# Midterm report - second delivery

## IN5480 - Fall 2020

*Group 5:*

*Mariann Gundegjerde, Snorre Ødegård, Thea Aksdal Nordgulen,*

*Barbro Årnes, Claudia Sikora and Linda Østerberg*

## Table of content

IN5480 - Midterm report
Group 5

# About us

We are a group of 6 students:
Mariann Gundegjerde, margunde@uio.no
Snorre Ødegård, snorreod@uio.no
Thea Aksdal Nordgulen, theaaano@uio.no
Barbro Årnes, barbrora@uio.no
Claudia Sikora, claudisi@uio.no
Linda Østerberg, lindaeo@utio.no

# Area of interest

Our area of interest could lead us in many directions. We want to research what the attitude towards AI is among users and how this could change based on the language and behaviour of the AI. Do users have expectations for how an AI should act and what the limitations of these actions are. The expectations of a user could be based on earlier experiences, things they have seen in the media or on TV. We are hoping to get a clearer view of the practice and knowledge among users, and the research done around this topic.

Furthermore, in our area of interest we are concerned with how an AI behaves, and whether this behavior can be modified to appear more human-like. What is a human-like behavior and how could this be translated to a chatbot or a physical robot? And again, what are users' experience and feelings toward this.

Due to the chosen area of interest, we wanted to focus the scope on an appropriate user group. We discussed who might be relevant to the chosen area, and decided that young adults / students would be a suitable user group as informants. Most young adults use chatting services on a daily basis and are familiar with how to communicate online with other humans. In addition to this, many young adults are familiar with the concept of chatbots, which might make it easier to collect valid data from the informants due to reduction of the hawthorne effect. Students are also quite an accessible group, as all the group members are students themselves.

# Background

People's social responses to interacting with technology has been of great interest to many. "The media equation" is a theory developed by the research of Clifford Nass and Byron Reeves. This theory claims that people respond socially to computers, the same way they would treat humans - the rules of human-human interaction apply to human-computer interaction (Reeves & Nass, 1996, p. 23). This also holds true for experienced computer users, such as IT-experts and the like - this is an innate reaction beyond our conscious control. Research has also shown that gender and ethnic stereotypes also apply to AI-systems (Nass & Moon, 2000, p. 81) respond to flattery and that the formulation of error messages affect how "friendly" they find the system to be.

This may be explained by evolutionary psychology (Nass & Gong, 2000, p.38). To our brains, there is no differentiation between a robot and a live being. Humans have complex cognitive systems dedicated to understanding speech and other forms of incoming communication, which leads to several implications in the design of human-AI systems. For instance, when one fails to be understood, people tend to hyperarticulate their speech. The implication of this with human-AI interaction is that if the AI in question learns from human input, this may lead to the AI learning from speech that is not natural to the person normally. This behavior may likely happen in text-based communication as well, with users simplifying their choice of words and grammar during moments of frustration when a chatbot fails to understand the users' communication.

Personality factors are also mentioned - people generally prefer both people and systems with personalities similar to their own, and have greater levels of trust towards systems they believe are similar to themselves (Nass & Moon, 2000, p. 92). The solution may not be so straightforward as to simply letting users choose a "personality profile", since many people do not necessarily know themselves very well, and are often not aware of what their personality is like. It is important to note that social responses towards technology are more likely to happen the more human-like characteristics the technology possesses (Nass & Moon, 2000, p. 97). This can inform the design of chatbots, depending on what kinds of interactions one wants the user to have - a chatbot meant to assist in tasks and a conversational chatbot for home entertainment may benefit from using this principle in different manners.

“Presence” is a topic of interest to many researchers, referring to “a psychological state in which virtual (para-authentic or artificial) objects are experienced as actual objects in either sensory or nonsensory ways” (Lee, 2004, p. 37). According to Lee (2004), there are three different types of presence: physical, social, and self presence. A user experiences physical presence when said user does not recognize the artificial nature of a virtual object. Social presence is experienced when a user interacts with a virtual social agent and perceives them as real. Self presence occurs when a user interacts with a virtual representation of themselves within a virtual world. Presence is an important concept when developing chatbots, because the feelings of presence affects the psychological fidelity of a system (Sharples & Wilson, 2015, p.208). Psychological fidelity may affect whether or not the chatbot is taken seriously, which in turn affects how the user behaves with the chatbot.

Lombard and Ditton (1997) discuss the concept of presence in greater depth, emphasizing the diverse nature of technologies and how they all affect presence in different ways. Physiological factors and psychological factors are investigated, anything from motion sickness to how increased presence can enhance enjoyment, increase “believability” in a medium and improve performance on various types of tasks and enhance learnability of training.

Since we live in a society where interacting with machines and AI is becoming part of our everyday life, the emerging field of AI and machine-behaviour has drawn an increased interest the past few years. This is a field not only including machine and data science, but also aspects of sociology, ethnography and phycology. Reflecting on questions such as, what trust do we put in AI, what do we expect and how does the interaction and use affect the society and the people as individuals? (Rahwan et al. 2019).

Research has been done on humans' emotional response when encountering non-human technologies. Shank et al. writes about how humans “ emotionally process the gap between nonhuman technologies and having a mind, essentially feeling our way to machine minds.” (2019). In other words the interaction between human and machine could be seen as equally complex as complex as the subjective feelings of the human that is interacting.

Developing an AI that responds to human nuances and manners really understanding the intent behind their words, are much more demanding and time consuming than developing an AI able to interact in a litterall and straightforward way. Communicating using this type of direct language would often be perceived as rude if the dialog were performed by two humans, while if it was performed with an AI it might be a question of effectiveness. As ISO 9241-11 states, usability concerns "Degree to which a product or system can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use."

With this in mind we wish to further explore users expectations and perception of AIs personality related to efficiency and the context of use.

## Questions

During our second iteration of the project we have changed one of our research questions. The reason for this was based on our findings from the two interviews we have conducted. The previous question was "Is it possible to make an AI more human in the way it acts?". The participants expressed that they saw no need for making an AI-system act in a human-like way, and this made us change our question to "Do users expect/prefer AI to behave more like humans?".

*"What are the expectations of how an AI should behave?"*
With a rapidly increasing amount of different AI's made for different purposes the expectation for what a given system is able to do, varies depending on the users earlier interactions with similar systems. We want to examine what the expectations of users are and how interacting with AI-systems behaving differently than what they expect affects them.

*"Do users expect/prefere AI to behave more like humans?"*
We are wondering if users expect or prefer AI to behave like humans in the terms of creating an illusion of the AI having a personality and opinions. Does it for example make it more approachable for new users or is it distracting? Things we could test could be if the language of the AI was more human-like, for examples using dialects and giving the AI more of a unique personality. Here we would also have to examine what makes an AI seem more

human and what differentiates for example a conversation with a human and a conversation with a chatbot.

# Methods

The questions we ask in this paper can be summarized into two main concerns: personality and human-robot interaction. We think a combination of literature review, interviews and user tests will be the best approach, considering the scope, theme and complexity of this study. And at this point of the project we have done a few interviews and researched some literature.

## Literature review

In robotics and AI, personality refers to a robot's ability to interact with users emotionally and logically (Rouse, 2017). They not only have to act logically, but do it with "style" - whether through voice (Nass & Gong, 2000, p. 38), variations in emotion (Hu et al., 2013, p. 7), and a multitude of other channels for conveying personality. The literature review will serve to make us acquainted with past researchers' attempts at giving robots a personality, and people's reactions when interacting with this kind of robot. From this literature review, we may then identify and specify areas of interest. While our paper will be based on some user test we will conduct ourselves, most of our base knowledge will come from the literature we find and this will guide us in how to make decisions and how to find answers to our research questions.

## Interview

Further we would need to examine how users actually experience interacting with different kinds of AI. The best way to examine this is through interviews, and at this point in the project we have interviewed 2 people. Interviews are an effective way of gathering rich information about users, which may include information about unexpected topics (Lazar et al., 2017, p. 188). Our interview tried to find out what the users expectations were when using an AI, and why they had these expectations. We also had in mind that different users have different knowledge about AI and have different amounts of experience with using AI, so we made sure to customize the questions based on this.

### Further testing

We would like to conduct user tests as well, but there are restrictions which need to be considered due to COVID-19 and we are hopeful that we can adapt this method so that it will be possible to implement. A possible user test we could do is testing two chatbots who have the same goal, but speak with different types of language. One could have a normal-type of language and one could contain all of the unexpected ways of communicating. Both of these would be based on what the users from the interview depicts as normal and unexpected. There are many possibilities here, but the benefits of conducting a user test would be to evaluate how users actually experience and behave when testing an AI.

### Prototypes

For the next phase of the project we will produce prototypes based on the information from the interviews and our literature research. We are planning on making two different chatbot-prototypes, as described above, where one will be based on the expectations presented in the interviews and one will be different and try to question and disrupt our users expectations. Hopefully these prototypes will produce an interesting opportunity to test with users and gain more knowledge about people's expectations and how these can be challenged.

## Findings

During our first round of interviews, we focused on getting knowledge about different types of AI users, and their experience with using AI. Our questions were open, and let the participant elaborate on their own experiences and thoughts around AI. We saw it as important to remember that not all users have the same knowledge as us, and it was therefore important not to lead the participants in directions we found interesting, but rather let them lead the conversation.

As of now we have just conducted 2 interviews. This is not enough to start drawing conclusions of peoples thoughts and expectations towards AI, but it has at least given us some interesting insights in some of the things we wanted to explore.

The participants of the interviews so far have been what we call novice users, meaning people that may use AI from time to time but who do not currently use it in a professional setting or have any sort of education that would give them any special insights into AI´s. Examples of AI´s centered more towards novice users that the participants mentioned that they had used during the interviews are Siri, Google Home and a couple of different chatbots.

It was clear that the users had different expectations of different types of AI. One participant said that AI usually have little to no emotions and gives you the answer to what you're looking for and nothing more. While another participant said he had different expectations for an AI meant for solving work tasks and one meant for consumer use.

When asked if they were afraid of getting replaced in their occupation in the near future, both of the participants answered that they did not think that AI would be able to replace them. Both stated that though AI can help effectivise a lot of their workload connected to things such as research and filling out paperwork, it would not be able to replace the human element needed in their occupation. In their opinion most AI-systems lack the skills to replace a human-to-human interaction, which is an important part of many occupations.

Perhaps the most interesting finding from our interviews was when we asked our participants how they would go about making an AI more human-like. To our surprise all of the participants answered they did not see a reason why there was a need for this. For example, why should a chatbot communicate with us in a more human way, when all they wanted was for the chatbots to be quick and effective, and do the task at hand. The participants stated that having to talk with a chatbot in the same way you would talk with a human would not be effective.

# References

Lee, K. M. (2004). Presence, explicated. Communication Theory, 14, 27-50.

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. Journal of computer-mediated communication, 3(2), JCMC321.

Nass, C., & Gong, L. (2000). Speech interfaces from an evolutionary perspective. Communications of the ACM, 43(9), 36-43.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. Journal of social issues, 56(1), 81-103.

Reeves, B., & Nass, C. I. (1996). The media equation: How people treat computers, television, and new media like real people and places. Cambridge university press.

Sharples, S., & Wilson, J. R. (2015). Evaluation of human work. CRC Press.

Rouse, Margaret. (2017, March). Robotic personality. Retrieved from https://whatis.techtarget.com/definition/robotic-personality

Lazar, J., Feng, J. H., & Hochheiser, H. (2017). Research methods in human-computer interaction. Morgan Kaufmann.

Hu, X., Xie, L., Liu, X., & Wang, Z. (2013). Emotion expression of robot with personality. Mathematical Problems in Engineering, 2013.

Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, JF., Breazeal, C., Crandall, J., Christakis, N., Couzin, I., Jackson, M., Jennings, N., Kamar, E., Kloumann, I., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D., Pentland, A., Roberts, M., Shariff, A., Tenenbaum, J., Wellman, M. (2019). Machine behaviour. Nature. Vol.568 (7753), p.477-486

International Organization for Standardization. (1998). *Ergonomics of human-system interaction.* Part 11, *Usability: Definitions and concepts.* ISO 9241-11. New York: American National Standards Institute.

Shank, B., Graves, C., Gott, A., Gamez, P., Rodriguez, S. (2019) Feeling our way to machine minds: People's emotions when perceiving mind in artificial intelligence. Computers in Human Behavior, Volume 98, Pages 256-266,

Goodwin, Dr Morten. “Interacting with Artificial Intelligence” (2020). Centre for Artificial Intelligence Research. University of Oslo.

# Appendix 1: Chatbot design task

An assignment during module two was to create a chatbot. It was a two weeks task and the group decided on creating a bot for dinner-inspiration. We ended up making a chatbot we called “Middagshjelp (Dinner help)”.

## The process

As mentioned, the purpose of the chatbot was to provide inspiration for choosing what to eat for dinner and the response from the chatbot was thought to be based on the user's personal preferences. After agreeing on the purpose of the chatbot the group performed a joint brainstorming concerning basic functionality necessary for a first prototype. The key tasks, like greeting the user, explaining functionality, asking for preferences and recommending a dish, were identified and some basic flows were sketched. At first all group members had a try at working with their own chatbot to get an understanding and some inspiration, for thereafter discussing further implementations to be made on the prototype chatbot. Amongst available tools for developing chatbots, different group members tested Chatteron and Dialogflow. The group decided on Dialogflow since we found it to have a more intuitive developing-UI and that it did not require connecting the bot to a Facebook page. The chatbot was tested and altered during development to improve user interaction. Some of the adjustments made were dividing tasks into separate steps and improving the language.

## Reflection

We found the assignment interesting and ended up with a chatbot we were happy about. One frustrating part of the project was to cooperate on making the chatbot. Because of the current situation with the pandemic, many of the group members work from home and are used to having the ability to cooperate on tasks online. There were minimal options for cooperating on the chatbot-task, since everyone had to work on their chatbot individually. So we ended up letting just a few members of the group finish the chatbot, and then share the result with the rest of the group to give feedback. This worked out well in the end, but it would have been more instructive and educational if more of the group members could have worked more closely together to create the chatbot.

Lastly we reflected on the limitations we had when making our chatbot. Our finished product was much easier than what we first prototyped and discussed. The chatbot only had one path, and a very predetermined path. This of course is due to our abilities, and the abilities of Dialogflow. Although this was frustrating at times, it gave us the basic knowledge of what goes into developing a working chatbot.

# Appendix 2: Machine learning task

In this appendix, we explored different capabilities of the python-script 'MovieChatbot.py' and the accompanying text-file 'movie_lines.txt'. The aim of this assignment was to customize our own model of the Chatbot-script, by appending or subtracting the number of neurons, as well as changing the input-text. We did this by editing the script in JupyterLab provided by Cair-hub. Throughout this assignment, we got to explore the different functionalities of machine learning by using the programming language Python and knowledge conveyed in lecture on the subject "Interacting with Artificial Intelligence" (Goodwin, 2020).

## The process

We started the process by preparing the JupyterLab with copy and paste from the assignment description. The next step in this process was to understand how the script and the text file was connected and what kind of interaction we could expect when executing 'MovieChatbot.py'. By exploring the default version of the program, we got to understand what the different numbers presented to us in the terminal meant, which was the starting point of getting to know the algorithm. We noted the results of the first execution and discussed expectations of how changed in the script would affect the flow of the system.

The very first change we made was editing the amount of neurons in the function 1 (fc1) and function 2 (fc2). We decided that we would start by appending even more neurons to the functions, and discovered that the loss decreased from 0.23 to 0.20. In the lecture, we learned the loss should aim for being as close to zero as possible and accordingly we concluded that the decrease in neurons increased the loss of the algorithm. We increased the number of neurons by adjusting the algorithm multiple times, where the outcome confirmed our

assumption. An increase in the number of neurons would make the algorithm gain more knowledge through each iteration.

An interesting aspect of the discovery of the increase in loss, was how adding more neurons when reaching a significant number (in our case 100 000) made a smaller impact on the learning of the algorithm. There was moderately variety in how the chatbot responded as well through the increase, but this observation might be biased by us not knowing what the algorithm actually does. We wanted to explore the concept of overfitting, but as the amount of neurons was increased we did not see the loss increasing again after decreasing. There is probably a way to provoke overfitting in this particular case, but we did not figure out how to make this happen. Maybe by adding more layers?

Further we tried to experiment by decreasing the amount of neurons. The findings of the process was how having quite few neurons would affect the learning curve of the algorithm by making it vastly shallow. The first loss number would be similar to the results of the algorithm handling many different neurons, but the number would not increase throughout the iterations of the system. By discussing this change and through knowledge about machine learning, we assume that this shallow learning curve is caused by fewer progrations and paths through the network of layers. Though the response of the chatbot was not perfect in the first place (with many neurons), the appearance of the chatbot was even more off and random now than ever.

## Reflection

In this assignment, we addressed the challenge of trying to configure a system which deals with a machine learning algorithm. The main challenge of the exploration of the system was understanding what was really happening. We tried to understand how the algorithm was learning by revisiting lectures, watching videos on the subject and doing research online. We felt like we got to understand the concept of machine learning, but not how it was done in practice. This is a challenge mentioned in the curriculum, where a significant amount of research on this field is based on designers not really understanding the capabilities of artificial intelligence and machine learning. Thus we understood more by doing research, but it was challenging to understand exactly what was happening without the technical competence of machine learning.