for older adults: Some phone manufacturers still use buttons and touch screens that require fine motor skills, accurate timing of actions, and either great vision or extraordinary memory. This makes message sending inaccessible.
We have been thinking about speech recognition and its possible benefit to communication with friends and family. Speech recognition has been suggested as an opportunity to enable

and include older adults in the evolving communication praxis in contemporary day-to-day conversational exchange.

1. Could AI-powered speech recognition extend older adults' communication capabilities with family and friends?

2. Additionally, speech recognition is not perfect. There will be errors on the part of the AI's recognition of speech. How are these incomprehensive handled by systems and users today? How should/could errors be handled?

As a group, our impression is that there is an appeal for more work surrounding these difficulties and opportunities for older adults and AI. We aim to contribute to this area within AI and universal design with our project work.

Background

Our group has been intrigued by the extensive inclusion that has been made possible with AI-infused systems. Recently, speech recognition and speech based interfaces have grown in popularity as they offer a different approach to HCI. Over the past few years, conversation has become a key research area within the field of HCI as many authors argue that conversation is the most natural form for interaction (Luger & Sellen, 2016, p. 5286). Arguably, humans interact most naturally with each other through verbal speech which is why there has been a rise in so-called Conversational Agents (CA) in the HCI field (Luger & Sellen, 2016, p. 5286). CAs can be defined as "dialogue systems often endowed with humanlike behavior" (Luger & Sellen, 2016, p. 5286). Examples for CAs are chatbots, interface agents or virtual assistants such as Alexa, Google Home and Siri. Speech-based interaction offers several advantages compared to conventional GUIs. In addition, speech-based interaction feels natural for humans, communicating with technological devices via speech can also increase efficiency. For instance, the average person can speak 150 words per minute compared to typing 40 words per minute (Boyd, 2018) . Further, speech based interaction may improve user experience in multitasking situations such as car driving. CAs may also let more users use services that previously have only been available on smartphones and less accessible technology. Especially people with cognitive or sensory impairments could benefit from such systems.

However, the dynamics of how and why CAs are used are still poorly understood. There exists also uncertainty around why certain CAs meet user acceptance whereas others are rejected by users. Nevertheless, AI still has opened up possibilities to include and empower people with cognitive, developmental, intellectual, mental, physical, or sensory impairments. We want to look closer at Universal Design of AI-infused systems and if Conversational Agents could improve older people's communication and Universal design.

Methods

Overall approach

To answer the questions that lead our inquiry we will initiate a design process where we apply AI technology to extend the communication capacities of older adults. With this design as a case, we will be better equipped to discuss the intricacies and potential application of AI in design for older adults and their communication. By working with a concrete case we will discover challenges and opportunities that would not necessarily have surfaced from a literature review.

The initial phase of the project will focus on getting to know the current state of use of AI and older adults as users, specifically speech recognition. This will be done by examining existing literature and by interviewing older adults and experts. With this improved understanding we can uncover areas with challenges or potential applications of AI that apply to specific participants, and initiate a design process that aims to directly improve their lives/everyday.

We do not aim to contribute with generalized theories or frameworks, rather with a case study with exemplifying challenges and potential solutions in applying AI. The concrete design process might also serve as inspiration for further beneficial application of AI to extend the capacities of older adults.

The primary questions we seek to answer by examining literature and through interviews of older adults and experts are -

- How do our participants prefer to communicate?
    - What could be better?
- Is AI used in design for older adults?
    - If not: Why?
    - Is speech recognition used in design for older adults?
        - If not: Why?
        - Within communication?
- Are the existing speech recognition technologies inclusive of older adults as users?
    - Are the voice characteristics of older adults compatible with existing speech recognition technology?
        - If not, why/how?

Data collection method

Rather than making a chatbot, we hope to make a prototype that can let users explore and experience interaction with speech recognition. We aim to create a prototype that ideally can be tried out by real users, and iterate on the prototype to overcome initial challenges and discover further challenges and potential. We want to use this prototype to conduct different qualitative data collection methods, combined with a literature study, and a quantitative data collection where we also look at errors in speech recognition that effect the overall user experience.

Due to the pandemic it might be irresponsible to meet in person with older adults. Proxy users might be a safer way to explore this area of Interaction with AI with older adults.

In the exploratory phase of the project, we identified a possible participant and reached out to one of the elderly user's (EU) family members to gain more insight about how they communicate, how they would like to communicate, and in what context.
To ensure the safety and wellbeing of the participant and prevent possible infection the interview with the family member (FM) was conducted via phone. We followed a semi-structured interview guide and rounded off with a small exercise. The exercise was inspired by a future workshop approach where one invites the participant to consider a certain situation in an imaginary setting with no limitations and everything is possible. The purpose with this exercise was to better understand the ideal form of communication for the participant, and prevent the participant from withholding ideas and input due to them not imagining a way to implement these. Even ideas and input from the participants that is not technologically feasible to implement can help us understand their goals, values, and preferences. Both the interview and exercise were useful to better understand how our participant communicates with her mother, in order to find potential beneficial/useful applications of AI. In addition - by tailoring the application of AI to the participants of this design process, we further ensure that participants also get something out of the design activities.

Further planned research

What we plan for future research for the final delivery is to hopefully have another interview, this time with the elderly user (EU) to collect more information on her thoughts and experiences on communicating with family members, her preferred ways of communicating and her thoughts on AI and voice activated systems. Further on we would like to give the participant some experience interacting with AI. This is thought to be done with a simple dictation interface, speech synthesis, or other applications that can contribute to expanding the design space and serve as examples. Further we would like to supplement this with another interview concerning the participants experience with AI. Also, we're quite interested in having the participant interact with a potential prototype and using it as a tool for a longer period of time with some complementary tasks, so we get an understanding of EUs ability to use the prototype and potential learning curve with the prototype. The task would consist of sending one or more messages each day using the prototype. These messages could be sent to a family member by email/message where we could store the audio and text, and later analyze this data. This will identify potential challenges for EU in the interaction with the AI-infused system.

This is how we hope to proceed with further research, but we will take into account that changes may happen along the way considering how the participants will respond, how much they can take part, and the project's length.

Preliminary findings

In this section we present the initial outcomes from the interview with a family member (FM).

From the conversation with FM we identified aspects of communication between FM and EU that could be better, what motivates participation in their communication activities and also unwanted forms or aspects of communication.

Aspects/elements of communication that could be better

- The elderly user (EU) has access to two devices. One feature phone and one ipad. She can use the phone unassisted but need assistance to operate the ipad.
- The feature phone does not have the capability to recieve or send images which hinder EU to get updates from her immediate family utilizing rich media.
- FM states that this is a feature that EU would be likely to appreciate.

Unwanted forms or aspects of communication

- FM and EU frequently communicate with sms.
- FM explains that she avoids making phone calls to EU because they take up too much time. The effectiveness and possibilities for more frequent communication is mentioned as an upside of using text messages.

Motivation/value to participant in communication activities:

- FM would much rather make an appointment by sms and have a longer conversation face to face with EU.
- Furthermore FM emphasizes that text messaging would be beneficial for EU concerning communication with the grandchildren
- FM explains that EU probably would like to write longer messages, but due to the keypad /interface on her feature phone it is tiresome for her.

Appendix 1 - Chabot design

What did we do?

We first had a group meeting where we decided which of the chatbot tools we wanted to use. We then had to decide on the purpose of our chatbot and how much time we wanted to spend on learning the tools. Harald, one of the group members showed how we could manage the task in the different tools available (Chattron, Chatfuel, and DialogFlow), and the group agreed on using DialogFlow because it better supports word recognition. In addition, we liked the user interface of DialogFlow better. One group member suggested making a chatbot about Stian and his handball team, the rest of the group agreed. The purpose of the chatbot is to give the user info about Stian and his endeavors playing and practicing handball.

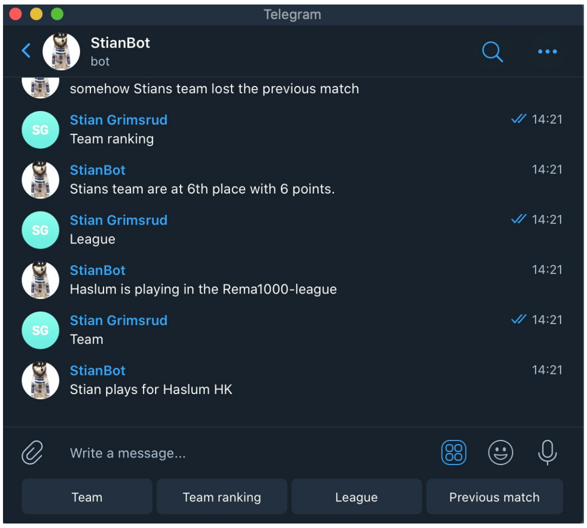
How did we do it?

Before we started with the implementation of the chatbot, we had to identify some key tasks that the chatbot should handle. We identified three key tasks that the chatbot should handle in a good manner, and these were;

- give the user insights to Stians matches and give links to Stians stats
- give links to upcoming matches on the streaming platform sumo
- give links to Stians team (Haslum) stats

The first task is for the user and will give insight into Stian’s matches, like the results and statistics for Stian in the given match. This will give the user options to ask the chatbot, named StianBot, about his stats from the previous matches. The next key task we identified was the chatbot to give suggestions on where to watch Stians matches, etc. TV2 Sport 2 or handballtv.com. This so the user easily can access Stians matches. The last key task for StianBot was to give the user statistics for Stians team (Haslum HK), so the user could see how the team performed and see where in the table they are.

After we identified these key tasks, we started with the implementation of the chatbot. We then started with making entities and intents, so the chatbot would have some “keywords” it looked for, for example, match, stream, number, etc. We also used the default intents; Welcome intent and Default Callback, so we didn't need to implement these by ourselves. After some testing of the chatbot, we identified a need for buzzwords, because of the lack of knowledge of what a user could ask the chatbot. We then choose four buzzwords the user could interact with to get an understanding of what the chatbot could answer. We then found appropriate places to “set” these buzzwords, for example at the beginning of the conversation and when the chatbot didn't get what the user asked for. After making the chatbot we wanted to integrate it into a conversation app (like Telegram and Messenger) and choose to integrate it in Telegram, because of the encryption of the messages.



Reflections on the outcome

While we were implementing the chatbot, all of the group members had the same experience in the making of the conversation. We had to design the chatbot for specific conversations and had to guess what a user actually would ask the bot. This is something we thought was a bit hard, to clearly identify what types of questions and words a user would use in the conversation. We then identified we needed to design the answers, in the same manner as we had to design the questions.

When thinking about the above, we also found out that a user could interact with the chatbot in a matter we didn't think about. This could result in a conversation with the chatbot that didn't give any answers to the users' questions. So we had to think about the whole conversation, so a user could get all the answers it needed.

Appendix 2 - Machine learning, MovieChatbot

The second task in Module 2; "Design of interaction with AI" was to take an existing chatbot program and test it with different learning attributes and compare these. The goal of the assignment was to give us a deeper understanding of how AI and chatbots work internally.

We first began with the code we were given and tested this. Then we changed the ML model with 10 layers of neurons and a pyramid of these(128, 256, 512, 256, 128). That didn't help the model, it actually got dumber. Then we tested with 5 layers of neurons with the same pyramid structure. It helped, but not as much as we wanted. From there we changed the structure of the neurons and added one layer at a time until we got to 8 layers and linear structure.

We tested different combinations of training iterations and layers neurons and different combinations of neurons to see the difference in the outcome.

```
0.24798004
0.12711085
0.08269722
0.063733995
0.05397605
0.048242543
0.04454889
0.04200633
0.04016741
0.038785547
0.037715115
0.0368651
0.036176357
0.035608448
0.035133433
0.03473102
0.034386307
0.034088142
0.03382811
0.03359948
0.03339719
0.03321715
0.03305599
0.03291079
```

```
self.fc1 = nn.Linear(128,256)
self.fc2 = nn.Linear(256, 512)
self.fc3 = nn.Linear(512, 1024)
self.fc4 = nn.Linear(1024, 2048)
self.fc5 = nn.Linear(2048, 4096)
self.fc6 = nn.Linear(4096, 8192)
self.fc7 = nn.Linear(8192, 16384)
self.fc8 = nn.Linear(16384, num_classes)
```

This is our result after trying out eight layers of neurons and making these layers linear, with 600 training iterations. Its 3.2% margin of error.

```
n_steps = 1200
for i in range(n_steps):
    y_pred_train = n(x_train)
    loss_train = loss_fn(y_pred_train,y_train)
    optimizer.zero_grad()
    loss_train.backward()
    optimizer.step()
    if (i % 25) == 0:
        print(loss_train.detach().numpy())
```

```
0.054432813
0.04851372
0.044721466
0.04212207
0.040248066
0.03884355
0.037757687
0.036897074
0.036200695
0.03562721
0.035148077
0.03474246
0.034395278
0.034095276
0.03383367
0.03360397
0.033400718
0.033219766
0.033057917
0.032912336
0.03278069
0.03266121
0.0325523
0.032452863
0.032361384
0.032277264
0.032199632
0.032127835
0.032061145
0.031999145
0.031941302
0.031887244
```

```
self.fc1 = nn.Linear(128,256)
self.fc2 = nn.Linear(256, 512)
self.fc3 = nn.Linear(512, 1024)
self.fc4 = nn.Linear(1024, 2048)
self.fc5 = nn.Linear(2048, 4096)
self.fc6 = nn.Linear(4096, 8192)
self.fc7 = nn.Linear(8192, 16384)
self.fc8 = nn.Linear(16384, num_classes)
```

These are the results with the same neurons, but we doubled the training iterations. We got a little bit less of a margin of error. But not as big as we hoped.

Our conclusion after trying out these different training iterations and layers of neurons are that we didn't see a clear connection between the margin of error and how well the chatbot's responses when we talked to it. For instance, we had a 3% margin of error, but the chatbot still repeated the same answers.

References

AI Design Guidelines (2019) G9 - Support efficient correction. Make it easy to edit, refine, or recover when the AI system is wrong.

Foley A. & Ferri B (2012). Technology for people, not disabilities: Ensuring access and inclusion. Retrieved from https://www.researchgate.net/publication/263752183_Technology_for_people_not_disabilities_Ensuring_access_and_inclusion

Guidelines for Human-AI Interaction
In CHI Con- ference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland Uk. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3290605.3300233

Lohr, S. (2018, February 9). Facial Recognition Is Accurate, if You're a White Guy. The New York Times. https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html

Luger, E., & Sellen, A. (2016). Like having a really bad PA: the gulf between user expectation and experience of conversational agents. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (pp. 5286-5297). ACM. (https://www.microsoft.com/en-us/research/wp-content/uploads/2016/08/p5286-luger.pdf)

Maartmann-Moe, Harald
https://www.duo.uio.no/handle/10852/70963

Schulz, T. (2020). Exploration of Moving Things in the Home. 136.
WHO (2020). World report on disability. Retrieved from https://www.who.int/disabilities/world_report/2011/report/en/

Boyd, C. (2018) Speech Recognition Technology: The Past, Present, and Future. Medium.com. Read: 22.10.2020 Retrieved from: https://medium.com/swlh/the-past-present-and-future-of-speech-recognition-technology-cf13c179aaf