Course name: Human Computer Interaction

Assignment no.: 8

Title: Human-AI Interaction and Analysis of ChatGPT with Design Guidelines

Subtopic 1. Human-AI Interaction

AI is short for Artificial Intelligence, which refers to the ability of computer systems to mimic or perform intelligent human tasks. AI aims to give computer programs the ability to perform tasks such as thinking, learning, problem solving, and judgment. Recently, the field of artificial intelligence has been making rapid progress, especially with the popularity of ChatGPT on OpenAI. As a result, many companies have been releasing their own AI-enabled systems, creating a buzz. Of interest to me is Nvidia, which has released a generative AI for visual applications called Picasso. As AI develops, new jobs will arise and many will disappear. In the future, people will need to be able to utilize AI, so we need to know the difference between basic human-computer interaction and human-AI interaction.

In human-computer interactions, the user is primarily in control and makes decisions, whereas in human-AI interactions, the AI system has varying degrees of autonomy and decision-making power. Users can interact with the AI system to provide input or guidance, but the AI system itself could make choices or take movements independently. The next difference is that AI can adapt and learn. AI systems have the ability to learn and adapt over time. They can continuously analyze data to improve their performance. Human-AI interactions should take these characteristics of AI systems into account and provide mechanisms for users to influence and guide the learning process. \

As a unique aspect of human-AI interaction, humans must trust the AI. Of course, no AI is perfect, but without trust, successful interactions are impossible. Another unique aspect is that humans need to understand the AI. Rather than understanding how the code is put together to make it work, AI is analyzing natural language processing and gestures as they are put together. If we are aware of this and understand the user input and context, and provide input accordingly, it will be a successful interaction. Additionally, while AI learns and grows, we need to keep it from stepping outside of its ethical framework. For products that allow users

to interact directly with the AI system, there is ChatGPT. This will be introduced in more detail in the next subtopic. For input interfaces in Human-AI interaction, there are voice, text, camera, and touch based. For outputs, we have mainly visual and auditory data.

There are many things to consider when designing human-AI interactions. For starters, it needs to be easy for users to use. It's a big mistake to assume that something built with AI will be used by a computer science major. It needs to be easy for someone who is not familiar with the AI to use it. It also needs to be able to take feedback and make changes. No AI is perfect, and interaction is only meaningful if it grows from feedback. Interactions need to be sensitive to people. Ethical and religious issues should be avoided, and privacy should be respected. By considering all of these factors, you can design a more intuitive, trustworthy, and effective interface for users to interact with your AI system.

In conclusion, AI is a technology that makes computers smart and capable of tasks like thinking and problem-solving. When interacting with AI, we need to guide and trust it, understand how it analyzes language and gestures, and make sure it stays within ethical boundaries. Interfaces should be easy to use for everyone, and we should listen to feedback and make improvements. Respecting privacy and considering ethical concerns is also important. By doing these things, we can create interfaces that are easy, trustworthy, and effective for people to interact with AI.

Subtopic 2. Analysis of ChatGPT with the design guidelines

ChatGPT is an AI chatbot created by OpenAI and is one of the Large Language Models. It is called ChatGPT by combining the Chat and initials of "Generative Pre-Trained Transformer". Currently, GPT-3.5 is free and GPT-4 is available for a fee. We released GPT-1 in 2018 and have been steadily upgrading it, with GPT-5 coming later this year. As January this year, the company claims to have over 100 million monthly users. This is said to be the fastest for any social network in history. Let's analyze ChatGPT considering each of Microsoft's 18 guidelines.

1. Make clear what the system can do.

Make it clear to users that ChatGPT can do things like answer questions or perform tasks on a variety of topics. When I entered to homepage, it introduced examples, capabilities, and limitations to understand how to use it.

2. Make clear how well the system can do what it can do.

Clearly communicate to users that while ChatGPT can work well in common conversational situations, it is not perfect. The performance of the model may vary depending on the input given. Not always satisfied.

3. Time services based on context.

ChatGPT can provide time-related services that take into account the given context. For example, the question "Tell me the weather today" can be answered with current weather information.

4. Show contextually relevant information.

ChatGPT can provide relevant information based on the context of the conversation. It may answer questions or provide additional information about something mentioned in a previous conversation or related information. However, it sometimes tends to rely too heavily on outdated information.

5. Match relevant social norms.

I'm not sure about this one, but it's possible if we tell the system that it's consistent with social norms.

6. Mitigate social biases.

ChatGPT is trained to minimize social bias. Prevent hateful or discriminatory remarks from being made through the Moderation API.

7. Support efficient invocation.

ChatGPT supports fast request processing to enable efficient calls. It interacts with the interlocutor in real time and generally provides fast response times. However, this also varies by language and can often take a long time.

8. Support efficient dismissal.

Just turn off the window and you're done. It's efficient because the information is still there when you close it.

9. Support efficient correction.

It provides corrections. You can tell it that it's wrong. But it's not efficient: sometimes I forget what I corrected and have data from before.

10. Scope services when in doubt.

I'm not sure about this, but in GPT-3.5 it doesn't seem to be limiting the scope, as it often misrepresents information it doesn't know.

11. Make clear why the system did what it did.

ChatGPT can clearly explain why a model generated a particular answer. If you ask further, it will let you know. Provides supporting papers or data. But again, most of this didn't exist.

12. Remember recent interactions.

ChatGPT can remember recent conversational interactions. If you keep having the same chat, it will remember.

13. Learn from user behavior.

In the short term, it learns from what you teach it in chat and performs well. In the long term, as the model is upgraded, it learns to fill in the gaps.

14. Update and adapt cautiously.

Based on the differences between GPT-3.5 and GPT-4, it seems that updates and adaptations are done carefully.

15. Encourage granular feedback.

ChatGPT encourages specific feedback from users. But the feedback doesn't come in the way we want it to.

16. Convey the consequences of user actions.

ChatGPT can communicate the consequences of a user's actions. For example, it can explain what outcome a user can expect if they request a specific action or inform them about the potential impact of a particular behavior. Of course, it's a prediction, and sometimes it's wrong.

17. Provide global controls.

We can't, and maybe OpenAI can, but we do have control in the chat room.

18. Notify users about changes.

We make it easy for users to understand what's changed with each update with performance comparisons, videos, and articles.