

# **NostalgAI: Using machine learning to support caregivers in storytelling with dementia patients at care homes**

Ahmed Gamal Mohamed Sayed Ahmed Ibrahim

Department of Human-Technology Interaction, Eindhoven University of Technology, a.ibrahim@student.tue.nl

Harry, HW, Wang

Department of Industrial Design, Eindhoven University of Technology, h.wang3@student.tue.nl

Leonardo, LCS, Chávez Sáenz

Department of Industrial Design, Eindhoven University of Technology, l.chavez.saenz@student.tue.nl

Vere, V, Vreeswijk

Department of Industrial Design, Eindhoven University of Technology, v.vreeswijk@student.tue.nl

This report discusses the development of a design solution that supports caregivers of people with dementia in social interactions to improve their quality of life. A prototype was made specifically using AI and machine learning. Our solution, NostalgAI, uses AI and machine learning to collect memories and analyzes if details are missing from future retellings. These details are used to provide caregivers with questions to increase social engagement. The prototype is tested on various inputs. The result is a robust tool that the caregiver can use without many issues or interference of the dementia patient. At the end, a multitude of potential improvements were found that can elevate the design for further developments.

CCS CONCEPTS • **Artificial intelligence** • **Machine learning** • **Interaction Design** • **Dementia** • **Human-Technology Interaction**

## **1 INTRODUCTION**

Advancements in health and medicine have contributed to an increase in life expectancy worldwide [22]. However, a concerning trend emerges as the global prevalence of dementia is projected to surge by 85%, soaring from the current 82 million individuals to an estimated 150 million by 2050 [8]. Notably, the likelihood of experiencing dementia rises with age, with a growing percentage of the elderly population being affected. In the absence of a cure, addressing dementia transforms into a risk management approach aimed at enhancing the quality of life for the elderly [6]. Dementia encompasses a variety of cognitive diseases and brain injuries. With Alzheimer's, being identified as the most prevalent form contributing between 50 and 75% of recorded cases [5, 18].

Research reveals that individuals suffering from dementia often contend with isolation, loneliness, and depressive feelings [10]. Miesen [14] emphasizes that those with dementia frequently encounter unfamiliar situations, triggering a strong need for attachment. This shows that inclusion, denoting the ability to engage fully with society, is crucial for individuals with dementia. To foster their participation in groups they can identify with and prevent withdrawal and apathy.

Lastly, identity, the ability to construct a narrative about one's life, is profoundly impacted by dementia. As cognitive decline occurs, identity is sometimes compromised due to both neurological and social-psychological factors [15]. Following the aforementioned information and in the absence of a cure, dealing with dementia turns into risk management to facilitate a higher quality of life for the elderly [6]. Where risk management pertains to promoting social interactions that allow people to feel a sense of attachment [13].

In this paper, the focus pertains to improving the mental state among dementia sufferers, by promoting social interactions. These social interactions are based on storytelling, encouraged by caregivers, to be performed by dementia sufferers. This would be achieved with the help of caregivers and machine learning (ML) models collaborating to talk about and record memories that dementia sufferers bring up in conversations. Caregivers would start a conversation with dementia sufferers regarding certain memories that have been pre-verified by family or friends. After this, the caregiver would start a conversation with the dementia sufferer based on those memories. Where the memory would then be transcribed using Whisper, an ML model that handles speech-to-text transcription [19]. The transcribed memory would then be compared to the true memory collected from the family and friends using a large language model (LLM), such as ChatGPT [16]. Where the LLM would find deviations regarding said memory, that can be in the form of missing or additional information. The caregivers would then use this information to start asking follow-up questions regarding this memory to encourage 2-way conversations.

This method helps dementia sufferers on 3 frontiers. First, it would solve the isolation and loneliness problems that they face. As it allows for easy conversation starters to be utilized by the caregivers, as they have access to certain conversational pretenses. Second, it allows for their identity to surface back up again as they recall certain memories in their lives. Lastly, it could be used as an extra method of tracking memory and cognitive decline. By using an LLM the frequency of missing or incorrect information can be recorded. This would then allow for accurate memory decline metrics to be created and tracked, based on the accuracy of the LLM used.

## **2 RELATED WORKS**

### **2.1 Use of Whisper in Healthcare**

In the study of Spiller [21], the transcription of audio files has historically posed challenges due to its labor-intensive and error-prone nature. Addressing this issue, recent advancements in speech-to-text technology, exemplified by Whisper [19], offer a promising solution. The presented work outlines a comprehensive tutorial on implementing a transcription pipeline using Whisper for psychology, psychiatry, and neuroscience research. This pipeline covers system setup, recording, preprocessing, transcribing, and post-processing of audio data. A practical example in a Python environment demonstrates the effective application of Whisper, showcasing its capability to transcribe a 1-minute audio file with minimal inconsistencies. Despite certain limitations, the integration of Whisper for transcription in mental health research stands out as a viable strategy to significantly reduce the time-intensive nature of transcription work, thereby facilitating the analysis of audio data. This work empowers researchers to make informed decisions about incorporating AI-driven transcription into their methodologies and harnessing the full potential of audio data in their studies.

The study of Vijayshri [12] explores the efficacy of OpenAI's Whisper module in automatic speech recognition (ASR) for detecting chemical word entities related to ovarian cancer from human speech. Recognizing the critical importance of early detection in the successful treatment of ovarian cancer, the proposed ASR system relies on deep learning models proficient in deciphering complex speech patterns and distinguishing between various chemical terms associated with the disease. Beyond mere recognition, the study demonstrates the practical application of the detected chemical entities in web

content search and retrieval, providing a valuable tool for uncovering relevant information related to ovarian cancer. This research not only emphasizes the potential of ASR technology for early detection and accurate identification of disease-related chemical entities but also paves the way for the development of intelligent systems dedicated to disease diagnosis and treatment.

## **2.2 ChatGPT in in Healthcare**

In the study of Cheng [7], ChatGPT [16] is currently being used in practice in psychiatry, where it has the power to support psychiatrists in routine tasks, streamline clinician-patient communication, improve academic outcomes, and facilitate program design and research analysis. Looking to the future, GPT technology empathy, emotion recognition, personality assessment, and detection of mental health warning signs will also play a role.

## **2.3 Dementia care home AI technology and usage**

There are also other forms of AI technology being used within the healthcare industry. As seen in the study of Battineni [4], the aging population in industrialized Western societies has led to an increased incidence of non-communicable diseases, including neurodegenerative disorders and adult-onset dementia. Timely and accurate diagnosis is crucial for effective interventions, and conventional magnetic resonance imaging (MRI) is commonly used for this purpose. Integrating artificial intelligence (AI) with MRI has shown promise in enhancing diagnostic accuracy for different types of dementia.

This study systematically reviews AI algorithms for the early detection of adult-onset dementia disorders and discusses their performance metrics. A search across PubMed, Web of Science, and Scopus databases yielded 29 studies meeting quality criteria. The most common adult-onset dementia disorders were Alzheimer's disease and vascular dementia. AI techniques combined with MRI demonstrated increased diagnostic accuracy. These findings emphasize the potential of associating AI with conventional MRI techniques for precise and early diagnosis of dementia disorders in old age.

# **3 DESIGN SOLUTION**

## **3.1 User experience design**

The design's main purpose is to improve the mental state of dementia sufferers through AI. The final design is an application (Figure 1). The benefit of an application is that the caregiver can easily use it anywhere using any tablet or computer.

Using it, the user, in this case, the caregiver, can open the application and look in the database for stories that the current patient has stored. Then, the caregiver can press the record button before the patient tells their story. In this way, the patient does not have to interact with the technology and can focus on the storytelling instead. Once the story is done, the caregiver uploads the story at the press of a button. Once the button is pressed, the application uses speech-to-text AI to turn the story into text. Then, the text is parsed through an LLM comparing the current story to previous recounts of said story. The LLM can recognize the story that is being told and will generate a response containing discrepancies and new additions to the story. If it is a new story, the caregiver can add it to the existing stories. The stories are stored in a database. If the caregiver needs to, they can edit the stories later. Using the discrepancies given by the application, the caregiver can continue the story by asking further about the details that were missing. This should result in a fluid conversation where the patient is stimulated to recount memories.



Figure 1 Application design for the algorithm (detailed version in Appendix 8.3)

### 3.2 Specification of the program

A diagram is created to show the workflow of our program (Figure 2). First, the story of the dementia sufferer's memory is recorded as an audio file and stored in a database. Subsequently, the audio file is uploaded, and Whisper is then used to convert the audio from speech to text. When the audio is converted, the chosen LLM, namely GPT-3.5-turbo, is prompted to transfer the text. According to OpenAI, the GPT-3.5 model can understand and generate natural language or code and

has been optimized for chatting [17]. Finally, in the app, GPT-3.5 will identify errors and omissions by checking the text against an associated external database.

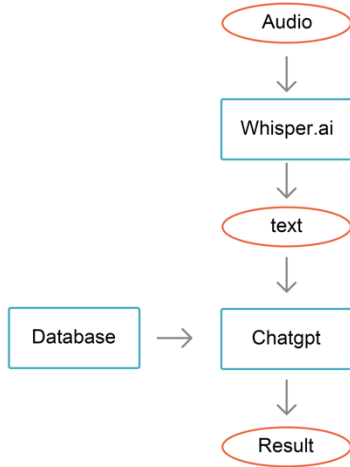


Figure 2 Workflow diagram of the application.

### 3.3 Realization of the program

To realize the design, we had to set boundaries to the functionality of the code. The most important part was to test the machine learning parts integrated in the project. This meant that the realization of databases was excluded and instead emulated through a Wizard-of-Oz methodology. Instead of using a database, the saved story would be hardcoded within the prompt of the LLM. This will give the same results as long as the inputs stay loosely within the boundaries of the hardcoded story. The final design of the prototype was realized in Python 3.11. Python was chosen as it is widely used as the basis for many existing AI tools. The chosen packages included, namely Whisper, OpenAI and Gradio, are also made for Python.

The program works as follows. Gradio is an open-source Python package that allows researchers to rapidly generate a visual interface for their ML models [1]. In this project, Gradio is used for web hosting and the interface. Once the user goes to the generated URL, the application will be visible to them. There is a button to record audio. Once clicked, the user can tell their story. Once you stop the audio recording, the user can listen back to the recording and submit it if they are satisfied. Once submitted, the application will first use Whisper to process and transcribe the recording locally to text. Afterward, GPT-3.5 will generate a response based on existing stories that are stored in the code. After GPT-3.5 is done, the results will be displayed in the corresponding output boxes. The view of the application before and after submission can be seen in Figure 3.

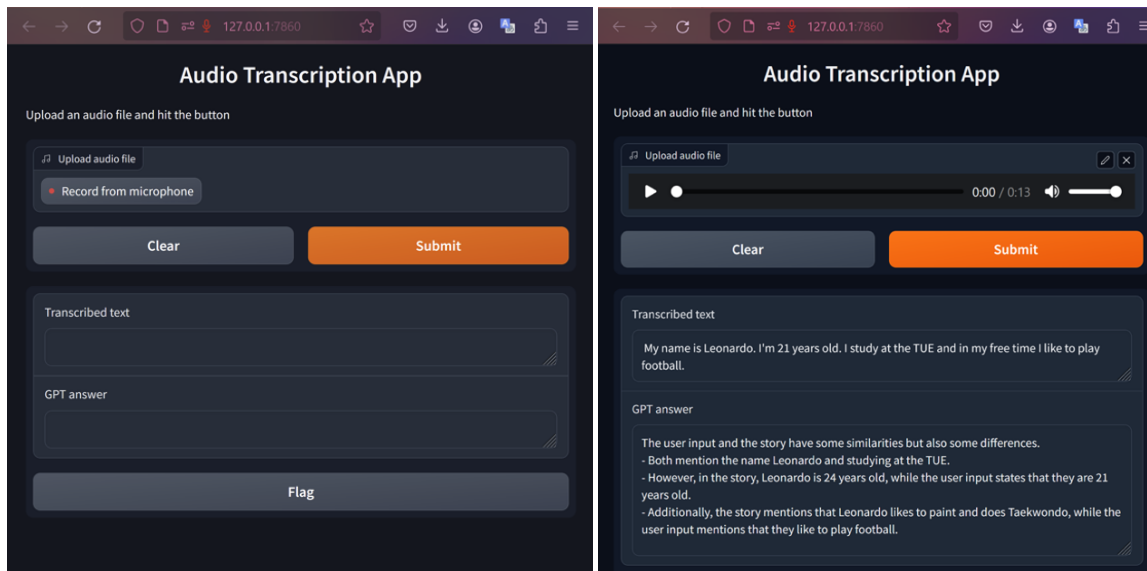


Figure 3 The application before and after submitting a recording.

### 3.3.1 Whisper

As mentioned before, a speech-to-text model was required to convert the conversational spoken words of dementia sufferers. Due to this requirement, Whisper was then chosen as the ML model of choice. Whisper is a general-purpose Automatic Speech Recognition (ASR) model trained on a diverse variety of data sets. Allowing it to perform multilingual speech recognition, speech translation as well as language identification [19]. Whisper uses a transformer-based sequence-to-sequence model based on a neural network architecture [23]. The difference between this model and other neural networks is the addition of the attention element for both the encoder and decoder elements. This allows for the model to weigh different parts of the input and output so that it can focus on the relevant information. Making it ideal for aligning the relevant audio features with the corresponding words in the transcription.

### 3.3.2 Large Language Model

The Large Language Model or LLM for short is a machine learning model that aims to predict and generate plausible language [9]. In the case of the design, it is used for comparing blocks of text representing stories. As LLM's are catered to generating text based existing text, using an LLM seemed quite fitting. In this project, the API GPT-3.5-turbo from OpenAI was used. GPT-3.5-turbo can be programmed to get a description from the system which will determine the way the AI generates its output. The designed code uses this feature to limit the output and to understand what the stored stories are. For testing purposes, the following system description was used (X was replaced with made-up data):

"Compare the user input compared to the following story:

Story 1: My name is X, and I am X years old. When I was X years old, I went to X and did X

Story 2: When I was X years old, I went to X and did X together with X and Y

Story X: ..."

The system description has been formulated through iterative testing with different stories. To fine-tune the application, the application was tested extensively using different prompts. All these prompts were written to have distinctive stories and structures. Based on the response, the GPT-3.5 system prompt was finetuned to give better results.

### 3.3.3 Interface

Gradio was used as the interface to be used by the average caretaker. The package offers a variety of possibilities in terms of input and output. Gradio requires a function to be given to know where to put the input into and where to take the output from. Gradio is also able to receive multiple inputs and multiple outputs. In this project, voice recording was used as an input and two text fields were used as outputs. Alternatively, there was the possibility to use file uploads for the audio files, but the easy use of the recording button without having to access anything outside the app, made the recording button our preference. The appearance of the standard Gradio interface using the required elements for NostalAI can be found in Figure 3.

### 3.3.4 Full integration

Combining the three tools was done in multiple steps. First, a working version of code with Whisper was made, with the output printed in the console. Then Gradio was integrated to make the two tools work together. Once this was done, GPT3.5-turbo was next. It was made to work on its own using a generated token, followed by integration with Gradio. Finally, everything was put together. This resulted in a robust application using compact working code. The final code can be seen in appendix 8.2.

## 4 RESULTS

Overall, the application seems reliable, being able to perform all the tasks it needs. The UI is scalable and available on all devices with internet access, and it works in multiple languages. The LLM with the resulting system input worked adequately with comparing stories, as this is also a form of looking at existing text. However, there are some details that can be ironed out with distant outliers in inputs being able to "break" the prototype. This is because GPT-3.5 is trained on a wide variety of data. When giving "wrong" inputs into this prototype, GPT-3.5 generated irrelevant answers. For example, if the user keeps asking the AI questions unrelated to any stories.

When used correctly, GPT-3.5-turbo is quite capable of comparing different stories. From testing results, given the right system prompt; it can find new details in the story, and it can even recognize if any spelling mistakes were made through the speech-to-text translation. In most cases, it is also capable of comparing stories told in a different language. What we found was able to break the code that we found was able to break the code is by giving a question to the AI instead of feeding a story. This problem is partially mediated by the users not directly inserting prompts to GPT-3.5, as the client must put more effort into "breaking" the system. Some of the key results of testing inputs are shown in table 1 below. For further testing on inputs and outputs please refer to appendix 8.4.

Table 1: LLM examples of iterative testing inputs and outputs

System input	User translated input	Generated output
Compare the user input compared to this story: My name is Leonardo and I'm 24 years old. I study at the TUE. In my free time I like to paint and I do Taekwondo.	Hello my name is Leo and I'm 21 years old. I study at the TUE and in my free time I like to play football	The user input and the story have some similarities but also some differences. - Both mention the name Leonardo and studying at the TUE. - However, in the story, Leonardo is 24 years old, while the user input states that they are 21 years old. - Additionally, the story mentions that Leonardo likes to paint and does taekwondo, while the user input mentions that they like football
Compare the user input compared to this story: My name is Leonardo and I'm 24 years old. I study at the TUE. In my free time I like to paint and I do Taekwondo.	Hallo mijn naam is Leonardo en ik ben 24 jaar. Ik studeer aan de TUE. In mijn vrije tijd, schilder ik graag en ik doe aan Taekwondo	De gebruikerinvoer is identiek aan het verhaal.
Compare the user input and state the differences compared to the following story: In 2009, my husband John and I went on a three-week trip to Japan. We went to Osaka and Tokyo, and we tried so many delicious foods – you wouldn't believe it! The only thing was, it rained a lot. But we didn't let that stop us. We even did a boat tour in the rain, and it turned out to be so peaceful and beautiful. Despite getting a bit soaked, the whole experience made our trip super memorable. Good times!	In 2009, my husband and I went on a trip. We went to Japan. It was raining all the time. However, we still went on a boat tour which was very beautiful. Despite getting soaked, we also ate a lot of delicious food. It was a very memorable trip.	The user's input is almost identical to the original story. The only minor difference is that the user did not mention the cities they visited in Japan (Osaka and Tokyo), unlike the original story.
Deconstruct the user input and state all the differences compared to the following story after deconstructing it as well: In 2009, my husband John and I went on a three-week trip to Japan. We went to Osaka and Tokyo, and we tried so many delicious foods – you wouldn't believe it! The only thing was, it rained a lot. But we didn't let that stop us. We even did a boat tour in the rain, and it turned out to be so peaceful and beautiful. Despite getting a bit soaked, the whole experience made our trip super memorable. Good times!	In 2009, my husband and I went on a trip. We went to Japan. It was raining all the time. However, we still went on a boat tour which was very beautiful. Despite getting soaked, we also ate a lot of delicious food. It was a very memorable trip.	Differences compared to the story: 1. The name "John" is not mentioned. 2. The duration of the trip is not specified (three weeks in the original story). 3. The specific cities visited, Osaka and Tokyo, are not mentioned. 4. The level of rainfall is not specified in the user input. 5. The phrase "you wouldn't believe it!" is not included when talking about the delicious foods. 6. The boat tour is described as beautiful and peaceful in the user input, without mentioning the rain. 7. The phrase "despite getting a bit soaked" is not mentioned in the user input. 8. The phrase "good times!" is not included in the user input.



## **5 DISCUSSION**

### **5.1 Outcomes**

This project focused on how machine learning can be used to support caregivers in supporting dementia sufferers in their storytelling. The created program transcribes stories and compares them to previous versions of the story to formulate questions about the missing details. By supporting the continuation of the conversation, through questions about missing details, the conversations will become livelier for the elderly and can support caregivers in having multiple conversations per day with multiple elderly. It will become easier to ask questions relating to said story. At its current form this program solves 2 of the 3 aspects mentioned earlier. It solves both loneliness and identity elements. First it solves the isolation element as it allows for more talking points to be accessed by the caregivers. This way the caregivers can have more elaborate conversations with the dementia sufferers about their own memories. Second, this program allows dementia sufferers to regain their identity for a short period of time. As they are guided to recall their old memories, their identities are allowed to re-surface during their recollection period. As for the final element, the memory tracking aspect, would need to be investigated during further research. The project in its current state is not viable for use in care homes, with the reasons being addressed in the following sections.

### **5.2 Limitations**

#### *5.2.1 User experience*

The project has illustrated the user experience (UX) flow through the Figma visuals (Figure 1). However, the UX has not been designed with feedback or co-creation methods for active engagement of the end-user. Besides, the project has been based on literature and does not include in-depth interviews with its stakeholders. Working together with care home personnel is vital to get a better understanding of the problem and the needed solution. Therefore, it is not yet possible to claim that the UX is completely satisfactory for the caregivers. Further research and user tests will improve the product to stakeholder needs.

#### *5.2.2 Usability*

Gradio has been used to create the interface of the algorithm to demonstrate the functionality of the algorithm. However, the designer interface was not created to provide proper usability to the user in the care home. Therefore, it is necessary to improve the project and create a better user-fit interface. This can be either done by building an interface from the ground up to create a mobile application, or more easy and more limited by using the markup customizability within Gradio.

Additionally, Whisper used in this project is limited to a file size of 25 MB. However, it should be possible to increase this file's size to allow users to tell longer stories and not be limited to the file size.

### **5.3 Future work**

#### *5.3.1 Inclusion of semantic textual similarity model*

When the design is put into use, the focus will be on collecting memories and analyzing if details are missing. Stories are often told differently, but the details are a core of the story so they should be mentioned in each version of the story. The algorithm should support the caregiver in analyzing details that are missing. To analyze the stories, a semantic textual similarity (STS) model could be used. The purpose of STS models is to generate a score indicating the similarity between texts [20]. In this use-case the STS model can help make a distinction between the 'original' memory, containing all the

details, and the similar memory told later in time. If the STS will compare both memories, a distinction between the number of details might be found. The STS will communicate this by showing a lower score for the most recent memory, which will then put the rest of the code into action.

To ensure that the code will be put in action depending on the STS score, the threshold for this score needs to be researched. For the STS threshold research, the same memory should be collected multiple times to research how the difference in memories affect the STS threshold. In addition to this, the same set-up should be collected over a wide variety of memories to define when the STS threshold is at the sweet spot of indicating missing details.

### *5.3.2 RAG*

As shown in this project, the large language model (LLM) like ChatGPT provides a clear answer to the caregiver. In this project the algorithm has been designed with one hardcoded memory. However, for this project to work in care homes the memories collected need to be stored somewhere. For this purpose, a retrieval augmented generation (RAG) could be used. A RAG model can connect a LLM with an external database to supply the LLM with additional information to strengthen answers [11]. By storing the memories in a database instead of hardcoding, this project has the potential to be used by many more people by collecting and storing more memories.

### *5.3.3 Pattern Recognition*

The next step regarding this project would be the implementation of some form of unsupervised ML model for pattern recognition. The hope is for there to be some form of pattern that highlights the decline in memory and cognitive abilities. That pattern could then be explored by fine-tuning an unsupervised ML model using the data collected and stored in the dementia sufferers' database. This could then help us understand the progression of dementia by tracking the memory recollection abilities of dementia sufferers. Implementing this program, which allows for memory collection and annotation, and having the pattern recognition model run alongside would be the path moving forward for helping experts understand how fast dementia progresses during its various stages.

### *5.3.4 Expanding to other fields*

The design decisions made during this project have been based on the literature on memory loss from dementia sufferers. The current design focuses on providing support in conversations with dementia sufferers and caregivers. Besides providing support, there is another opportunity for NostalgAI to support caregivers. Through data collection, it could be evaluated over time how many details a dementia sufferer is missing. This could help caregivers in creating a clearer picture of how the patient's dementia is increasing over time.

Another opportunity to expand to other fields unrelated to dementia. There are other fields where people suffer from memory loss, such as old age or after excessive alcohol usage [2, 3]. Therefore, NostalgAI could be developed further, for example by using other aesthetics and positioning itself as the application to support engaging conversations with people with memory loss.

## **6 CONCLUSION**

This project's purpose was to design a solution to support social interactions for dementia sufferers regarding 3 main areas of interest; reducing loneliness, supporting their identities and recording their memory and cognitive decline. This was performed using story telling as it's a vital social interaction for dementia patients provided by caregivers. This report introduces NostalgAI, an application that collects memories, compares versions of these memories over time, and provides

the caregivers with questions based on missing details through AI analysis. NostalgaAI's algorithm is created in Python 3.1.1 with OpenAI's Whisper and GPT-3.5-turbo with an interface made through Gradio. A user-experience prototype has been created to illustrate the functionality of NostalgaAI. Preliminary testing has shown that the concept has potential to support the intended social interactions.

However, this paper has also shown that NostalgaAI needs further improvement before it can be used in care homes, such as the implementation of STS and RAG models. The STS and RAG models are existing technologies that could be integrated into the project's current state. Further development would require support from experts on dementia care and existing technology to implement this properly. Besides, a large study should be conducted to formulate the STS threshold where a large collection of memories over time is necessary. Following their implementation would then allow for the full automation of the database creation and maintenance. This would allow the caregivers more time for the social aspect of conversing with the dementia sufferers and less time spent on the backend side of the program. Finally, the project was not based on user insights which should be included to evaluate the user experience aspect of the project. This project should be viewed as proof of concept rather than a final product. With a final aim of gaining more insights into memory and cognitive decline, while fully supporting any caregiver as a conversation starter with dementia sufferers.

## 7 REFERENCES [ACM]

- [1] Abubakar Abid, Ali Abdalla, Ali Abid, Dawood Khan, Abdulrahman Alfozan and James Y. Zou. "Gradio: Hassle-Free Sharing and Testing of ML Models in the Wild." <https://doi.org/10.48550/arXiv.1906.02569>
- [2] Alcohol hulp. N.d. Alcohol en geheugenstoornissen. Retrieved January 23, 2024 from <https://www.alcoholhulp.be/alcohol-geheugenstoornissen#:~:text=Overmatig%20alcohol%20drinken%20verhoogt%20de,goed%20aanpassen%20aan%20nieuwe%20situaties.>
- [3] Alzheimer Nederland. N.d. Oorzaken van vergeetachtigheid. Retrieved January 23, 2024 from <https://www.alzheimer-nederland.nl/dementie/oorzaken-preventie/oorzaken/oorzaken-vergeetachtigheid#:~:text=Als%20iemand%20vergeetachtig%20wordt%2C%20betekent,lichamelijke%20of%20psychologische%20oorzaak%20hebben.>
- [4] Gopi Battineni, Nalini Chintalapudi, Mohammad Amran Hossain, Giuseppe Losco, Ciro Ruocco, Getu Gamo Sagaro, Enea Traini, Giulio Nittari and Francesco Amenta. 2022. Artificial Intelligence Models in the Diagnosis of Adult-Onset Dementia Disorders: A Review. *Bioengineering*, 9(8), 3701
- [5] Berlin-Institut für Bevölkerung und Entwicklung. 2011. Most common forms of dementia. (February 2011). Retrieved January 22, 2024 from <https://www.statista.com/statistics/271934/forms-of-dementia-worldwide/>
- [6] Helen B. Burks, Jude K.A. des Bordes, Riya Chadha, Holly M. Holmes, and Nahid J. Rianon. 2021. Quality of life assessment in older adults with dementia: A systematic review. *Dementia and Geriatric Cognitive Disorders* 50, 2 (2021), 103–110. DOI:<http://dx.doi.org/10.1159/000515317>
- [7] Szu-Wei Cheng, Chung-Wen Chang., Wan-Jung Chang, Hao-Wei Wang, Chih-Sung Liang, Taishiro Kishimoto, Jane Pei-Chen Chang, John S. Kuo, & Kuan-Pin Su. 2023. The now and future of ChatGPT and GPT in psychiatry. *Psychiatry and Clinical Neurosciences*, 77(11), 592–5962
- [8] John Elflein. 2019. People with dementia forecast 2018-2050. (September 2019). Retrieved January 22, 2024 from <https://www.statista.com/statistics/264951/number-of-people-with-dementia-from-2010-to-2050/>
- [9] Google. 2023. Introduction to large language models. (August 2023). Retrieved January 21, 2024 from <https://developers.google.com/machine-learning/resources/intro-llms>
- [10] Karin Holmén, Kjerstin Ericsson, and Bengt Winblad. 2000. Social and emotional loneliness among non-demented and demented elderly people. *Archives of Gerontology and Geriatrics* 31, 3 (2000), 177–192. DOI:[http://dx.doi.org/10.1016/s0167-4943\(00\)00070-4](http://dx.doi.org/10.1016/s0167-4943(00)00070-4)
- [11] Chandler K. 2023. Diving Deep with RAG: When AI Becomes the Ultimate Search Assistant. Retrieved January 19, 2024 from <https://medium.com/@crskilpatrick807/diving-deep-with-rag-when-ai-becomes-the-ultimate-search-assistant-e55b06dada3>
- [12] Vijayshri Khedkar. 2023. INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING. *International Journal of Intelligent Systems and Applications in Engineering* 4
- [13] T. Kitwood. 1997. The experience of dementia. *Aging & Mental Health* 1, 1 (1997), 13–22. DOI:<http://dx.doi.org/10.1080/13607869757344>
- [14] Bère Miesen. 1992. *Care-giving in dementia* (1992). DOI:<http://dx.doi.org/10.4324/9781315800196>
- [15] Marie A. Mills. 1997. Narrative identity and dementia: A study of emotion and narrative in older people with dementia. *Ageing and Society* 17, 6 (1997), 673–698. DOI:<http://dx.doi.org/10.1017/s0144686x97006673>
- [16] OpenAI. 2023. ChatGPT (Mar 14 version) [Large language model]. <https://chat.openai.com/chat>
- [17] OpenAI. 2023. Models. (2023) Retrieved January 23, 2024 from <https://platform.openai.com/docs/models/GPT-3-5>
- [18] Martin Prince, Emiliano Albanese, Maëlen Guerchet, and Matthew Prina. 2014. World Alzheimer Report 2014: Dementia and Risk Reduction. An Analysis of Protective and Modifiable Factors. *Alzheimer's Disease International (ADI)* (2014).
- [19] Alec Radford, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey, and Ilya Sutskever. 2023. Robust speech recognition via large-scale weak supervision. In *International Conference on Machine Learning* (2023), 28492–28518.
- [20] Aliaksei Severyn, Massimo Nicosia, and Alessandro Moschitti. 2013. Learning Semantic Textual Similarity with Structural Representations. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*. Sofia, Bulgaria, 714–718.
- [21] Tobias R. Spiller, Finn-Lennart Rabe, Ziv Ben-Zion, Z., Nachshon Korem, Achim Burre, Philipp Homan, Ilan Harpaz-Rotem, Or Duek. 2023. Efficient and Accurate Transcription in Mental Health Research—A Tutorial on Using Whisper AI for Audio File Transcription [Preprint]. *Open Science Framework*.
- [22] UN DESA, & Gapminder. 2022. United States: Life Expectancy 1860-2020. (June 2022). Retrieved January 22, 2024 from <https://www.statista.com/statistics/1040079/life-expectancy-united-states-all-time/>
- [23] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2023. Attention Is All You Need. <https://doi.org/10.48550/arXiv.1706.03762>

## 8 APPENDIX

### 8.1 Work distribution and individual reflections

#### 8.1.1 Ahmed Gamal Mohamed Sayed Ahmed Ibrahim

##### Individual contribution

I contributed to the literature research regarding the topic of dementia and the negative consequences it has on the sufferers' lives. By providing extensive knowledge of the topic, the group was able to focus primarily on the implementation of our AI-based solution. After that, I was involved during the ideation phase on how we can implement our vision and the steps needed to move forward. I believe I contributed greatly during this phase as I already had an idea of what was needed from start to finish, as a result of my work on the individual poster. My focus then shifted to first understanding how to implement ML models locally, starting with Whisper. After that, I also implemented the STS model to work, but it was then decided that we needed to drop its use within the context of this project. My focus then shifted to understanding how Gradio works and implementing it to work with Whisper. I was also involved in the creation and presenting of our solution during the presentation aspect of this project. I was also able to convince the team regarding the need for an interface for our program outside of our Gradio interface. While also preparing the script and acting it out for our product video. Finally, I handled the introduction, whisper implementation, and part of the discussion for the report.

##### Individual Reflection

I come from a non-design background. I did my bachelor's in Psychology and Technology and I'm currently doing a master's in Human-Technology Interaction (HTI), where it's predominantly research-oriented. I see myself as having a great background in conducting research, quantitative and qualitative, but lack the design skills to be able to create prototypes that support my research. I enrolled in this course as I am interested in design within the context of AI, and I wanted to learn the different design methodologies one can use. I wanted to learn how to practically implement technologies and not just research them. In this course, this was in the form of learning how to use Machine learning (ML) models from an implementation perspective and not theoretical or mathematical perspectives.

During this course, certain learning goals were set as a way to evaluate how much one has learned after following this course. One key element was to first ask oneself whether a problem requires an AI-based solution or not. I appreciated hearing this out loud as I was fatigued by the number of times I have heard AI used as a marketing tool for all kinds of products. A key element that was taught during this course was to first take your time about the problem at hand and how to solve it. After which you can reflect on whether an ML model is necessary. To be able to reflect on this matter, one must first have an overview of the available types of ML models currently available. Which was what the first couple of weeks of this course were set aside for. I was first introduced to the different types of tasks that ML can be used for and the relevant information required to understand how those tasks are performed. I was introduced to the different types of supervised and unsupervised approaches, and I learned about the differences between them. As well as how model evaluations are performed, the different types of networks as well as generative models among other things. This course provided me with an overview of the different models available as well as other key elements that need to be understood, before working with ML models. This method of teaching provided me with a nice overview and guidelines on what I need to look for when working with ML models.

Following on the aforementioned information, I now understand how important data pre-processing is. I learned that good ML models are as good as the data that they are trained on. The better the annotated data is, the better the ML model runs. While I did not gain any insights on how to pre-process data for training purposes, I at least know how an ML model may fail in achieving high confidence percentages. I was also introduced to the Hugging Face website which is great for running demos of ML models. As well as providing you with access to implementations of different ML models.

I then started on my own journey of learning how to use Python to a sufficient level to be able to implement ML models locally. I was introduced to the process of creating new environments and how to install the dependencies needed per ML model used on said environments. While it was challenging at times, it's always great to see the fruits of one's labor realized. In this context, it relates to my increased understanding of installing and coding the different ML models. Especially when it came to running Whisper and all-MiniLM-L6-v2 locally as it was a constant struggle to match dependency versions to the required ones needed to run them error-free. It was also great to learn how one can use Gradio as an initial prototyping interface to showcase ML models. As it makes it a lot easier to have layman interact with ML models without having any programming knowledge to do so. I believe that I was able to challenge myself thoroughly when it came to implementing ML models. Even at times when I thought I should just give up and try to find someone else to help me with my challenges, I maintained my drive to learn how to solve the problems I faced. To summarize my learning goals, I believe I have achieved greater insights at an overview level of all the available ML models introduced in this course as well as how to handle the coding aspects of practically implementing them while creating an initial interface for their use.

As for the group work aspect, it was relatively smooth but not as great as I hoped it would be. As my idea was chosen for the group work, my excitement and proactiveness towards it was high. Sadly, I didn't feel that the rest of the team was as excited or proactive as I was. It mostly felt like it was just an idea that they needed to implement well enough to pass this course and move on. While I viewed this idea as something that could be worth working on extensively. This slightly dampened my proactiveness towards the work needed for this group project, but I still maintained the necessary level of proactiveness. Other than this fact, the team communication was on point and everyone for the most part contributed to the final report. I have worked in multi-disciplinary teams prior to this one and while it can sometimes be a challenge, that wasn't the case this time. I believe that was a result of already having a goal to work towards, while also having the relevant roadmap needed to reach there. Thus, eliminating the ideation phase where conflict tends to occur. The group work was also helped by the fact that everyone knew their strengths and weaknesses and worked around them.

Based on the aforementioned insights, I believe that I am learning new and relevant skills that allow me the flexibility to prototype ideas. As I am interested in UX research within the sector of warm technologies (IJsselsteijn et al., 2020). One of the main concepts regarding warm technologies is a user-centered, personalized approach to design when it is related to activities of daily living (ADL) (Spector et al., 1987). To achieve a user-centered and personalized approach, first I need to be able to collect data in a naturalistic social environment while focusing on rich qualitative insights. I then need to create prototypes that are based on user data and feedback. Which entails the requirement of being able to create said prototypes. Since a user-centered approach is dependent on an iterative process of research and design. After learning how to implement ML models to run locally and how to run them on 3rd party servers. I believe that I am at the start of my journey of being able to use ML models as a tool to implement my envisioned ideas as well as help me create tangible prototypes using them.

To conclude, I believe that I have gained a great tool for my design toolkit as AI is increasingly becoming more dominant in society. This course allowed me to take the first step into learning how to use ML models to not be left behind technologically, as a researcher and designer.

#### Course Overview

Overall, I'm happy with what I have learned in this course. The focus on having an overall understanding of AI and skipping the in-depth mathematical knowledge is on point from a design perspective. As the focus should be more on the implementation aspect and nothing else in my opinion.

#### References

IJsselsteijn, W. A., Tummers-Heemels, A., & Brankaert, R. (2020). Warm Technology: A Novel Perspective on Design for and with People Living with Dementia. In Human-computer interaction series (pp. 33–47). [https://doi.org/10.1007/978-3-030-32835-1\\_3](https://doi.org/10.1007/978-3-030-32835-1_3)

Spector, W. D., Katz, S., Murphy, J. B., & Fulton, J. P. (1987). The hierarchical relationship between activities of daily living and instrumental activities of daily living. *Journal of Chronic Diseases*, 40(6), 481–489. [https://doi.org/10.1016/0021-9681\(87\)90004-x](https://doi.org/10.1016/0021-9681(87)90004-x)

#### 8.1.2 Harry Wang

In this course, I went through a profound academic journey through self-study and group work to finalize a design project on the use of AI to assist in the care of people with dementia. Below is a summary of my learning goals and individual contributions

#### Learning goals and outcomes

##### 1. Increase knowledge about AI

Through in-depth study, I managed to expand my knowledge of AI. Understanding how AI works and the range of applications provides a solid theoretical foundation for solving practical problems.

##### 2. In-depth study of how AI works:

The focus of the course enabled me to gain a deeper understanding of how AI works, particularly in the area of speech recognition. This provided more accurate technical support for the care systems we designed.

##### 3. Exploring the possibilities of AI in the design field:

During the project, we explored the innovative possibilities of integrating AI into the care field. Through speech recognition, we provided caregivers with a smarter and more efficient way to provide care.

#### 4. Integrating AI into Designer Workflows:

I have worked to make AI an integral part of design workflows to improve design efficiency and creativity. This experience has enabled me to better understand how to embed AI technology into the actual design process.

### **Individual Contributions to Group Projects**

For the group project, my personal contribution focused on searching for relevant cases and integrating those cases into our design. I delved into case studies related to nursing, speech recognition and dementia patient care, which provided valuable references and inspiration for our design.

#### 1. UX Design

I was involved in the user experience (UX) design of the project, ensuring that our system interacts with caregivers in a user-friendly way that enhances the experience of using it.

#### 2. Video editing:

I was instrumental in creating the project's demo video, which vividly demonstrated the workings and benefits of the AI-assisted care system we designed.

#### 3. Product Workflow Framework:

I was responsible for organizing the product workflow framework of the project to ensure the organic connection of each link and improve the overall effectiveness of the system.

In summary, through this course, I not only deepened my understanding of AI, but also contributed to the team project and accumulated valuable experience in applying AI to actual design work. This learning experience has not only improved my professional skills, but also laid a solid foundation for my future design work.



### *8.1.3 Leonardo Chávez Sáenz*

#### **INTRODUCTION**

The advanced AI course contains many interesting topics to delve into. It can be taken from different angles depending on what the individual wants to get out of it. I personally chose this course because of my interest in innovation in the tech industry. I have a background in Creative technology at the University of Twente. There I learned multiple things about integrating technology within design. Now that I follow Industrial Design at the TU/e I want to continue to improve in these skills by getting expertise within ‘Math, Data and Computing’ and the ‘Business and entrepreneurship’ expertise area. AI still has immense potential to grow, and it also contains more use-cases than I can currently think of. I have little experience with machine learning, so this was the perfect opportunity to delve deeper and explore it.

#### **INDIVIDUAL CONTRIBUTION**

In the progress of creating the dementia caretaker application the team made sure to work together by planning weekly meetings and dividing tasks. As someone with decent knowledge in programming, I took some of the programming tasks upon myself.

Together with Ahmed, I worked on integrating the API’s together in a python script. The application contains three tools in total. First was Whisper, a speech to text tool, which was mostly handled by Ahmed. Then Gradio was integrated next. This tool is mostly used to provide an easy to make interface for other AI tools. It also handles multiple forms of output, which was very convenient to receive audio for Whisper. Once this was done, I worked individually on integrating the last and biggest piece of programming, which was OpenAI’s GPT API. I chose to go for gpt-3.5-turbo which was the latest version of GPT 3. One problem that arose was that this version used the OpenAI 1.0.0 python package which was relatively new, so it had less documentation.

After the coding part was finished, we focused together on finetuning a system prompt, so that the output was up to our standards. And next to that we made sure that everyone provided enough contributions to the final report and video. For the video, I did some filming and the voiceover. For the report, I personally worked the most on the Design solution and Results chapter, but just like everyone made sure to read over other parts and provide feedback or any necessary additions.

Overall, I am happy with my contributions. I was also happy with the result of the project. In the team, there was frequent and clear communication, but a relatively substantial portion of work was done individually between the sprint meeting days.

#### **INDIVIDUAL REFLECTION**

##### **The Advanced AI course**

The advanced AI course provided interesting opportunities and materials to get a better understanding of AI. I learned more about how AI works, what kind of AI is out there and diverse ways it can be trained. While I do have programming experience, I have never programmed or designed something specifically using machine learning or AI. I always saw AI as a giant hurdle that requires a high understanding of knowledge to program with it. This was not the case. In fact, using APIs from existing AI tools is relatively easy, as the designers of the tools made them easy to adjust to the users’ preferences. I also realized the advantages and opportunities advanced AI tools bring. Machine learning provides multiple specialized tools that are excellent at analyzing. The most important part is that the designer needs to understand the tool and guide it towards correct behavior.

### **Assignments**

During the course, there were individual and group assignments. I personally felt challenged by them but also enjoyed working on them. During the individual assignment I had an unfortunate case of COVID-19 which took more time to recover from than I hoped. For the rest I had some struggles to come up with an interesting use-case for a project. As I am new to industrial design there were not a lot of previous projects I could pull inspiration from.

The group project was personally more enjoyable. The project went overall smooth. From previous projects I was worried about effective communication and in this project, it showed yet again how important it was. Having regular meetings worked wonders and the only regret that I had was not guiding the project just a bit more and pushing for more regular meetings. Luckily, my teammates were very motivated individuals which all contributed greatly in their individual strong points.

### **Contribution to future career**

As both a designer and a programmer, AI is a powerful tool to use. I have a vision of becoming a designer with expertise in programming and technology. My new understanding of machine learning and AI tools will be extremely useful for future ventures. I want to contribute to society by being a leading force of innovation through creative but useful designs. AI is a useful tool that deserves more time investment. What was new for me was understanding to what extent AI tools are developed and what possibilities are still out there to use AI. In the team project, I found an innovative way of using AI in a use-case that does not use AI just for the sake of it. Instead, it uses AI because it is the best fitting tool for the job. With this new outlook and deeper understanding I will be able to more effectively know when to use AI and where to improve design using AI.

### **FUTURE SUGGESTIONS**

Overall, I am happy with the knowledge gained from this course. The material was interesting and the way the group project was structured was helpful. Yet, like with anything, there is always space for future improvement. What I missed in the course was a more engaging way of starting the course of. Most of the course was done through the Flipped Classroom approach. This contained a fair amount of unsupervised learning through video lectures. I found that this approach didn't stimulate any engagement with the material or the teachers. While the material was interesting, I found myself looking at it less and less as not all material was integrated into the tasks given. This reminds me of the way lectures were given during COVID time. During this time, I worked as a teacher assistant and noticed that because of the lack of contact and group engagement there was a large discrepancy between understanding and progress of different individuals. Overall, I enjoyed the course and I'm looking forward to seeing improvements.

### 8.1.4 Vere Vreeswijk

#### **INTRODUCTION**

To be able to reflect on this course, it is essential to mention the reason why I chose to follow this course. Throughout my bachelor's in Industrial Design (TU/e), I have followed multiple courses in the expertise of Math, Data and Computing. However, none of them have focused on artificial intelligence. Artificial Intelligence (AI) has become more important and integrated into our society, not only accessible to experts but also to all citizens. Therefore, as a designer, I highly value knowing what AI is, how it functions, and how and when to implement it. The purpose of following this course was to gain background knowledge in AI and skills in working with it to become an all-round designer with knowledge in one of society's most significant innovations.

#### **CONTRIBUTION TO THE GROUP PROJECT**

In the group project, I noticed at the start that I could maintain a helicopter view of the process and the code we were trying to create. The group has made a code out of multiple existing algorithms. Therefore, it was necessary to keep an overview of what was lacking, which component was redundant, and if the code would achieve our intention. With this overview, I supported the group in shifting our focus from a detailed perspective (focusing on the code) to an overarching perspective (focusing on the user experience and purpose). This has allowed us to ask critical questions supporting our development. Besides maintaining an overview of our project, I have also steered our process through making task proposals, note keeping and dividing workload.

Besides keeping an overview and managing our process, I have also worked on our prototype. I have focused on researching the threshold of the STS model in combination with LLM. This exploration has given us more insight in how the STS and LLM could work together and how we could define the threshold. I have also researched other methods to strengthen our concept, such as using the retrieval augmented generation (RAG) model. In addition to the research, I have looked into integrating the LLM into the code. However, as I could not manage them to work, my other team members took over. I continued with researching the prompts we could use in our code (Appendix 8.4). Additionally, I have created the user-experience prototype in Figma (Appendix 8.3). The prototype is an interactive prototype displaying how the user could interact and use our design.

Lastly, I will describe my contribution to the group report. I took the lead in writing the discussion and conclusion, reading through the chapters, providing feedback and critical questions, improving chapters, keeping an overview of the work that needed to be done and focusing on using the layout correctly.

#### **REFLECTION**

##### **Foundation in AI**

This course has supported me in creating a foundation for understanding how AI and ML work, the opportunities they can provide and how they can be implemented in the design. From the activities throughout the course, such as the individual assignment, I have learned that my initial design could have been made more structured and advanced if I had implemented the knowledge I gained through this course. It was a valuable assignment to reflect on a previous project and discover how AI could improve the quality of the project. In my future projects, I will have the skills and the awareness of AI implementation, which will help me make better design decisions as I can now consider more perspectives.

When looking at the gained hard skills in coding it is less than intended when choosing this elective. My group members were more skilled in coding, which has led to the fact that I also became afraid to 'trial and error' and learn from that, which is abnormal for me as I usually do not step away from a challenge. However, this has brought to my attention that math, data and computing remains an expertise area to further explore how to integrate it into my projects.

## **Communication**

Throughout the course, I have gained knowledge on the fundamentals of AI, how to work with them, and how to consider AI on a conceptual level. From my group contribution, it is clear that thinking about AI from a conceptual perspective is necessary to discuss what goal the project is trying to achieve. Through the group work, I have learned that it is one of my strengths to facilitate this conceptual thinking and help the group to switch between different levels of discussion, from project conceptual to detailed and from user experience to code.

From working in a multi-disciplinary group with different backgrounds, I have experienced that different perspectives come to the table. In the group, getting on the same page sometimes remained difficult as people had other interpretations of the same ideas due to different backgrounds. I have used various techniques such as acting it out, systematic overviews, drawings and more to find out where we were misaligned and to get on the same page. Throughout the rest of my master's, I have not often experienced this struggle as in most other projects the background knowledge was closer to each other.

As a designer with a potential career in a society driven by AI, it is vital to think on all these different levels to discuss design solutions with various stakeholders. Designers differentiate themselves by being able to communicate innovative design solutions to various stakeholders, which makes the skill of altering your 'story' to your 'crowd' a necessity. As a designer, you also often work within a team where people have various skills, making this a significant development as I can now communicate plans and activities on multiple levels. This course provided me with this insight as most of my projects do not require this shift between levels.

## **Development of Professional Identity and Vision**

A fundamental part of my professional identity is to become an all-round designer, i.e., knowing various design methods rather than being an expert in one. Becoming an all-round designer is important as I want to create fitting solutions to societal issues. Societal issues have various topics, stakeholders and opportunities, which requires designers to consider multiple perspectives when creating a solution. This course has added a new design perspective and method to my toolbox. I value understanding and working with AI as a new perspective, as a large part of society is becoming data-driven, and AI is used in solutions.

My vision focuses on climate action, behaviour change and positivity, which does not directly relate to AI. However, much scientific information is often presented through statistics and data collection when designing for climate change. As a designer, I switch between scientific information and information from stakeholders and try to combine them in my designs to create an impact on climate change. To do this, it is vital to understand what the scientific information is based on, such as the data collection, training of models, and how they make these predictions for the future, to communicate and translate this in my designs.

Additionally, this course has shown me that AI will not be a focus of my vision. I am content with the background knowledge gained through this course and being able to think about human-AI collaborations. However, I prefer to use AI on a conceptual level and communicate with experts about possibilities rather than becoming the expert.

## **CONCLUSION**

In conclusion, the purpose of following this course was to gain a foundation in AI and how to design with AI, which was successfully achieved. This course's group project has shown me that my qualities are in conceptual thinking about AI implementation and communicating with different disciplines when designing with AI. I have developed in becoming an all-round designer and the course has provided me with a AI-based perspective that I will consider when designing for climate action.

## 8.2 Python Code

The following is the code used to make the prototype of the design explained in the report.

```
import whisper
import openai
import gradio

from gradio.components import Textbox, Audio

openai.api_key = "INSERT_KEY_HERE" # Replace with own API key generated on OpenAI site

gpt_input = [{"role": "system", "content":
    """Deconstruct the user input and state all the differences compared
    to the following story after deconstructing it as well:
    In 2009, my husband John and I went on a three-week trip to Japan.
    We went to Osaka and Tokyo, and we tried so many delicious foods -
    you wouldn't believe it! The only thing was, it rained a lot.
    But we didn't let that stop us. We even did a boat tour in the rain,
    and it turned out to be so peaceful and beautiful. Despite getting a bit
soaked,
    the whole experience made our trip super memorable. Good times!"""]}

def transcribe_audio(audio_file): # Use Whisper to transcribe the audiofile
    model = whisper.load_model("small")
    result = model.transcribe(audio_file)
    return result["text"]

def process_GPT(user_input): # Generate a response from GPT using the transcribed input
    # GPT can receive different input sources. We append the user input to the system
input
    gpt_input.append({"role": "user", "content": user_input})
```

```

# request openai server to generate a response using the correct model
response = openai.chat.completions.create(
    model="gpt-3.5-turbo",
    messages=gpt_input
)
gpt_answer = response.choices[0].message.content
return gpt_answer

def process_all(audio_file): # function called by gradio to get the necessary output
    # Whisper part
    whisper_output = transcribe_audio(audio_file)
    print(whisper_output)

    # CHATGPT part
    gpt_output = process_GPT(whisper_output)

    return (whisper_output, gpt_output)

def main():
    # First we record the audio through Gradio
    # remove source="microphone", if you want to upload a file instead of voice record
    directly
    audio_file = Audio(label='Upload audio file', source="microphone", type='filepath')

    # Gradio textbox labels
    transcribed_textbox = Textbox(label='Transcribed story')
    gpt_textbox = Textbox(label = 'GPT answer')

    # Gradio interface calls the process all function.
    # It receives the audiofile as an input and gives back two outputs: the transcription
    and the answer by gpt-3.5

```

```
app = gradio.Interface(fn=process_all, inputs=[audio_file],
                        outputs=[transcribed_textbox, gpt_textbox],
                        title='NostalgAI',
                        description='Upload an audio file and hit the button')

# launch part
app.launch(share=False) #Set share to true to host publicly

if __name__ == '__main__':
    main()
```

### 8.3 User interface



Figure 4 Opening screen app



Figure 5 Menu screen - select story





Figure 6 - Screen of pre-recorded story while recording.

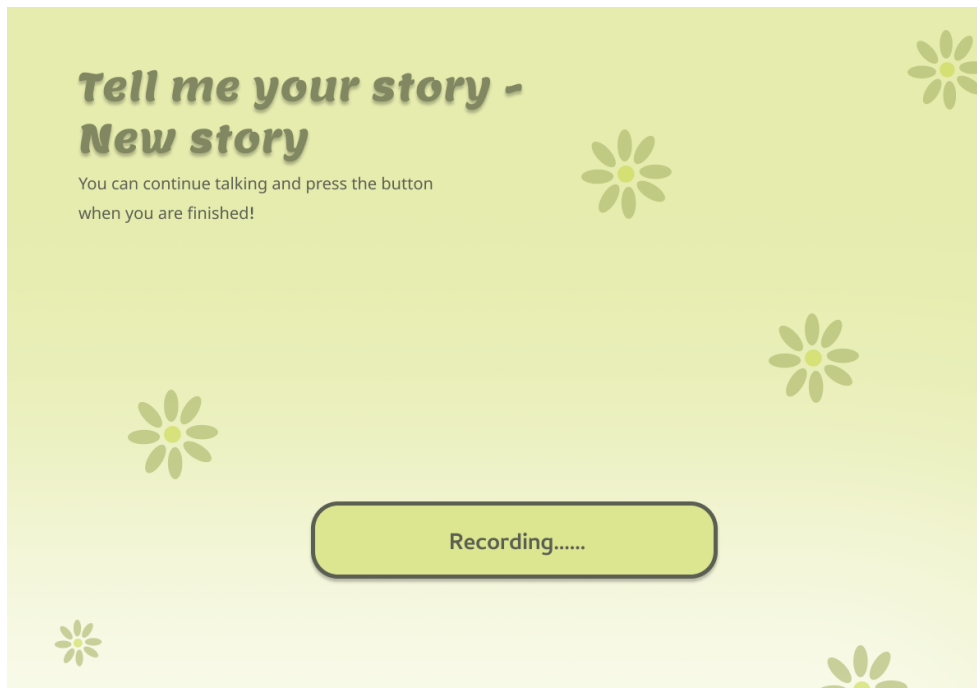


Figure 7 Screen when recording a new story



Figure 8 Screen for more information - Transcript of new story



Figure 9 Transcript of original story and recent told story



Figure 10 Screen with question suggestions

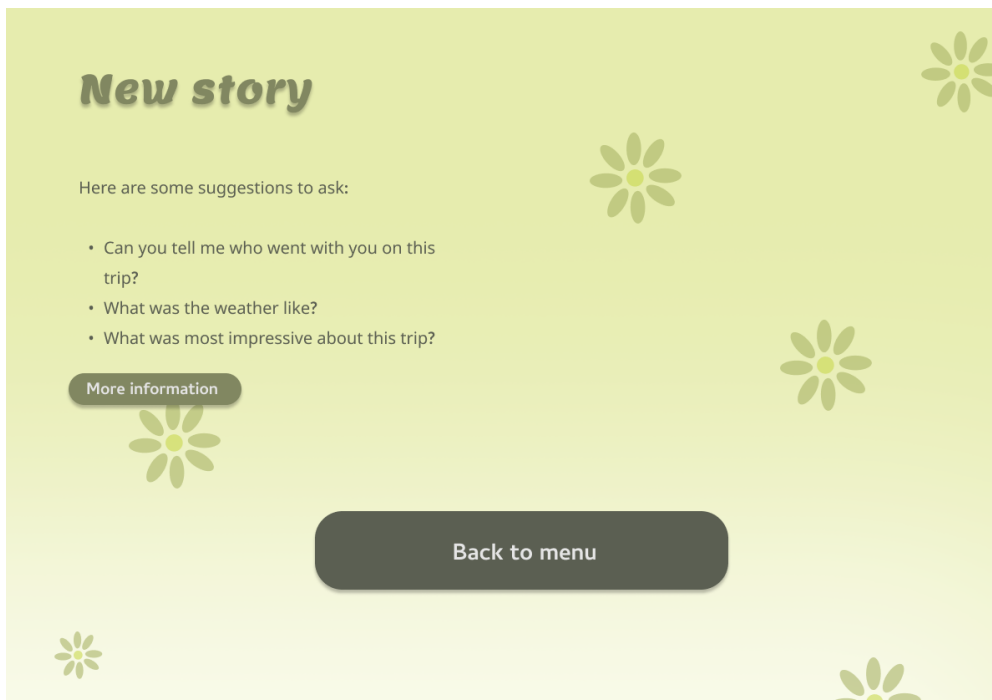


Figure 11 Screen with question suggestions

## 8.4 ChatGPT Prompt exploration

Table 2: ChatGPT Prompt exploration

Input	Generated output
Can you compare 'The cat chased the mouse around the garden' to 'The human caught the mouse with the cat' and explain what is wrong with the second sentence?	<p>Certainly! The first sentence, "The cat chased the mouse around the garden," is a straightforward and clear statement. It conveys a common scenario where a cat is pursuing a mouse in a garden.</p> <p>The second sentence, "The human caught the mouse with the cat," is grammatically correct but seems to describe an unusual or unlikely situation. In a typical context, humans do not catch mice with the help of cats. Cats are generally the ones known for chasing and catching mice.</p> <p>The issue with the second sentence lies in the unlikely scenario it presents, which goes against common understanding. It's not a grammatical error, but it is semantically confusing and doesn't align with typical real-world situations. If you want to convey a more realistic scenario, you might rephrase the second sentence to maintain a logical relationship between the cat and the mouse.</p>
Can you compare 'The cat chased the mouse around the garden' to 'The human caught the mouse with the cat' and explain why the content is not the same	<p>Certainly! The two sentences describe different scenarios:</p> <ol style="list-style-type: none"> <li>1. "The cat chased the mouse around the garden": In this sentence, the cat is the one actively pursuing the mouse. It implies a natural and expected behavior where a cat, as a predator, is chasing a mouse, a common prey, around a garden. The focus is on the cat's actions.</li> <li>2. "The human caught the mouse with the cat": This sentence presents a less common and somewhat unusual scenario. Here, it suggests that a human caught a mouse with the assistance or involvement of a cat. It creates a different dynamic where the cat is somehow instrumental in the human's act of catching the mouse. This situation is less typical and might be considered an imaginative or less realistic scenario.</li> </ol> <p>In summary, the content is not the same because the first sentence describes a standard predator-prey interaction between a cat and a mouse, while the second sentence introduces a scenario where a human is actively involved in catching a mouse, with the cat playing a role in the process.</p>
Can you deconstruct sentence 1: 'The cat chased the mouse around the garden' and deconstruct sentence 2: 'The cat chased the mouse'. Compare what is missing in the second sentence and tell in 6 words maximum what is missing.	<p>Sure, let's deconstruct both sentences:</p> <ol style="list-style-type: none"> <li>1. "The cat chased the mouse around the garden": <ul style="list-style-type: none"> <li>- Subject: The cat</li> <li>- Verb: chased</li> <li>- Object: the mouse</li> <li>- Adverbial phrase: around the garden</li> </ul> </li> <li>2. "The cat chased the mouse": <ul style="list-style-type: none"> <li>- Subject: The cat</li> <li>- Verb: chased</li> <li>- Object: the mouse</li> </ul> </li> </ol>

What is missing in the second sentence is the adverbial phrase "around the garden."

In 6 words maximum: "Location detail is missing in second."

Can you deconstruct sentence 1: 'The cat chased the mouse around the garden' and deconstruct sentence 2: 'The cat chased the mouse'. Compare what is missing in the second sentence. Only give as output what is missing in the second sentence in 6 words

Location detail around the garden missing.

Can you deconstruct sentence 1: 'The girl petted her horse with her sister Josie' and deconstruct sentence 2: 'The girl petted her horse with her sister'. Compare what is missing in the second sentence. Only give as output what is missing in the second sentence in 6 words

Sister's name "Josie" is missing.

Can you deconstruct sentence 1: 'The girl petted her horse with her sister Josie' and deconstruct sentence 2: 'The girl petted her horse with her sister'. Compare what is missing in the second sentence. Only give as output what is missing in the second sentence

Sentence 1: 'The girl petted her horse with her sister Josie'

Sentence 2: 'The girl petted her horse with her sister'

Comparison:

In Sentence 2, the specific name "Josie" is missing, which is present in Sentence 1.

Can you deconstruct sentence 1: 'The girl petted her horse with her sister Josie' and deconstruct sentence 2: 'The girl petted her horse with her sister'. Compare what is missing in the second sentence. Only give as output what is missing in the second sentence without showing what the sentences are

Josie

Can you deconstruct sentence 1: 'The cat chased the mouse around the garden' and deconstruct sentence 2: 'The cat chased the mouse'. Compare what is missing in the second sentence. Only give as output what is missing in the second sentence without showing what the sentences are

Sentence 1: 'The cat chased the mouse around the garden'

Deconstruction:

- Subject: The cat

- Verb: chased

- Object: the mouse

- Adverbial phrase: around the garden

Sentence 2: 'The cat chased the mouse'

Deconstruction:

- Subject: The cat

- Verb: chased

- Object: the mouse

Comparison:

What is missing in the second sentence: Adverbial phrase "around the garden"

Can you deconstruct sentence 1: 'The cat chased the mouse around the garden' and deconstruct sentence 2: 'The cat chased the mouse'. Compare what is missing in the second sentence. Output: what is missing in the second sentence without showing what the sentences are

In the second sentence, the prepositional phrase "around the garden" is missing compared to the first sentence.

Can you deconstruct sentence 1: 'The photograph shows 3 sons and 1 daughter named, Luc, Tom, Thijs and Anouk' and deconstruct sentence 2: 'I have 3 sons, Thijs, Tom, and Luc'. Compare what is missing in the second sentence. Output: what is missing in the second sentence without showing what the sentences are

In the second sentence, the daughter's name, Anouk, is missing.

# FutureNarratives

Vere Vreeswijk, s1314246

## Introduction

One of the issues with enabling people to act for climate change is that the effect of our present behaviour has no direct visible impact on the present, but on the future. The gap between the present and the future impacts personal relevance, affecting people's motivation to change their behaviour for a better future. Through storytelling, people are supported in thinking of their future selves and showing their possible future. The project has two parts. Firstly, an interactive story where users are challenged to think of their future. Secondly, the user is presented with a story structured like 'choose your own adventure' books but also based on the answers previously given. The current state of the prototype is limited, as all the 'choose your own adventure' options need to be considered by the designer. There is an opportunity to integrate machine learning (ML) to make the storylines more advanced, allowing for more personal relevant storylines.

## Implementation plan

With the integration of ML, the story could have more diverse paths. In the current prototype, all storylines had to be written manually, limiting the story's possibilities. Figure 1 shows a simplified version of the code structure. The code is repetitive and branches out quickly, limiting the variety of stories. In Figure 2, an alternative structure integrating GPT-4 (OpenAI, 2023) is given. It was decided to improve the project by integrating a GPT model, as GPTs are trained to produce creative and human-like texts because they are trained on a wide variety of resources (books, Wikipedia, etc.) (Amazon Web Services, Inc., n.d.; Klippa, 2023). With the integration of GPT-4, every time a person decides, this decision can be added to the existing prompt for the general storyline. The questions asked can also be created by the GPT-4 if the prompt guides the general storyline.

To successfully do this, the code should be rewritten from Ink (Inkle, n.d.) to Python and cleaned to extract the general storyline. In Python, the GPT-4 API should be integrated. Lastly, the code for the general storyline prompts and the prompts for questions should be created.

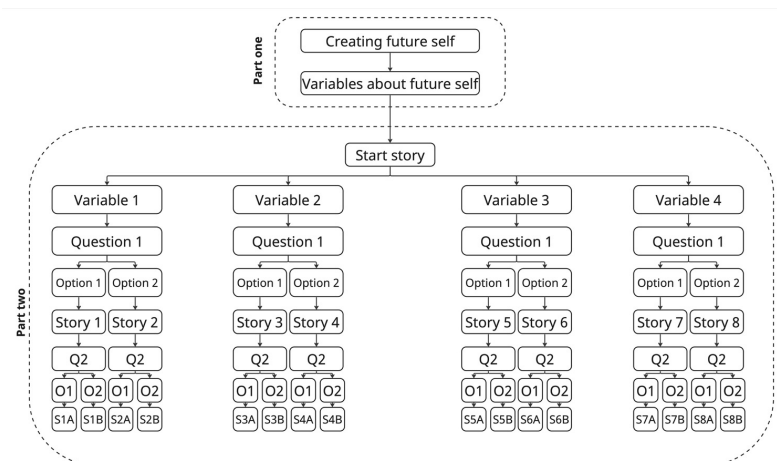


Figure 1 Simplified overview of structure of a part of the story

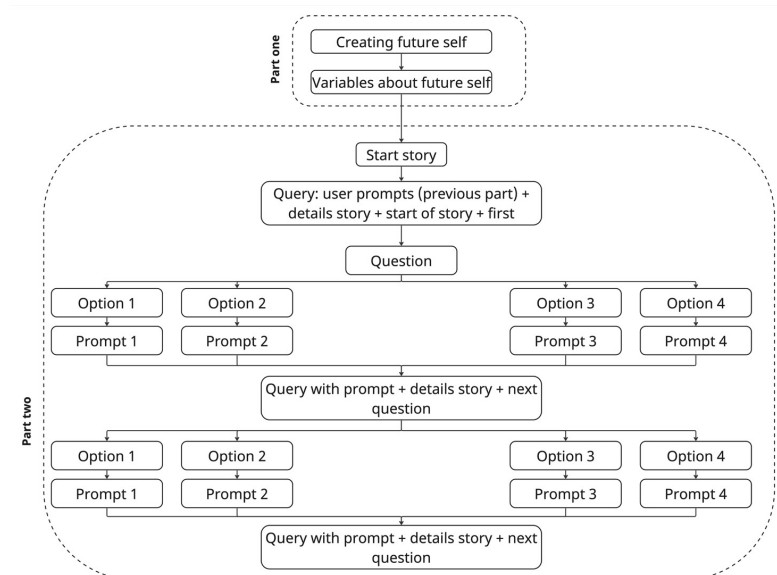


Figure 2 New structure with GPT-4 integrated.

## Results

From the perspective of an improved user experience, the user will have more control of the story. By having the GPT-4 write variations on the general storyline instead of a human considering possible directions, more diverse choices can be given, resulting in more story directions. A wider variety of story directions can increase the personal relevance of the story as it will start to feel more as if they are in control and feel more relatable.

## Discussion

How far the story should be differentiated to match the purpose of the story should be carefully considered. With giving the control of the story partially to the GPT-4, a risk could be that users end up with unintended stories. Another limitation can be that GPT-4 has difficulty overseeing the whole storyline and basing the generated text on that, as it works with a context window. Lastly, empathy and biases are often an issue in GPT's. The GPT-4 may not understand why people choose a particular direction, which might impact the story negatively.

### References

- Amazon Web Services, Inc. (n.d.). *What is GPT AI? - Generative Pre-Trained Transformers Explained - AWS*. Retrieved December 5, 2023, from <https://aws.amazon.com/what-is/gpt/#:~:text=The%20GPT%20models%20are%20general,and%20extracting%20data%20from%20documents>.
- Inkle. (n.d.). *ink - inkle's narrative scripting language*. <https://www.inklestudios.com/ink/>
- Klippa. (2023, March 22). *What is ChatGPT-4 & why is it important?* Retrieved December 7, 2023, from <https://www.klippa.com/en/blog/information/chat-gpt-4/#:~:text=As%20an%20AI%20language%20model,%2C%20generating%20text%2C%20and%20more>.
- OpenAI. (2023) GPT4 [Large Multimodal model]. <https://openai.com/gpt-4>

## Conclusion

Overall, the integration of machine learning in FutureNarratives could give a wider range of storylines, increasing the personal relevance of the experience. This method helps people understand the bridge between present behaviour and its impact on the future.

### Link to report

<https://drive.google.com/file/d/1VtOsEnjJTN8Fzr05QW62bUJYIU6cePg/view?usp=sharing>