AI Memory Sculptures: Creating Accessible Memory Preservation Method with Sentiment Analysis and Autonomous 3D Modelling

Submitted April 2023, in partial fulfillment of
the conditions for the award of the degree **MSci Computer Science with Artificial Intelligence.**

Filip Kornata

20192267

Supervised by Colin Johnson

School of Computer Science

University of Nottingham

I hereby declare that this dissertation is all my own work, except as indicated in the text:

Signature ____________________

Date _______ / _______ / _______

Abstract

Memories are precious, but millions of people cannot preserve them through photographs and videos - the media that have almost ruled out the development of alternative solutions. This project tackles the challenge of helping visually impaired individuals preserve and recall sentimental memories through the power of touch. Recent work in Psychology has shown that emotions play a significant role in storing and retrieving memories. An easy and accessible way to do this is recording voice memos. With the current state of Artificial Intelligence, it is possible to extract emotions from audio recordings with sentiment analysis tools. Findings in Human-Computer Interaction identified crossmodal correspondences between emotions and 3D shapes, which can be used to arouse specific emotions from tactile experience. This project explores the combination of this set of blocks to extract emotions using IBM's Natural Language Understanding API and build an autonomous algorithm for generating tactile 3D printed models, which tangibly express emotions to trigger easier recall of sentimental past events. I implement the solution as a web application considering all the accessibility requirements of visually impaired people. The algorithm effectiveness and the presentation layer were tested internally, and evaluated in a user study involving the target group. The results show potential successful associations between emotional events and generated models and a positive correlation between tactile experience and memory retrieval, but they are not statistically significant.

Acknowledgements

I would like to thank my supervisor, Prof. Colin Johnson, for his regular advice and guidance throughout the project.

I also extend my gratitude to my industry mentor, Prof. John McNamara, for his input in shaping the idea and his feedback on the project's progress.

I want to thank Mr Edgar Bodiaj and Dr Dominic Price for sharing their knowledge and 3D printing equipment, which was crucial to test the validity of my findings.

I would also like to appreciate the time that the study participants devoted to evaluate the project.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Most of our houses are filled with photographs and memorabilia reminding their residents about emotionally valuable people, events, and places. Their broadly popular forms are photographs, which we can experience exclusively through the sense of sight. According to recent estimates, more than 40 million people worldwide suffer from a complete lack of sight [5]. Out of their many challenges, this project minimises this problem by expanding their limited ability to preserve and recall memories using physical objects.

Figure 1.1: Snowmobile tracks and a landscape as an example of a hidden meaning in a photograph. [33]

Some approaches exist to solve that issue, but each has significant limitations detailed in Section 1.3. There is a shortage of accessible and affordable end-to-end solutions that could be equally or closely convenient to taking pictures.

As [33] suggests, the most valuable part of home photographs is not the image itself but the full range of experiences they recall, such as interpersonal interactions, surrounding circumstances, or events. Impartial viewers often fail to find the value of such photographs or find it in the wrong element. The reference deciphered by them in one of the studied images shown in Figure 1.1 was a beautiful landscape. At the same time, according to the owners, the real value came from the snowmobile tracks, which reminded them of a specific situation. The references, sentiments, and experiences brought by photographs have symbolic meanings. We can potentially separate them from the form of storing them, which is currently disadvantageous for visually impaired family members.

Findings in psychology suggest that emotions play a role in these elusive characteristics that help with memory preservation [35][29][30], and one of the other senses that can arouse a similar brain behaviour in this specific aspect is touch. Using autonomously generated 3D printed sculptures that help recall memories by triggering sentimental neural connections is a solution that could solve the problem. Research in Human-Computer Interaction identified "sensory substitution devices for visually impaired and deaf people" as one of the fields in computing technology that can be further explored with the recent findings in crossmodal correspondences [32].

Extracting emotions from images can be challenging and ambiguous. As shown in the above example, their sentiment is implicit. Recording voice memos is a much more reliable way of capturing the feelings experienced in a moment.

1.2 Aims and objectives

This project aims to design and develop a system for creating printable 3D models based on speech recordings and the emotions extracted from them, allowing people without sight to preserve memories in sculptures. The solution is expected to enable the target group to complete the entire process on their own, from recording a voice message to obtaining a file they can send to a 3D printing service.

The key identified objectives of the project are:

1. To develop a method for automating 3D model generation using emotions recognised by IBM Watson Natural Language Understanding tool [12].
2. To explore the relationship between emotions and shapes and apply the findings to obtain sculptures that can arouse relevant emotions and trigger memory retrieval.
3. To investigate the habits and preferences of people without sight using different devices and environments to identify the necessary tools and design a fully accessible web application integrating the functionality.
4. (Optional) To develop a technique for reversing that process, allowing for reproduction of the audio from a photo of a 3D-printed sculpture.

The requirements are further broken down in requirement specification in Section 1.4.

1.3 Related work

Some solutions already exist that allow people without sight to overcome that issue. One of the most common methods is printing graphics on manually manufactured convex surfaces [6]. However, that solution requires skilled professionals, which makes it expensive and time-consuming. It helps visually impaired people experience famous artworks but cannot replace photography. Moreover, blind people often fail to recognise facial expressions or the sexes of face sculptures [38], which can be even less successful for the 3D representation of more complex objects. Something different than explicit sculpturing of real objects or humans is needed for visually impaired people to raise emotions or convey some meaning. That is why, instead of trying to express the visual aspects, my method focuses on preserving the emotional load of a moment, making the experience more direct and lowering the likelihood of misinterpretation.

Recently, a method using convex collages of icons for representing photographs [34] has partially automated this process. These allow blind people to experience the main semantic aspects of paintings and images at a relatively low cost and short time for the first time. However, the generated graphics can only depict non-overlapping foreground elements in a fixed pose using a defined selection of icons. This solution can produce repetitive results and significantly reduces the emotional load of experiences included in the original images.

Another accessible way of preserving memories is using NFC-based sticky tags [2] that can be put on any item that brings memories. It can be a good extension for any memorabilia, but it requires a user to possess such an item in the first place. Photos are much more accessible and can be captured with a smartphone anytime and anywhere. The use case of these tags that appeals to me the most is recording a voice message when taking a picture and placing the tag with the recording on the photograph exposed at home. However, as a framed photograph is a flat surface, the memory is only recognisable through the digital voice recording, which, as a non-physical good, is generally less valued than a physical object [19].

AI Sound Sculptures aim to solve these issues. 3D printed models are already helping visually impaired people in many ways through enhancing engagement in Braille alphabet learning [40], building adaptations on everyday objects [25] and making interaction with mobile touchscreens more accessible [48] and more solutions reviewed in [24]. The unique form of a memory sculpture would allow a visually impaired person to recall a specific experience. As voice recordings can be captured with a smartphone, this new method is closely convenient to photography. The ownership that its physical form creates gives it an advantage over digitally stored voice recordings and allows users to display them at home as others do with nostalgic photographs. The cost-effectiveness is slightly lower than the traditional approach, but consumer 3D printers and printing services are becoming more common and affordable.

1.4 Description of the work

The first software product developed as part of this project is an algorithmic solution to modelling emotions based on voice recordings. It consists of the three following stages:

- Waveform processing: Building the sculpture's base using the waveform extracted from audio recordings makes the output models more personalised and unique. I described this stage in detail in Section 3.1.
- Emotion extraction: Using Speech Recognition and Natural Language Processing to perform sentiment analysis on the audio input (Section 3.2).
- Emotion shaping: Forming the top of the sculptures (Section 3.3) using the emotion extraction results and the academic findings exploring the associations between emotions, shapes, touch and memory presented in Chapter 2.

The second software product built for this project is a website, which lets the users access the sculpture generator (Section 3.4). Its primary focus lies in minimising the challenges that blind people encounter when navigating the internet.

Evaluating the products is separated into two stages explained in Section 5.1 and Section 5.2. In the first one, I test the technical reliability of the software by experimenting with various inputs and measuring the system's response time. The second stage focuses on the usability of the end product and the demand for it. It involves a study in collaboration with an organisation uniting visually impaired people. It aims to assess the target group's general interest and the solution's ability to achieve its goals. Each element of all the evaluation stages aims to satisfy a specific requirement from the list below.

The website will allow its users to use the following functionalities:

1. Easily navigate the website using screen reading software.
2. Upload their voice recording.
3. Download a file containing a printable 3D model.

The entire system should comply with these non-functional requirements:

1. Accessibility
   (a) The website will be designed in a way that doesn't require a blind user to seek assistance.
2. Quality
   (a) The software will be robust and fault-tolerant.
   (b) The website will be responsive to different device types commonly used by the target group.
   (c) The generated sculptures are able to arouse emotions that are relevant to the provided recording in a tangible way.

Chapter 2

Methodology

In this chapter, I describe the research foundations, which form the base for the design, as well as studies that justify the assumptions that make the general idea of this project scientifically valid.

2.1 Memory and recall

The common factor of the events people tend to capture in photographs is the sentiment evoked at the time of taking it. The preserved emotions play a significant role in the process of both storing and retrieving memories. Two brain memory systems are involved: the amygdala - responsible for emotional memories gathered unconsciously - and the hippocampal complex - consolidating conscious memories of facts, events and experiences into long-term memory. Despite functioning independently, "they act in concert when emotion meets memory" [35]. In addition, it impacts memory longevity positively [29].

Apart from the influence on their storage, another study has shown that the retrieval process of autobiographical memories can also be affected by emotions. More specifically, recalling some memories can be triggered by putting oneself in a specific mood [31]. According to a study of long-term memory storage of perceptual experiences [30], the durability of visual and haptic memory is comparable. That implies a possibility to replace photography as a medium with a tactile sculpture. Embedding the emotions accompanying important moments inside 3D printed models can help to retrieve them and make the memories accessible to everyone regardless of their vision condition.

2.2 Emotion-touch association

Crossmodal correspondence is an extensive research area in Psychology and Human-Computer Interaction. One of the most widely recognised concepts in the field is the "Bouba/Kiki" paradigm. People tend to strongly associate the word "Bouba" with the bulbous shape presented on the right in Figure 2.1 and "Kiki" - with the jagged shape on the left. Although this phenomenon used to be considered "fakery or simply [based on] memories" before, we now know it is a result of synesthesia - the interaction of different brain areas that are usually not acting together "to make abstract connections between seemingly unrelated inputs" [36]. Although most of the further research explores the social aspect of affective touch, which explores "brain mechanisms underlying social touch processing" [28], the emotional impact of touching different shapes was studied too.

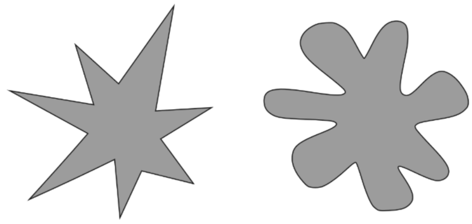


Figure 2.1: "Kiki" (left) and "Bouba" (right) shapes. [32]

A study by Olesya Blazhenkova and Melisa Maya Kumar measured 14 emotions aroused by eight 2D shapes in angular and round variants [21]. These are analysed and used to design the physical features of different emotions in Section 3.3. Apart from matching shapes to written emotion labels, the study also measured the correspondence between the shapes and photographs showing real experiences of different emotions.

More recently, researchers focused on the haptic experience of touching 3-dimensional sculptures based on "Bouba" and "Kiki" features. The findings of [32] showed that people associate complex and angular tangible stimuli with excitement, lack of control and unpleasant emotions. Simple and round structures, however, make us experience calm and pleasant emotions. The authors emphasise that their findings about 3D shapes contrast with previous effects observed with 2D shapes. Although that was only discovered for colour-shape correspondences, I assume that the studies involving 2D shapes should be treated with more uncertainty.

Symmetrical and round shapes are associated with pleasure, while asymmetrical and angular - correlate with the feeling of threat. The identified crossmodal correspondences are culturally independent [43].

The texture of objects is another significant element of brain activity related to tactile experiences [42]. Plastic surfaces have the lowest arousal score compared to other materials, such as metal or cotton [26]. Since my sculptures are made of plastic, there is no risk of a potential conflict of the emotions aroused by their texture with those aroused by their shape.

Chapter 3

Design

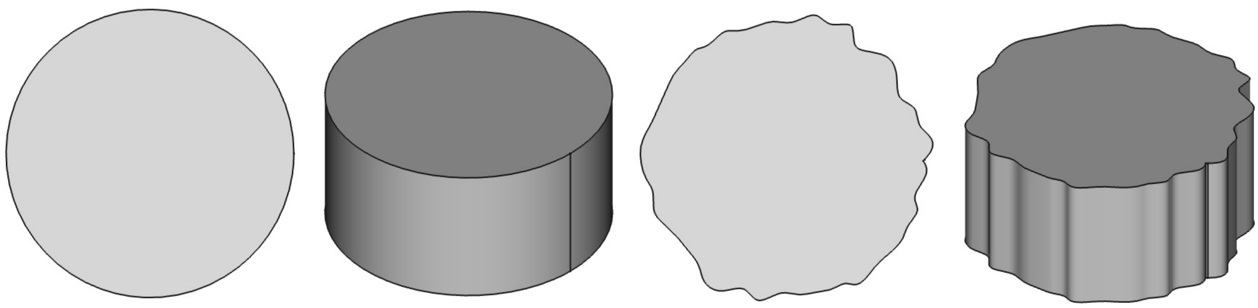
In this chapter, I discuss the design chosen for each of the project components starting from shaping the model's base using audio waveforms, through emotion extraction from voice recordings, to generating the model's upper body. I also explain how my choices are motivated, what issues they address and how the solution is delivered to the end-user.

There are many possible ways to shape the sculptures. However, the described design aims to represent emotions tangibly while keeping the models unique and minimising complications with modelling software and printing.

3.1 Waveform processing

The first stage adds uniqueness to the sculpture making it more personal. That will allow two different sculptures that convey similar emotions to be more distinguishable and retrieve relevant memories.

Initially, it loads a simple cylindrical object as presented in Figure 3.1a and 3.1b. The shape is then deformed using the waveform extracted from a voice recording. I achieve that by applying a set of processing techniques that extract amplitude information from the recording and fit the data to a range and length expected by the model generator. The process is explained in detail in 4.2.2. The resulting deformations are then smoothed, as



(a) (b) (c) (d)

Figure 3.1: (a) The base and (b) the sculpture in isometric view in its initial shape. (c) The base and (d) the sculpture in isometric view after deformation.

shown in 3.1c and 3.1d, to eliminate issues with printing preprocessing and removing support structures in post-processing. Even though these operations make the recovery of the audio recording from the structure's shape impossible, not implementing them would additionally result in a jagged base shape influencing the emotions aroused in tactile experience too strongly. It would also not be possible to recreate the audio file from such small sculptures, as the printing accuracy of low-cost methods is too low.

3.2 Emotion extraction

Implementation of emotion extraction uses two IBM APIs: Watson Speech to Text (STT) [11] and Watson Natural Language Understanding (NLU) [12]. The former takes audio recordings as the input and returns text in natural language as the output. The latter can perform many NLP-based operations on text, but the only subcomponent necessary for this project is Tone Analytics (TA), which performs sentiment analysis on the input. Training such models requires large datasets and powerful resources, which could only be possible with a large trained model, like the one provided by IBM. The outputs for an example audio input are presented in Table 3.1.

Table 3.1: Outputs from IBM tools for example audio input with transcription errors in bold.

<table>
<tr><th colspan="7">Input (manually transcribed)</th></tr>
<tr><td colspan="7">We went to Málaga with my friends and we had wonderful time there. I remember that we went to the beach in the night and we swam in the sea and the water was warmer than the air. We were laughing and playing funny games on the shore later. I felt happy to be there and it is a memory that keeps me motivated and makes me grateful for the friends that I have.</td></tr>
<tr><th colspan="7">Output from Speech to Text</th></tr>
<tr><td colspan="7">we went to mallago with my friends and we had wonderful time there i remember that we went to the beach in the night and we swam in the sea and the water was warmer than the year we were laughing and playing funny games on the shore later i thought happy to be there and it is a memory that keeps me motivated and makes me grateful for the friends that i have</td></tr>
<tr><th colspan="7">Output from Tone Analytics</th></tr>
<tr><td>excited</td><td>satisfied</td><td>sad</td><td>polite</td><td>sympathetic</td><td>frustrated</td><td>impolite</td></tr>
<tr><td>58%</td><td>15%</td><td>11%</td><td>5%</td><td>2%</td><td>2%</td><td>0.1%</td></tr>
</table>

IBM's STT is one of the most popular Automatic Speech Recognition systems publicly available. Its architecture is not published, but it achieves satisfying results for the intended use case [27].

Although the architecture of IBM's method used for TA is neither officially published on their websites, a research paper released around the same time as the tool introduces a solution resembling its characteristics significantly [47]. It uses Tone Latent Dirichlet Allocation (T-LDA) to assign continuous probabilities to each of the following emotions: anxious, excited, frustrated, impolite, polite, sad, satisfied, and sympathetic. Latent Dirichlet Allocation is an unsupervised probabilistic generative model, which identifies topics in documents based on word distribution in each topic [22]. T-LDA extends that by introducing probabilistic sentiment for each word, which allows it to return continuous values as the output for each emotion. Although the primary purpose of IBM'a TA is analysing customer reviews, it has been applied to problems in different disciplines such as judicial decisions [44], music lyrics [18], and scientific writing [41]. It can, therefore, determine the emotional tone in a broader, general context.

Table 3.2: Text input used for matching emotions from [21] to Tone Analyser's outputs.

Emotion	**Input to Tone Analyser**
Pride	I felt an overwhelming sense of pride as I watched my daughter receive her diploma on stage. She had worked tirelessly for years to get to this moment, and seeing her achieve her goal was a moment of pure pride.
Joy	The feeling of joy was infectious as my family and I danced and sang along to our favorite songs at a concert. It was a moment of pure happiness that I will never forget.
Amusement	I couldn't help but laugh as my friend recounted a funny story about their weekend. Their animated gestures and comical delivery had me in stitches.
Pleasure	The first bite of the deliciously creamy cheesecake brought an immediate wave of pleasure. I savored every bite, enjoying the rich, smooth texture and sweet, tangy flavor.
Relief	As the doctor walked in with the test results, I felt a weight lifted off my shoulders. The relief of knowing that my health was not in danger was overwhelming.
Interest	The intricate details of the sculpture caught my eye, and I couldn't help but feel a deep sense of interest. I spent hours exploring the nuances of the piece, marveling at the skill and creativity of the artist.
Surprise	As I opened the door, I was greeted by a group of friends shouting 'Surprise!' My heart raced as I realized they had thrown me a surprise birthday party. It was a moment of shock and joy that I will always treasure.
Anxiety	My heart raced as I stepped on stage to deliver my speech. The anxiety of speaking in front of a large audience was almost overwhelming, but I took a deep breath and began to speak.
Fear	As I stepped into the elevator, I felt a wave of fear wash over me. The enclosed space triggered my claustrophobia, and I struggled to calm my racing thoughts.
Disgust	The sight of the overflowing trash bin made me feel disgusted. The pungent odor and unsightly mess made me want to leave the area as quickly as possible.
Sadness	Tears streamed down my face as I hugged my friend goodbye. We had shared so many memories together, and the thought of them leaving filled me with deep sadness.
Despair	As I sat in my car, staring at the 'For Sale' sign on my business, I felt a sense of despair wash over me. After years of hard work, the business had failed, and I wasn't sure what my next steps would be.
Irritation	The constant beeping of car horns in traffic irritated me to no end. I could feel my temper rising as I sat in the backed-up traffic, wondering how much longer it would take to get home.
Anger	As I read the news article detailing the latest instance of police brutality, I felt a wave of anger wash over me. The injustice and senseless violence filled me with a burning fury.

3.3 Emotion shaping

This design stage focuses on encapsulating the identified emotions in 3D models. I do this by analysing the findings described in Section 2.2 with respect to the emotion set returned by IBM's Tone Analyser (TA). I dismissed the outputs for *polite* and *impolite* classes, as they don't contain any explicit emotional load that could be associated with shapes based on the literature.

3.3.1 Literature analysis

Since the emotions studied by [21] differ from the model's output, they required some mapping. I achieved it by creating a separate short story expressing each of the emotions explored in the mentioned study presented in Table 3.2. The text was then passed into TA, and the results, shown in Figure 3.2, indicate which of the emotions available in TA are the most relevant.

I further analysed the crossmodal correspondences established by Olesya Blazhenkova and Melisa Maya Kumar by extracting the shapes associated with the relevant groups of emotions in Table 3.2. This resulted in obtaining

Model output	Studied emotions
Excitement	Surprise Interest Amusement Joy
Frustration	Irritation Anger Despair Disgust
Sadness	Sadness Despair
Satisfaction	Pride Joy Pleasure Relief
Sympathy	Sadness Interest

Figure 3.2: Mapping of the emotions from [21] to Tone Analyser's output.

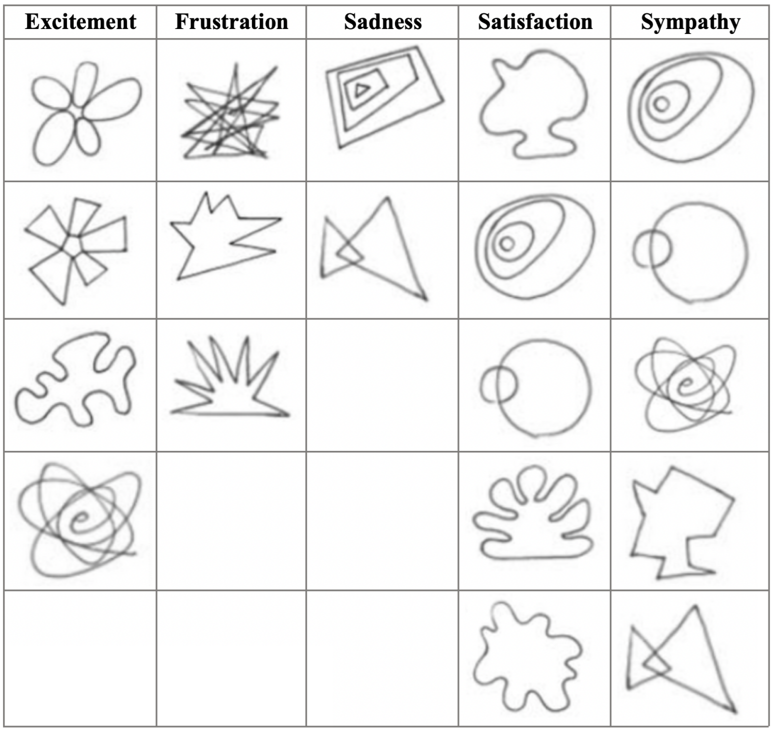


Figure 3.3: Visual associations between 2D shapes and emotions from Watson Tone Analyser's output based on analysis of the findings from [21] sorted by relevance.

Figure 3.3 with shapes from [21] assigned to each of TA's output emotions with the most relevant on top.

The range of emotions studied in [32] is smaller and easier to match to the emotions explored in my study. It helped me to draw the following conclusions:

- *Excitement*, *frustration* and *sadness* fall into the "unpleasant/excited" category of the studied associations and, therefore, should be expressed with high angularity and complexity also associated with lack of control.
- *Satisfaction* falls into the other side of the spectrum as a pleasant and calm emotion and can be associated with control, low complexity and roundness.
- *Sympathy* falls in neither of these extremes and positions itself around the middle of the spectrum.

Symmetry constraints of *satisfaction* are motivated by its pleasantness. Unpleasant emotions that are associated with the feeling of threat, like *frustration* and *sadness*, are expressed in asymmetrical shapes [43].

3.3.2 Modelling System

I identified recurring patterns and dependencies from the assumptions above to create a set of rules forming the sculptures. The three most confident emotions from TA's output are passed further for modelling. If the *confidence* value of any of these emotions is lower than 15%, they are replaced with the most confident one. The same happens to emotions which *confidence* value is lower by more than 30 percentage points compared to its immediate predecessor in a dictionary sorted by *confidence* in decreasing order.

Most of the modelling characteristics of a specific emotion are expressed inside a *layer*. Each *layer* contains the following parameters:

- *type* - describing the type of modelling that will be performed (*1* or *2*),
- *confidence* - describing the confidence of TA's prediction of this specific emotion,
- *points_num* - describing the number of points located circularly on the plane of the sculpture's base,
- *radius* - describing the radius for generating the points,
- *deviation_range* - defining the range of horizontal deviation that will be applied to the generated points,
- *symmetry* - defining whether the points distribution across the plane will be symmetrical.

Other parameters are specific to *layer's* type. *Sadness*, *satisfaction* and *sympathy* set the *layer* type to 1. These emotions are calmer than *frustration* and *excitement*, and, therefore, they are shaped using the base of vertically extruded polygons with low starting complexity [32]. Tactile interaction with these layers is focused on the edges without immediately sending arousal signals by the *layer's* flat and relatively boring form. The visual association of inward nesting with these emotions has been inferred from [21] and shown in Figure 3.3. These *layers* are defined by these additional parameters:

- *vertex_fillet* - defining the level of smoothing applied to the vertices of the resulting 2D polygon plane,
- *edge_fillet* - defining the level of smoothing applied to the edges of a resulting 3D structure,
- *bot_fillet* - defining the level of smoothing applied to the surface underneath the resulting 3D structure.

Frustration and *excitement* set the layer type to 2. Such layers are formed using sets of multiple small geometric shapes protruding from an underlying plane. Their complexity immediately triggers arousal [32] before any other element narrows the brain's emotional response to a specific sentiment. Their visual associations with multiple distinct deformations in Figure 3.3 also motivate the design decision of constructing many separate forms. These *layers* are defined by the *polygon_range* parameter, which determines the range of the number of points used to form the regular polygon base of each small prism.

Each sculpture is built using two or three separate *layers* of possiblt various types. They are sorted by the *confidence* parameter. The *radius* parameter is decreased in each successive layer to avoid overlapping of the structures. The *points_num* parameter follows the same pattern making the model more diverse and the smoothing operation more straightforward. Type 1 *layers* are stacked onto each other vertically. In contrast, type 2 *layers* are circularly placed onto a plane without vertical progression, where the middle part is left empty and can be filled with the next *layers*. An example of applying type 1 layers on a model base is presented in Table 3.3.

The height of type 1 *layers* is multiplied by $1 + round(\frac{confidence}{4})$ if their *confidence* exceeds 50%. For type 2 layers, the height of each of the small structures is randomly selected between the value of 0.14 and the result of $\max\left(\frac{1}{2} \times \mathit{confidence}, 0.15\right)$ and multiplied by a global height parameter. The heights are additionally adjusted

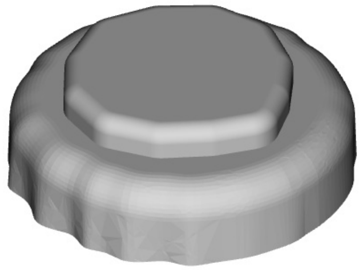
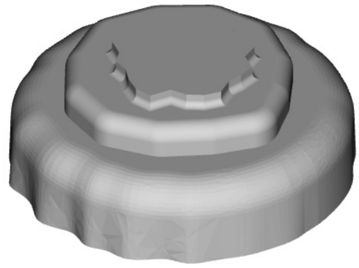
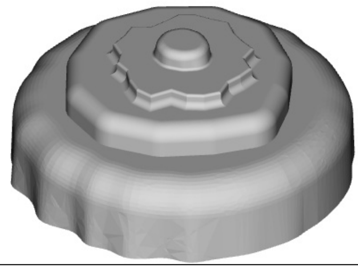
Layer Number	Emotion	Layer Parameters	Model
1	Satisfied	*type*: 1 *confidence*: 0.61 *points*_num: 12 *radius*: 30.0 *edge_fillet*: 6.0 *vertex_fillet*: 12.0 *deviation_range*: 2.2 *symmetry*: True *bot_fillet*: False	
2	Sad	*type*: 1 *confidence*: 0.48 *points_num*: 26 *radius*: 21.0 *edge_fillet*: 0.0 *vertex_fillet*: 0.0 *deviation_range*: 1.0 *symmetry*: False *bot_fillet*: True	
3	Sympathetic	*type*: 1 *confidence*: 0.32 *points_num*: 8 *radius*: 9.0 *edge_fillet*: 6.0 *vertex_fillet*: 6.0 *deviation_range*: 1.0 *symmetry*: False *bot_fillet*: True	

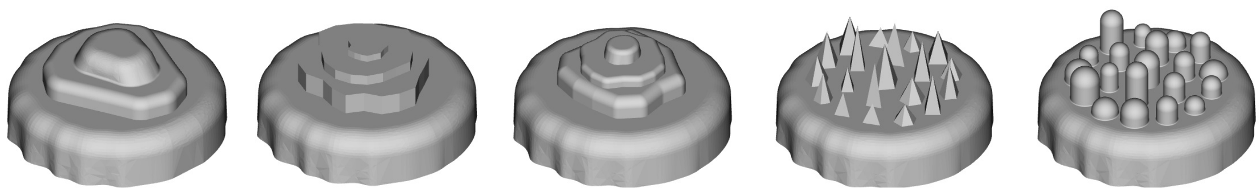
Table 3.3: Example of consecutive layers' construction with emotions and parameters.

based on the layer's level. These modifications will make more confident emotions impact a larger proportion of the final models.

Shape implications for each emotion are described below, and example models created using each single emotion for all the layers are shown in Figure 3.4.

Shaping satisfaction

This emotion is mainly associated with simple, round and symmetrical forms evoking a sense of control, calmness and pleasantness [32][43]. To contain it in the *layer*, I set *points_num* to a random, relatively low value between 12 and 15, both the fillet parameters to the possibly highest numbers and *symmetry* to *True*. The *layer* should



(a) (b) (c) (d) (e)

Figure 3.4: Models generated for (a) satisfaction, (b) sadness, (c) sympathy, (d) frustration, and (e) excitement, where all the layers in a single model are constructed using the same emotion.

also incorporate some ample smooth curves, as shown in the top shape in Figure 3.3. That effect can be achieved by increasing the deviation range, which - with low *points_num* and *symmetry* set to *True* - adds some deformations while keeping it simple and satisfying.

In addition, if *satisfaction* is among the three selected emotions, it can influence other *layers*. Type 1 *layers* have their *bot_fillet* set to *True*, which adds a satisfying factor to it while keeping the primary emotion dominant. In type 2 layers, the *symmetry* parameter is set to True, which does not significantly change the main sentiment of *frustration* and *excitement*, but it adds an extra input. These two design decisions are motivated by the fact that *satisfaction* cannot be created as a third layer because of its high *deviation_range*.

Shaping sadness

Although *sadness* is not generally linked to high arousal, it is an unpleasant emotion. It should be represented with complex and angular edges [32] with asymmetrical elements in the resulting structure [43].

To achieve high complexity, *deviation_range* is set to 1.0 and *symmetry* to *False* to make the polygon irregular. *Points_num* parameter is set to a high random value between 17 and 35. However, its *confidence* values exceeding 50% result in multiplying *points_num* by $1 + round(\frac{confidence}{4})$. To increase angularity, *edge_fillet* and *vertex_fillet* are set to 0. Additionally, if *sadness* is a dominant emotion, it enforces *bot_fillet* to be set to *False* in the first layer, which maximises the angularity.

Shaping sympathy

Cross-modal correspondences of *sympathy* and its associations with arousal and pleasantness are not evident, which makes it more problematic than the others. Given its position in the middle of the spectra, its shape characteristics include some elements from both the extremes. I set *points_num* to a value between 10 and 25. High fillet parameters of 6.0 are counterbalanced by asymmetry and high deviation resulting in moderate arousal and valence.

Shaping frustration

Frustration is the most unpleasant emotion linked to the highest arousal. I achieved high angularity by placing multiple pyramids on a bottom plane. That design resembles the visual associations in Figure 3.3. Their regular polygon bases have different, random numbers of vertices ranged by values 3 and 6 declared in *polygon_range* parameter. Their individual radii range from 8% to 14% of the entire model's radius. These adjustments increase the model's complexity. The number of prisms, defined in *points_num* parameter, is based entirely on *confidence* value and ranges from 10 for *confidence* below 50%, to 13 if it exceeds 80%, expressing intense frustration with even more complex result. As in other emotions, it decreases with *layer's* level, but it's always set to 1 for the third level. *Deviation_range* is set to 0 because of implementation issues described in Section 4.2.5. The symmetry of their sizes is set to *False* to increase its tactile association with threat [43] unless the appearance of relatively strong *satisfaction* changes it, as mentioned above.

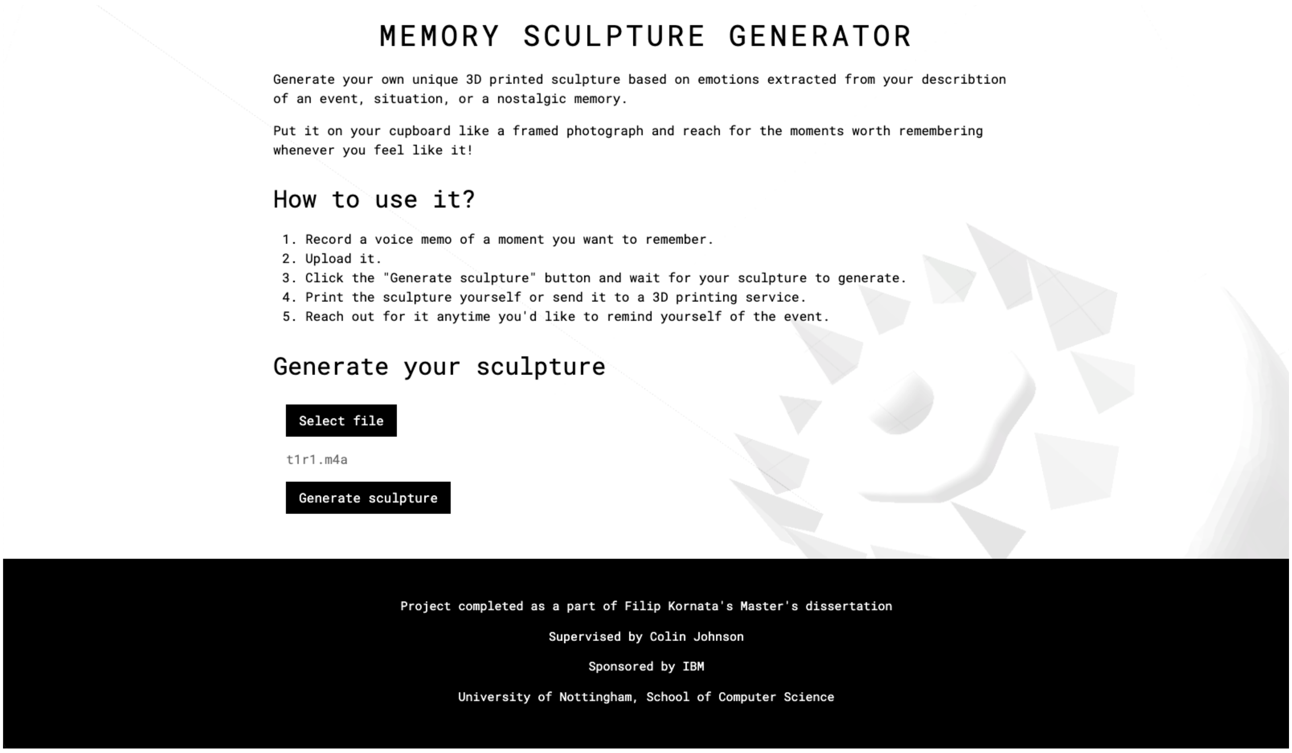


Figure 3.5: Website design.

Shaping excitement

Although [43] associated *excitement* in the high complexity and angularity category, these two factors were not separated in the study. Following the findings in [21], as far as the correlation is positive for complexity, it is the opposite for angularity. The latter decreases the pleasant experience significantly more than the former [39]. Therefore, the process of shaping this *layer* follows the same process as shaping *frustration*, with the exception of using circles as a base for the small prisms and ending them with a hemisphere on the top. That keeps the required complexity while making its tactile experience more pleasant. Symmetrical shapes increase arousal but decrease valence [45]. Since *excitement* is associated with high coefficients of both of them, *symmetry* is determined by the presence of *satisfaction* among the top output emotions. Additionally, it implements the *bot_fillet* functionality under specific circumstances limited by the challenges discussed in Section 4.2.5.

3.4 Website

The main aim of the website design is simplicity and accessibility. It includes the title, a short project description, a short user manual, and two buttons for uploading an audio file and downloading an STL model file. It only contains one page, shown in Figure 3.5.

I designed its visual aspect concerning the experiences of people using screen reading software studied by WebAIM in [15], which mentions website design as the aspect with the biggest impact on accessibility. In recent years, assistive technology has been helping blind people to browse the web. Screen readers, the most common assistants, scan and read websites' content aloud. Their users tend to use websites more often than mobile applications. For that reason, the front end is implemented as a web application.

Heading levels were identified as very useful by most study participants. The website uses a consistent, well-structured heading system starting from the project title as a level 1 heading. Then, a level 2 heading follows for a section explaining the process required to generate a model. The third one is also a level 2 heading and indicates the sculpture generation section, where the audio file is uploaded and the model is downloaded.

WebAIM provides an additional document highlighting important principles of accessible design [16]. I identified the relevant elements and, with the help of Web Content Accessibility Guidelines (WCAG) [14], designed the following solutions:

- Well-planned heading structure addressed above.
- Logical reading order is implemented by starting with a short project description followed by a list of steps required to complete and the final section where the actions are made.
- Good contrast is achieved using two colours: black and white.
- Adequate font size of a minimum of 16px was selected, and the content can be easily magnified by 200% thanks to relative units used in styling.
- Keyboard navigation with a screen reader is simple because of the use of HTML tags like *form*, *section*, and *div*.

Accessible Rich Internet Applications (ARIA) [1] labels are used as an additional facilitation for the target group, allowing screen readers to read some extra information aloud like "Select a file from your device" when focusing the "Select file" button. The website's language is specified in code to ensure screen readers know what to expect.

Chapter 4

Implementation

4.1 System Architecture

All application functionalities are executed on the server side, consisting of the back-end and front-end. The back-end functionality generates a 3D model based on audio recordings, and it is split into three stages: waveform processing, emotion extraction and model generation. The front-end side contains a simple website that makes the software easily usable. The system architecture presented in 4.1 shows this process including the flow of information.

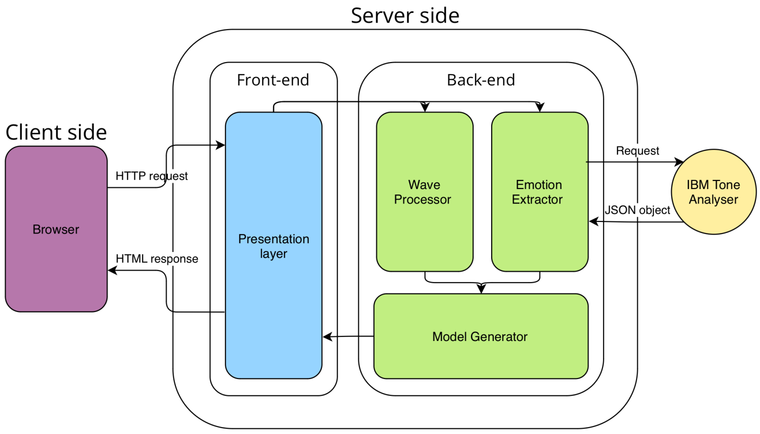


Figure 4.1: High-level system architecture diagram

The cycle starts and ends with user interaction. To initialise the process, a user must upload their file and click a button to generate a model. The audio file is then passed to *Wave Processor* and *Emotion Extractor*. The former returns a processed array that can be used to form the base, and the latter returns a dictionary of emotions and their confidence values. These outputs form the input of *Model Generator*, which turns them into a file containing a model returned for the user to download.

4.2 Back-end

4.2.1 Overview

Python is the programming language I used for the back-end side. This decision is motivated by my previous experience with it and its wide selection of libraries containing all the necessary tools. As a reputable and widely-used language, it also ensures the software will be easily maintainable and extendable, should the need arise. I commented the code for each implementation stage, and added function and class descriptions using Google docstrings style to make it easier to understand. I tracked and managed changes in the project using Git.

I used object-oriented design to create a modular, maintainable and extensible codebase. Each back-end element is implemented inside a separate class, each utilising a number of libraries specific to the needs. All the functionalities are called from *app.py* script, which uses Flask to communicate with the front end. When a user presses "Generate sculpture" button on the website, the function presented in Listing 4.1 is executed. This and many lower-level functions use abstraction to make them easier to understand. The reason why *Emotion Extractor* does not take the same input as the *Wave Processor* in practice is discussed in Section 4.2.2.

```
def upload_file():
file = request.files['file']
# Process audio and return audio array and the path to WAV file
p_audio_array, wav_path = process_audio(file)
# Extract emotions from the audio recording using IBM Speech-to-text and IBM Natural Language
    Understanding tools
response = extract_emotions(wav_path)
# Generate the sculpture
outputfile = generate_sculpture(response, p_audio_array)

return send_file(outputfile, as_attachment=True)
```

Listing 4.1: Main Python script executed on file upload.

4.2.2 Waveform processing

This stage performs various numerical and audio operations to extract the waveform array from a file. The audio metadata, including the sample rate, duration and frequency, and sample array, are extracted using pydub [37] library, which ensures adaptability to different audio formats like MP3, WAV, FLAC, M4A, AIFF and AAC.

That information is then used to perform the processing with NumPy [9] library. It starts with reducing the length l of the audio array a using Algorithm 1 to match it with the number of deviations in the sculpture's base p and produce the array a_p of length p. I calculate each value of the new array a_p using the average of n consecutive samples from a. After doing so, I normalise the array to the scale resembling the deflections of the circular sculpture base.

Although that approach was working at first, it produced modelling errors and weirdly shaped deflections for some specific recordings. I created additional functions that allow to plot or listen to the processed recordings using Matplotlib [7] and PyAudio [10] libraries. After some experimentation, I found out that the issue was caused by outliers produced by background noise or touching the microphone when capturing audio. To avoid these problems, I handle the outliers by replacing them with the mean a_p value using Algorithm 2. I tested different values of constant m and finally set it to 7.

Algorithm 1 Reducing the length of the audio array

Require: $length(a) \geq p$
  Input: audio array a, deviations number p
  $n \leftarrow \lfloor\ length(a)/p \rfloor$
  for $i \leftarrow 0$ to $p-1$ **do**
    $a_p[i] \leftarrow \frac{1}{n} \sum_{j=i\times n}^{(i+1)\times n-1} a[j]$
  end for
  Output: a_p

Algorithm 2 Imputing mean values for outliers

  Input: audio array a_p, constant m
  $d \leftarrow \mid a_p - median(a_p) \mid$ ▷ Calculate the absolute difference between each element and the median.
  $b \leftarrow median(d)$ ▷ Calculate the median of the absolute differences
  if $b \neq 0$ **then**
    $c \leftarrow \frac{d}{b}$ ▷ Calculate the modified z-score for each element
  else
    $c \leftarrow$ array of zeros with the same size as a_p ▷ Set c to an array of zeros
  end if
  for $i \leftarrow 0$ to $length(a_p)$ **do**
    if $c[i] \geq m$ **then**
      $a_p[i] \leftarrow mean(a_p)$ ▷ Replace $a_p[i]$ with the mean of a_p
    end if
  end for
  Output: a_p

The functionality of *Wave Processor* is extended with a function which compresses and exports the audio file to MP3 and stores it locally to allow the *Emotion Extractor* to function correctly.

4.2.3 Emotion extraction

As mentioned in Section 3.2, two IBM APIs are used to perform emotion extraction: Speech to Text (STT) [11], and Natural Language Understanding (NLU) [12]. Both use the standard API call procedure, where an HTTP POST request is sent to IBM's endpoint, and a JSON response is returned. The API keys are stored in a separate config file to ensure confidentiality.

Both the APIs use Transport Layer Security (TLS) for securing the connection between the client and server. The IBM service supports European Union General Data Protection Laws (GDPR) to ensure information security.

STT tool initially introduced some issues with the input files, but it works properly after the compression to MP3.

4.2.4 Parametric Modelling

My initial selection of the 3D modelling tool, Blender Python API [3], is an excellent manual model production tool, but it is not perfectly suited for working autonomously only based on numerical parameters, which is a crucial element of this project. It is considered to be more complex and requires more time to learn. The final choice of software, CadQuery [17] Python library, is designed explicitly for parametric modelling, and its learning curve is more reasonable considering the duration of this project. It allows to load, manipulate and export models to STL format easily with Python.

CadQuery aims to allow its users to build models similarly to how people describe objects. It usually starts from defining a *Workplane* object and using method chaining until its expected shape is achieved. An example in Listing 4.2 starts with selecting a plane where the operations will be performed. Then, it creates a box with provided parameters, selects the plane to modify and executes *hole()* function to put a hole inside.

```
result = Workplane("front").box(width,width,thickness).faces(">Z").hole(thickness)
```

Listing 4.2: Example of a rectangular prism with a hole in the middle created with CadQuery.

Although the library allowed to deliver the project entirely, it caused many issues on the way due to its strict environmental requirements, poor documentation, small user base and its purpose for the production of industrial parts rather than expressing emotions. The challanges are further discussed in Section 4.2.5.

Three crucial parameters must be defined to start building any model:

- r, which defines the radius of the base circle before deformation,
- h, which setting the sculpture's height, which is rather an approximation given the various possible forms of it,
- *p_num*, which is the number of deformations in the base that was used for wave processing before.

The bottom part

The base of the sculpture can be formed in four different ways, with each next executed upon failure: using the full audio array returned from *Wave Processor*, using an array with the last element removed, using one with two last elements removed, and using a random array. Although it never reached the random array generation during tests, it is a safety measure for potential incompatibility. To construct the deformed circular base, I calculate each point using Algorithm 3. It finds the position of each successive point in a circle and deforms it using the processed array a_p.

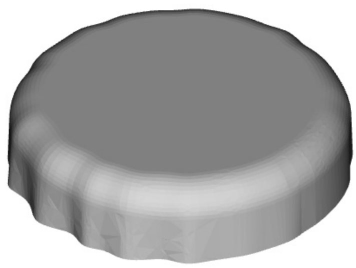


Figure 4.2: The bottom part of an example sculpture.

Algorithm 3 Forming the bottom base

Input: processed audio array a_p, number of points p, radius r
for $n \leftarrow 0$ to $p-1$ **do**
  $d \leftarrow a_p[n]$
  $points[n] \leftarrow ((1+d) \times r \times \cos(\frac{(n+1)\times 2\pi}{p}) - r, (1+d) \times r \times \sin(\frac{(n+1)\times 2\pi}{p}))$
end for
Output: $points$

The points are then smoothly connected by a line to form a closed plane. A similar plane is created at the height of half the h parameter, and they are connected to form a 3D structure shown in 3.1. Its upper edge is rounded to create a shape shown in 4.2.

The top part

Generating the rest of the sculpture starts with setting all the layer parameters, as explained in Section 3.3.2. After all the necessary information is contained in a dictionary data structure, the function loops through all the layers and calls a function relevant to the layer type.

Type 1 layer generator function creates a *deviation_array* based on *deviation_range* and *points_num* layer parameters. It is then used along with *radius* and *symmetry* parameters to get the exact locations of the points that will form a polygon base. The exact function responsible for that will be described later in this section. When the points are formed, their 2D shape is smoothed based on *vertex_fillet* parameter. Then it is extruded to relevant height based on *confidence* and layer level, and *edge_fillet* is applied to smooth the top edges. If possible, *bot_fillet* smooths the structure's contact points with the bottom plane.

For non-final layers, the procedure of shaping type 2 layers begins with generating *heights* and *sizes* arrays based on *confidence*, *points_num* and *symmetry* parameters. The process follows the same procedure as type 1 layers to create *deviation_array* and find the locations where multiple small prisms are placed. A relevant function is then called to build the round or spiky elements based on the created arrays, layer's level, and *polygon_range* parameter for the latter. For the final layers, the arrays are unnecessary because only a single shape is created in the middle of the plane with an increased polygon's or circle's *radius* and *height* parameters.

The widely used function for getting polygon points in the modelling space is based on the solution proposed for the generation of the bottom base (Algorithm 3). However, this time *polygon_p_num* parameter is responsible for the number of points similarly distributed across the circumference of a circle. They can be deviated similarly to how it was done before. Large asymmetric deformations may alter the centre of mass of the sculpture. The points are, therefore, moved towards it after the deformation to avoid this issue. An additional parameter of *symmetry* is considered for this algorithm. Setting its value to *True* results in using only the first half of the deviation array, which is then mirrored for the negative side of the Y axis.

4.2.5 Challenges

Although I successfully implemented the project in the end, some challenges caused numerous delays during the development process. Issues with the NLU API and CadQuery caused some of the most significant ones.

Initially, I was not planning to use the TA component of NLU API, but a slightly different sentiment analysis tool called Watson Tone Analyser. However, it was discontinued in March 2023 and replaced with the TA component inside NLU API. The issue was that the new alternative used a different set of emotions than the old solution, and the literature analysis in Section 3.3.1 was prepared with respect to the emotions of the terminated tool. I solved it by finding a relevant model ID, *tone-classifications-en-v1*, which allowed me to perform the analysis with the expected output emotions.

Some other challenges occurred during the development of the modelling algorithm. CadQuery is best suited for Linux and Windows users with x86 processor architecture. For this reason, as a user of an ARM-based Mac, I struggled with many environmental problems. Initially, I developed the modelling algorithms using FreeCad [4], a software which allows to add CadQuery as a module and use it within the tool's environment. Although this solution works correctly on ARM processors, it uses an outdated CadQuery version and has limitations for installing additional libraries, which caused problems with implementing crucial features. With minimal options, I further developed the project on a public Lab machine with virtualised Ubuntu until I managed to install an emulated x86 version of Linux on the Mac. Ultimately, I found a solution that emulates x86 architecture inside a Miniconda [8] environment, which shows better performance and is more convenient than emulating the entire operating system.

Other issues had a major influence on the final design of the sculptures. CadQuery documentation lacks information about the parameters and specific use cases of more advanced functions, even though the general and widely-used concepts are clearly explained with suitable examples. The small user base and the unpopular purpose that the tool is used for in this project made it challenging to find relevant information on online forums. It made me struggle with implementing overlapping structures or finding the boundaries of irregularly shaped planes. The latter caused most issues with defining ranges, e.g., *deviation_range*, smoothing edges and the general placement of objects. Even though I solved most of them through experimentation and bypassing, such as keeping track of the exact height at which the next element can be placed or using preset parameters like the model's radius to calculate the positioning of a new object, some of them required trade-offs. The uniqueness and large diversity between the models could only be achieved by automatically trying to generate sculptures with minor differences in parameters until a viable solution is found. Every tested example managed to find it, but some of them required more time as measured in Section 5.1.

The last issue refers to public access to the final product. I tried AWS and Google Cloud App Engine hosting platforms to accomplish that. However, none offered access to install dependencies using Conda package manager. As Pip's version of CadQuery is strictly dependent on FreeCad, using Conda is the only way to use the latest version of the library. That issue can potentially be solved in the future using a Docker container, which would allow to install all the required dependencies.

4.3 Front-end

Web technologies used for implementing the front end include:

- HTML for structuring the interface of the website,
- CSS for styling the interface,
- JavaScript for dynamic and interactive elements,

Additionally, I used W3.CSS framework [13] to make the website design modern, responsive and usable on mobile devices, and Google Fonts to style the text.

The website is hosted without the back-end functionality using GitHub Pages under the domain https://filipkor.github.io.

4.4 Printing

The most probable use case scenario for this stage is using 3D printing services - these are available in many locations worldwide, and most of them can deliver printouts to places with limited access. It is also possible that people may use some commercial 3D printers on their own. Two widely available printing technologies have been considered: Fused Deposition Modelling (FDA) and stereolithography (SLA), also known as resin printing. Printers of both technologies are affordable, and their maintainability costs do not differ significantly.

The first method is more commercially common. Its printing process is based on building successive layers of heated thermoplastic filament. The popularity of this technique comes from two main factors: it is easy to maintain and does not require much postprocessing. However, it lacks in accuracy.

SLA is based on curing photopolymer resin with a laser. It provides better resolution, but such printers have to be thoroughly cleaned after each printing session. They also generate more supporting structures, and the printouts require extra effort to make them look as intended. SLA printers also involve more health risks caused by the toxicity of used materials.

The maximum printing diameter in commercial 3D printers usually ranges from 15 to 30cm. The diameter of the generated models is set to 12cm, which is enough to allow experiencing its tactile features while keeping it small enough to print with a vast majority of the printers on the market. I printed prototype models using FDA, and the accuracy of the results is satisfying. Therefore, both types of printers can be used without any tactile experience drawbacks. The exact parameters used for producing the prototypes are:

- Layer height: 0.3mm
- Fill density: 10%
- Shell Count: 3

The estimated cost of printing materials to print one model using the parameters I mentioned is 5.21 GBP, and an example online 3D printing service charges their customers 15.41 GBP to print the example model using the same layer height, but a higher fill density of 20%.

Chapter 5

Evaluation and External Aspect

5.1 Internal Testing

This section evaluates the software against the non-functional requirement 2a from Section 1.4. I tested the robustness and the response time of each back-end element. For the front end, I evaluated the accessibility of my implementation.

Wave Processor was tested manually by converting audio files to previously mentioned popular formats: AAC, AIFF, FLAC, MP3 and WAV. The software component passed the test successfully without any issues. I tested its compatibility with *Model Generator* by recording 20 voice memos with various voice tones and checking the modelled results. After implementing the complete processing pipeline, no more issues persisted. The execution time of this part was stable at about 0.5 seconds.

Emotion Extractor was tested using the MP3 files generated by *Wave Processor*. I manually compared the transcriptions of the Speech to Text (STT) model with the recording by listening to it for each input. Although it made some mistakes, as presented in Table 3.1 in Section 3.2, they do not seem to influence the extracted emotions. STT and Tone Analytics tools did not produce any errors after using compressed audio files as the input, as mentioned in Section 4.2.3. The execution time of that stage is stable and averages around 10 seconds without major outliers.

Model Generator was tested in two scenarios. The first one aimed to identify issues with modelling various permutations of emotions without focusing on their confidence. I did it by creating all the non-repeating permutations of the possible emotions of size 3, 2 and 1, which produced $\frac{5!}{(5-3)!} + \frac{5!}{(5-2)!} + \frac{5!}{(5-1)!} = 85$ test cases. I stored the success boolean and execution time of the results in a table and repeated the process for four different recordings with various voice tones. The counter was started after receiving a response from *Emotion Extractor* and stopped after a rendered model was saved on disk. The success rate of the final test was 100%, and the execution time averages were 7.23, 11.36, 7.69, and 5.14 seconds with a global average of 7.86 seconds. A boxplot picturing the distribution of these data is shown in 5.1a.

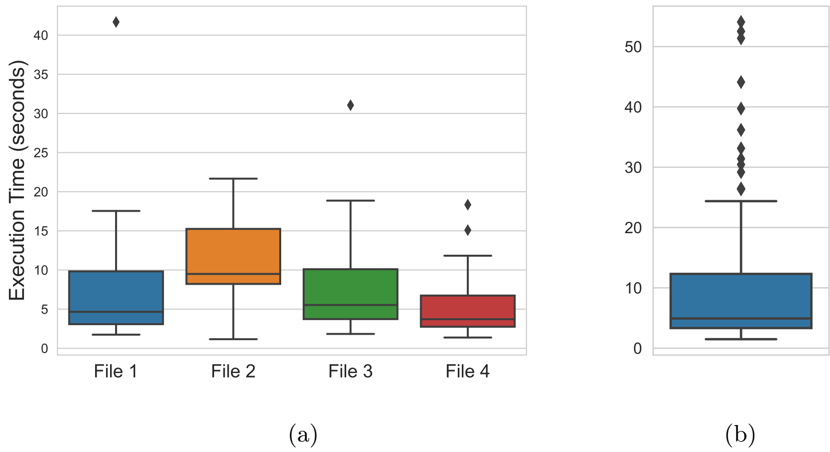


Figure 5.1: Boxplot of *Model Generator* execution time (a) for all possible emotion outputs for four different audio files and (b) for randomly selected emotions and confidences.

The other scenario focused on testing potential responses with different confidence values. I generated an array of 500 responses with random confidence values for randomly chosen non-repeating emotions. Similarly to the previous scenario, I measured the success rate and execution time. The code successfully generated the sculpture for each case, and the average execution time was 7.97 seconds, with the highest value of 54.07 seconds (Figure 5.1b).

The back end can generate a model with a 100% success rate, and its overall response time is estimated to be below 20 seconds. However, the entire process should take longer for the end user, as the file transfer speed has to be considered.

The fully functional website is currently hosted locally, with a front-end-only version hosted publicly using GitHub Pages for user evaluation of accessibility. It was manually tested by uploading different files and checking the returned model. The upload feature restricts the file type to the mentioned audio formats, which solves potential issues with uploading incompatible files. I manually tested the screen reader compatibility of the website using the most common solution for Mac computers, VoiceOver. The content is easy to navigate, and the system reads ARIA [1] labels correctly.

5.2 User Evaluation

As the product's purpose is to serve people, I evaluated it with the target group by conducting in-person and online interviews. Points 1a, 2b and 2c from the non-functional requirements specification in Section 1.4 formed research hypotheses for this stage. The study was approved by the University of Nottingham Research Ethics Committee. Participants provided written informed consent.

The study aimed to answer the following research questions:

1. Do visually impaired people need a tool for preserving memories?
   (a) Do they identify a need to preserve memories at all?

(b) Are there any alternative ways they do this?

2. Do they associate any physical object with past events?
3. Does interaction with tactile stimuli increase the recall rate of a described event?
4. Does the emotional association between the described event and the touched sculpture result in a better recall of the event?
5. Do the sculptures represent the expected emotions successfully?
6. Do the sculptures express the same emotions as the voice recordings used to generate them?
7. Is the website accessible to the users?

Points 1, 2 and 7 strictly required the involvement of visually impaired participants to identify the potential demand for the product among them and investigate their current practices for memory preservation. I accomplished this with the help of three organisations uniting people with vision health issues. Some of their members expressed interest in contributing to the project and participated in the study. Questions 3-6 examined emotions aroused by tactile experience. Although these affect all people, the sense of touch of some blind people is enhanced [20]. For this reason, these questions were answered with the same group of participants.

5.2.1 Methodology

Adults who struggled with blindness or severe visual impairment were invited to participate in this study. Severe impairment in this context is understood as vision loss that does not allow one to experience photographs fully. Four professionally active participants, mainly from highly technical fields, expressed interest and took part in the study. They answered a series of questions and completed three tasks. They used their preferred device with internet access to complete the website evaluation.

Part 1

Its first part aimed to answer the first two research questions, and it was conducted in the form of a semi-structured interview, where I asked the following questions:

1. What degree of vision impairment do you have?
2. Do you preserve memories of important moments in your life in some form?
   (a) If yes, how?
   (b) If no, is it important to you to preserve memories?
3. Is there any physical object that reminds you of a past event?

Part 2

The second part explored research questions 3 and 4, and it involved participants' interaction with the following study material:

- Three voice recordings describing fictional memories with information about:
  - a specific location,
  - accompanying people,
  - an event with one specific detail describing the experience,
  - another event.
- Two 3D models:
  - a model generated with the second recording,
  - a model generated with an unrelated recording.

All participants were asked to listen to the first recording without touching a 3D model. Then, they listened to the second recording and were instructed to touch a 3D model generated based on that recording. After that, they were asked to do the same for the third recording, where the model was generated using a different, unrelated recording. After the next part of the interview was finished, I asked the participants to describe the events from the recordings. For the recordings paired with models, I asked them to touch the sculptures during the recall process.

Part 3

The third part of the study focused on answering research questions 5 and 6 and involved interaction with the following material:

- Recording 1, describing an event associated with sadness.
- Model 1, generated using Recording 1 as the input.
- Recording 2, describing an event associated with frustration.
- Model 2, generated using Recording 2 as the input.
- Recording 3, describing an event associated with excitement and satisfaction.
- Model 3, generated using Recording 3 as the input.

I asked the participants about the emotions aroused from touching each model. They could pick any number from the studied set of emotions: *excitement*, *frustration*, *sadness*, *sympathy* and *satisfaction*. They were asked to sort the choices by relevance in decreasing order. Then, I played the recordings and instructed the participants to match them to the sculptures.

Part 4

The fourth part evaluated the accessibility of the website in a short task. I asked the participants to download a sample audio file and open a link. Then, I instructed them to generate a sculpture using the website without any help. Once they clicked the "Generate sculpture" button, the task was over. The following questions followed its completion:

1. Did you manage to perform the task?
2. Was it easy to navigate? Did you have any issues?
3. Is there anything that would make it more accessible for you?

Part 5

In the fifth part, I asked the participants about their general thoughts about the project and any feedback.

5.2.2 Results

The answers will be collected in written notes and analysed using a simple quantitative statistical analysis and a qualitative method of a theoretical thematic analysis [23], a commonly used technique for identifying interview patterns. It allows for extracting essential and repetitive themes from them in an organised way.

Qualitative analysis

The responses from parts 1 and 4 were analysed in a qualitative way. Research questions 1, 2 and 7 are considered in this stage.

Initially, I re-read the collected responses to familiarise myself with the data. As the research questions had already been prepared before, the study themes were extracted from them. I operated on codes to reduce unstructured data into small chunks of meaning. I used the open coding method to segment the data into themes, meaning I created the codes using irregularly sized text chunks. It could be a sentence, a lengthy statement, a line or a few words. The resulting list of codes for each theme is shown in Figure 5.1.

The study identified some currently used alternatives to the proposed idea. Some participants mainly recall memories using the sense of sound as the medium. Some participants collect physical objects for that purpose. Others only keep them in their heads and don't identify the need to use any medium.

Although some participants don't associate past memories with any physical objects, others do. They draw attention to their complete experience with the objects, including their shape, smell and texture. Their uniqueness plays an important role in the recall process of past events.

The participants didn't express the need for the proposed type of tool. The reasons they mentioned are:

- not being a sentimental person in general,
- model's generation price being too high,
- lack of emotional attachment to the tested models.

The website's accessibility was identified as high for both desktops and mobile devices. Some common design mistakes are not present. The good structure allows fluent navigation. The content and the way to use the tool

Table 5.1: Themes and associated codes identified after analysing participants' responses.

Memory preservation importance and used media	**Association of physical objects with past events**	**Accessibility of the website**	**The need for the tool**
Used to record voice messages.	Souvenirs like mugs of specific shape help recover memories.	Easy navigation and clear instructions.	Could be useful for more sentimental individuals.
Buys souvenirs on trips.	The association of souvenirs with the events is crucial.	Good structure and well-balanced paragraphs lengths.	No emotional attachment to the sculptures.
Records voice recordings or text notes.	Collecting bells as souvenirs.	No issues, short and concise.	Not a sentimental person.
Not that much anymore.	None.	Well-described content.	Too expensive to produce for now.
Used to write a journal but doesn't feel the need for it anymore.	None.	No unlabeled graphics.	Doesn't collect objects that are not practically usable.
Keeps memories in their head, and it's enough.	High importance of uniqueness.	Buttons are buttons and not clickable graphics, which is a common issue.	
	The texture and smell are important media as well.	The "Select file" button is not identified as a button by a screen reader.	

are easy to understand. An issue with the "Select file" button not being read as a button by a screen reader was mentioned.

Quantitative analysis

The influence of tactile stimuli on memory retrieval was evaluated in the second part of the study. This was measured using the information chunks that the recording was built on. Each successful recall of one chunk was awarded one point; wrong answers were penalised with a deduction of half a point, and missing answers didn't change the result. The average result for information retrieval without any tactile stimuli is 2.623. The result was higher for the stimuli of a model generated with an unrelated recording, which averaged 3.0. The highest average recall rate of 3.25 was measured when the touched model was generated using the relevant recording.

In the third part of the study, the emotional associations between the generated 3D shapes and emotions were measured. For the first sculpture, expressing sadness, one participant identified it as their second choice. The other three didn't provide an answer. For the second sculpture, expressing frustration, one participant responded with the correct emotion as their first choice. One participant associated sadness with it. The other two identified the emotion as negative but didn't make a more detailed choice. For the last model, associated with frustration and satisfaction, two participants selected satisfaction as their first choice. The other two didn't make any selection.

After hearing the recordings used for generating the sculptures, the participants were asked to match them. For the first recording, two of them selected the model correctly, one made a wrong choice, and one didn't answer. Three participants matched the second recording to the correct model, while one selected the wrong model.

Three participants selected the model correctly for the third recording, and one of them did not respond.

5.2.3 Discussion

The first analysis shows that some visually impaired people value memory preservation, but the method developed in this project is not satisfying or necessary for each of them. The currently used alternatives mostly satisfy the participants, and the physical form of the proposed method is not convincing enough to make up for the cost of 3D printing. Further exploration of the importance of the physical form [19] in this context could help better determine the need for this solution. The need for the product could be better evaluated on a wider spectrum of participants. As the participants mentioned, as highly technical people, they value utility over sentiment. Some additional media, such as smell or texture, could be introduced to further researched approaches to extend the experience. The study also lacks evaluation in a long-term use scenario, where a user would be able to experience the product the way it is meant to be used, as some of the participants expressed their inability to describe the sentimental expression of a sculpture due to the lack of retrospective autobiographical attachment to it. Some participants mentioned issues with identifying objects from icons and explicit miniature 3D representations of them, as explored in [38], which shows that shaping emotions is a better way to achieve the expected results.

The second analysis confirms that any form of tactile stimuli may positively correlate with a higher recall than sound on its own. Furthermore, touching shapes that are associated with the sentiment of retrieved information can potentially result in an even higher recall, which follows from the findings in the memory-emotion relationship from [35]. The explicit identification of emotions from the models is better than random. It is more apparent when the association is implicit as a match between an emotional story and a model. This shows that the sculptures can form a tangible expression of a sentimental story well, which confirms the associations identified in [21], [32] and [43] and suggests that the mapping and the generation algorithm achieved the goals set. The website is designed in an accessible way except for one minor issue with the "Select file" button. The discussed results are not statistically significant, and more research is needed to shed more light on the topics.

5.3 External Aspect

The original intentions of this project included external aspects in the following elements:

- using IBM's API for emotion recognition,
- collaboration with the external mentor to shape the idea and develop the system,
- a potential showcase of the project in IBM Innovation Studio,
- making the product openly available to users on the internet.

Two IBM APIs were used to implement the emotion recognition functionality. The mentor was involved in the early stage of the project to discuss different ideas. He significantly contributed to shaping the memory preservation use case of 3D models, which ultimately became its primary purpose. Although we had little interaction during the literature search and design phase, I regularly updated him regarding the project's progress towards the end development stages. I shared this information in the form of video recordings sent by email. The mentor then replied, sharing his feedback on the state of it.

The last two points are still in progress. I agreed with the mentor to share a demo of the project in a video form with him to pass it further after it is finished and submitted. The project is currently unavailable for public use due to issues mentioned in 4.2.5, but most of the project components are ready for deployment.

An additional external aspect of the project is the involvement of a community of visually impaired and blind people whose members significantly contributed to evaluating the project.

Chapter 6

Summary and Reflections

6.1 Project management

The Gantt chart from the project proposal presented in Figure 6.1 shows my initial project management plan. It was regularly revisited and updated if necessary after my weekly meetings with the supervisor on Thursdays. Its final shape is shown in Figure 6.2. The management methodology used in this project is the Waterfall approach with some extra flexibility and dynamic adjustments.

The duration of each task is measured in weeks to add some flexibility and allow for an overlap of some stages which could, or needed to be completed simultaneously. Some tasks took less time than pictured, and others took more. I set the length of the time windows considering other factors like other assessment deadlines or exams. The main goal was to dedicate as much time as necessary for each task and complete the milestones highlighted in yellow on schedule. I planned a 4-week gap in January to amortise tasks that took longer than expected. In the initial project plan, I expected the deadline to be later than it was finally announced. Even though the proposed chart included some extra time in the end for potential complications, it had to be removed later.

The first milestone included submitting the project proposal, research ethics documents, and the requirements specification document, as well as identifying necessary tools and provisional design. It was successfully completed on schedule without significant issues.

The second milestone involved developing a minimally viable back-end solution to generate models successfully. I encountered many issues at this stage, which resulted in filling the January gap with time dedicated to overcoming them. Firstly, the literature exploration of the project became deeper than expected since its motivation had to reach further than only to the emotions-touch association. I had to find relevant information about emotions' role in memory preservation and recall. Secondly, I planned to use a version of the Tone Analyser, which was soon discontinued and finding the solution (Section 4.2.5) delayed the wave processing stage. Finally, the issues with setting up a working development environment for CadQuery (Section 4.2.5)

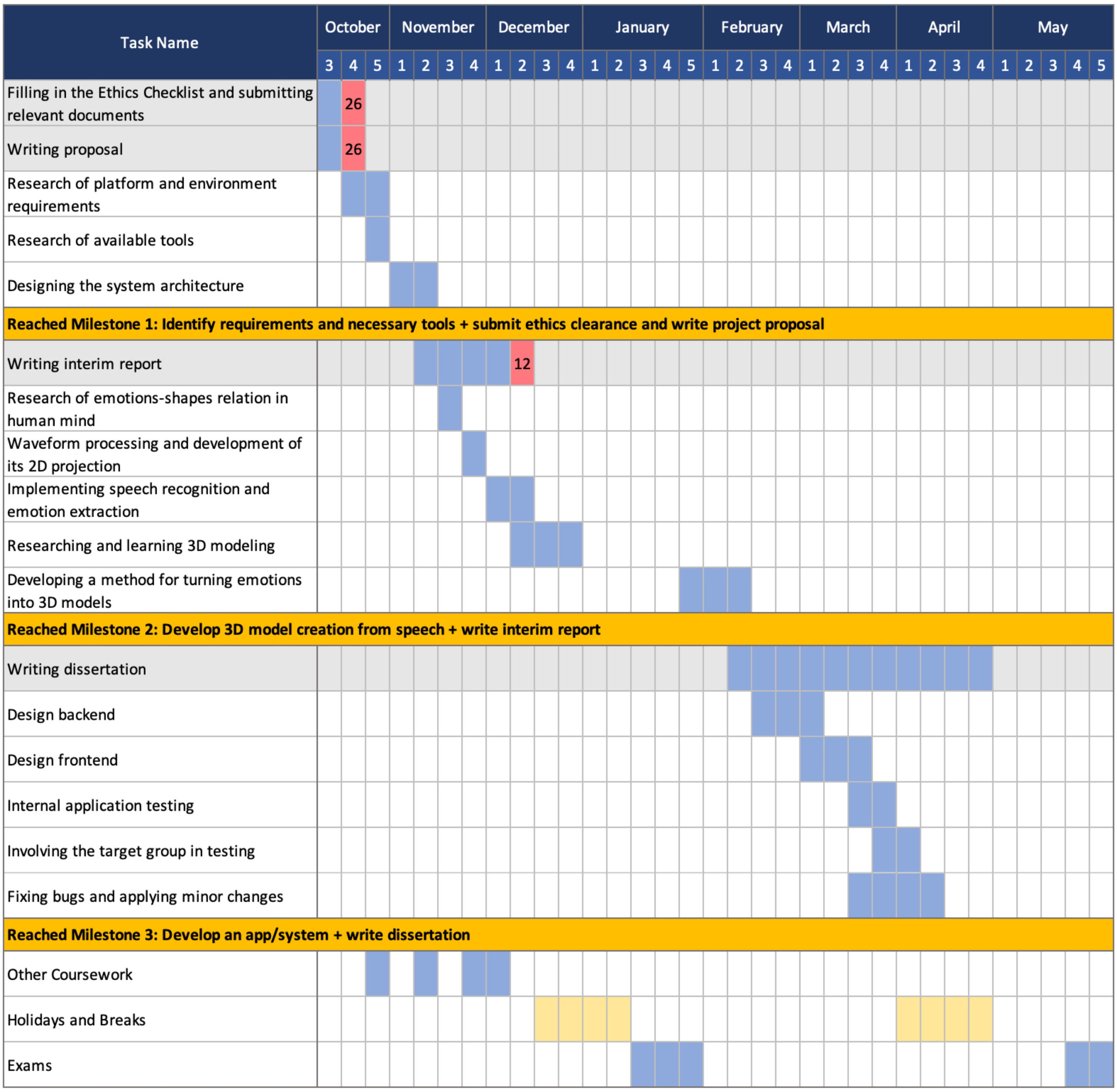


Figure 6.1: Initial project work plan.

significantly impacted the delay in the modelling stage. The minimum viable product was therefore implemented in mid-March.

To complete the project before the deadline, I modified the website design to save some time on features that were not crucial. These included:

- Removing "audio recording" feature. Now uploading a ready audio file was the only way to generate the model.
- Removing the split between pages and moving the functionality to the main page.
- Not considering the optional functionalities anymore and focusing on the core functionality.

Resulting from better awareness of time scales and project background, I made the following adjustments to the previous work plan:

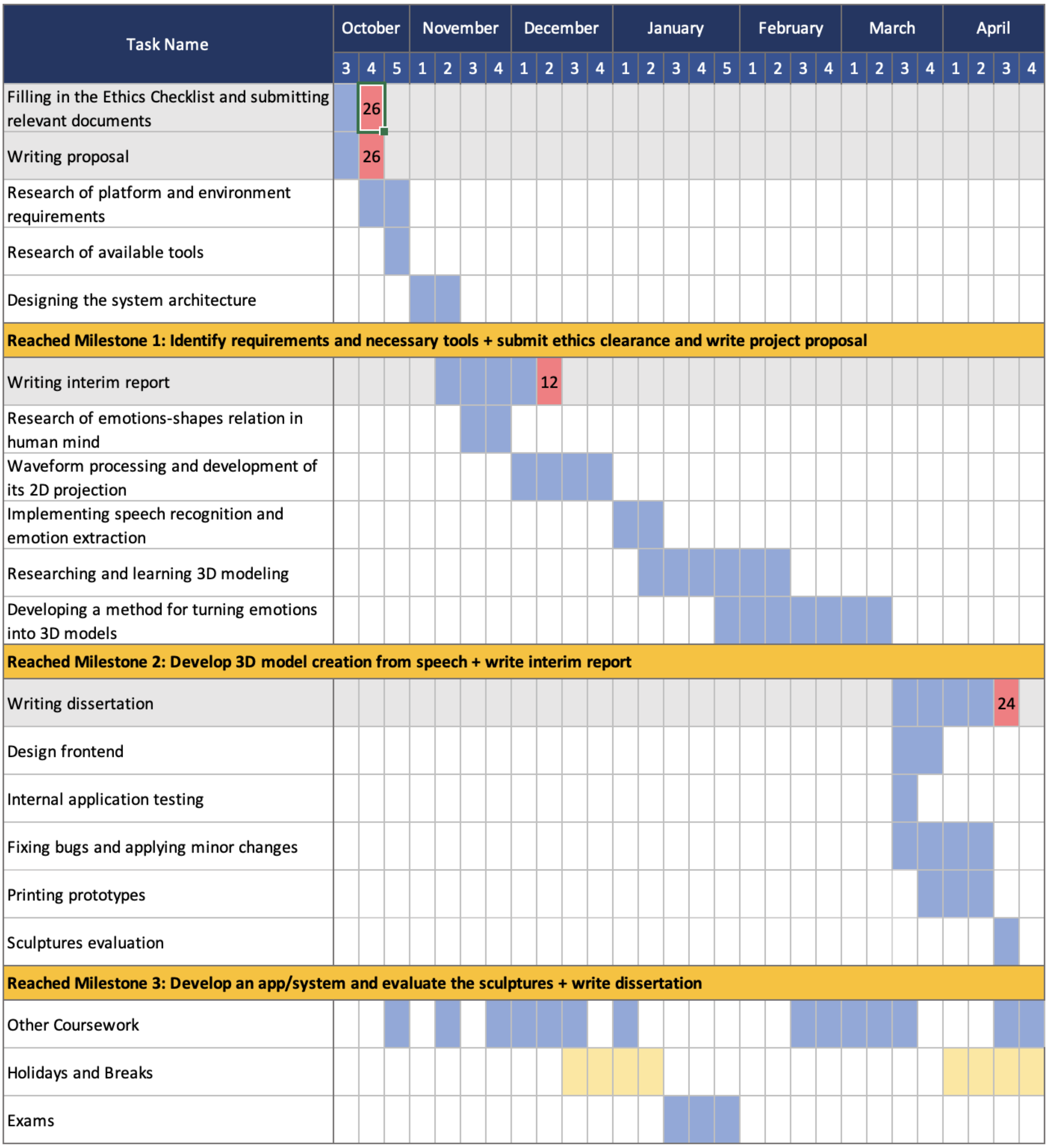


Figure 6.2: Final project work plan.

- Removed the "Design back-end" task. During the requirements specification and system architecture design stages, I did not identify a need for a database, which was initially expected to be necessary. Therefore, the entire back-end solution was already there at this point.
- At some point, I replaced "Involving the target group in testing" with "Sculpture evaluation", but I moved back to the initial idea of evaluating the entire project involving the website.
- Squeezed all the tasks in the third milestone to complete them before the deadline.

I completed the first two tasks of the third milestone faster than expected. In the meantime, I was improving and extending the model generator. One of the major elements I should have considered in the plan was printing the prototypes for evaluation. Although the first volunteers expressed their interest in participating in my study in late March, the printing process of the prototypes took more than three weeks, which delayed the evaluation.

6.2 Contributions and reflections

6.2.1 Research gap

The field of Human-Computer Interaction has already proven the potential of 3D modelling in accessibility issues. However, there needed to be more focus on the problem of memory preservation. Some existing solutions do not satisfy potentially important requirements [34][2], and others require too much time and resources to satisfy the needs [6]. The research gap allowed this project to bring up the issue and form new questions that can be explored further.

6.2.2 Achievements and implications

The project successfully achieved the aim and most of the objectives set. A potential visually impaired user can generate a 3D memory preservation model without help. However, the product is yet to be publicly available. Although its ability to help with memory preservation is not fully proven, the evaluation results show some potential for it. The web application is accessible to the target group. The model generation algorithm produces results with a 100% success rate on the tested examples in a reasonable time. The optional objective of reversing the process and obtaining an audio file from the model has not been implemented. As I mentioned in Section 6.1, the website functionality was narrowed to the main feature, and the optional ones like model preview, multiple pages and audio recording were omitted. The difference between models printed with SLA and FDM was not evaluated, even though it was suggested in the interim report. As I mentioned in Section 4.4, the resolution of FDM, being worse than that of SLA, was satisfying for me, which implied that the SLA method would be satisfying too. However, this was only my personal assessment, and some unexpected differences could arise. The idea of comparing different open-source sentiment analysis techniques was discontinued as more information about the Tone Analytics tool [47] and a benchmark [27] was found.

The results discussed in Section 5.2.3 suggest that the explicit association between emotions and the 3D printed models is identified, and a more substantial implicit correspondence is observed between short sentimental stories and their relevant models. There is a potentially positive impact on the recall rate of information when the same tactile object is involved during the absorption and retrieval of information compared to the same process without tactile experience. A higher recall rate was identified when the objects were emotionally associated with the given information. The cost of 3D printing was suggested to be yet too high to attract users. The results were obtained with a small study group and are not statistically significant. More exploration is required to form fully reliable conclusions. Although study participants had no issues with the file upload feature on the website, recording voice messages directly within it could make it more accessible for blind people who are less technically fluent.

The project combines various seemingly independent findings in Psychology, Human-Computer Interaction and Artificial Intelligence to propose an innovative idea that can be useful for various purposes and form the basis for other research. I believe that apart from memory preservation, this contribution can be used or extended for:

- Education - to familiarise children with emotions.
- Therapy or counselling - for individuals who struggle with expressing emotions.
- Measuring responses - as an extra measure of emotional response in different studies.
- Training AI models - the autonomy of this solution could form a base for training Computer Vision models to recognise the "Bouba/Kiki" effect, which is intuitive for humans.
- Artistic purposes.

6.2.3 Limitations

Although the models generated with my solution have already shown some potential in correct emotional associations, there is still room for improvement. The proposed solution could be designed with the involvement of the target group and psychology experts, making it more tailored to existing needs and professional. A broader set of positive emotions could be considered, as people tend to preserve these more often. The emotion recognition part could be extended by considering some acoustic features of the recording apart on the top of the transcribed text [46]. The mapping between the findings from literature in Section 2.2 and the emotions used by the Tone Analytics tool can lack precision and could be consulted with experts.

Because of many limitations of CadQuery identified during development, Blender could be explored to add more variation to the models. The implementation could be extended to introduce more complex, overlapping shapes. Although the algorithm overcomes the overlap issue mentioned in Section 4.2.5 by trying out different configurations and regenerating random numbers on error, using this approach is not the best practice in software development. Finding a better solution would reduce the response time and make the software more robust.

The method embodies some original ideas and creativity and could create new intellectual property. However, I did not discuss its potential ownership with my supervisor and the industry mentor. I will likely develop it on my own to allow others to use it via the website. I am not planning to exploit the work commercially or promote it in other ways in the nearest future.

6.2.4 Bigger picture

The research involves conducting studies with human participants ethically. In particular, all of the University of Nottingham's research ethics requirements have been followed. Participation in the study was entirely voluntary, and all of the participants were clearly informed about the process and what was expected of them. The participants' privacy was protected by removing any personal information shared during the interviews that could be used to identify them. Additionally, the participants were identified using alphanumerical IDs instead of their real names. It involved collecting general visual health data, but the mentioned measures and safe data storage on the University O365 OneDrive folder prevent identification and unauthorised access. Overall, the research was conducted ethically and responsibly.

The created website ensures accessibility to all users considering the accessibility implications of the Equality Act (2010) with updates from 2018. Its design more closely considers the needs of people without sight, and it complies with WCAG [14], as I mentioned in Section 3.4. There is no risk of potential criminal use of the software. It does not store any information, and the data transit and processing by IBM ensure security (Section 4.2.3). Although these services comply with GDPR, which implies adherence to the Privacy and Electronic Communications (EC Directive) Regulations (2003), some distinct requirements of the Data Protection Act that differ from the European Union's law may need further review. There is no potential exposure to harmful content or online abuse, so the Online Safety Bill, which can come into force soon, does not affect the project.

The work considers the broader ethical and social aspects. It aims to benefit visually impaired people, reducing the inequality in their accessibility to memory preservation. The project has no potential unintended consequences nor application to military or criminal use. It does not appear to affect the stakeholders, and no concerns regarding this were reported in the user evaluation. The work extends current research and provides a novel approach to an existing issue. Furthermore, it has significantly more benefits than drawbacks. These points make it worth the exploration that was done in this project. The proposed idea involves the production of physical objects, which requires the use of 3D printing materials. FDM printing can be done with biodegradable or recycled plastic. There are also sustainable materials that can be used for STL printing. Additionally, the sculptures are not intended for mass production, so their environmental impact is minor.

6.2.5 Reflections and personal growth

I am satisfied with the outcome of the project. It helped me develop better planning and management skills through its self-directed nature. Some inaccurate predictions of task time consumption helped me to identify the project elements that usually require more time than expected. The technical challenges I solved made me a more aware and confident software developer. I identified the importance of the involvement of the stakeholder in the design process, which is the primary factor that I will consider in the development of my similar future projects. I understood the vast benefits of regularity, which, through meetings with my supervisor, helped me with motivation and continuous progress during the academic terms. I noticed the side effects of its absence during the breaks when I struggled with the dynamic identification of new tasks to keep the same development pace. The work helped me grow as a researcher by familiarising me with all the laws necessary to conduct a user study ethically and to publish a secure software product. I learned to critically evaluate my solutions with the help of people involved in the project and external resources to improve the proposed solution dynamically.

Bibliography

[1] ARIA14: Using aria-label to provide an invisible label where a visible label cannot be used, Techniques for WCAG 2.0. Retrieved 13/04/2023 from https://www.w3.org/TR/WCAG20-TECHS/ARIA14.html.

[2] Assistive NFC Tagging System for Blind and Low Vision Users. Retrieved 24/11/2022 from https://www.wayaround.com/.

[3] Blender 3.3 Python API Documentation. Retrieved 07/12/2022 from https://docs.blender.org/api/current/index.html.

[4] FreeCAD: Your own 3D parametric modeler, Retrieved 10/04/2023 from https://www.freecad.org/.

[5] The International Agency for the Prevention of Blindness (IAPB), Global Estimates of Vision Loss. Retrieved 18/11/2022 from https://www.iapb.org/learn/vision-atlas/magnitude-and-projections/global/.

[6] Living paintings. Retrieved 24/11/2022 from https://livingpaintings.org.

[7] Matplotlib — Visualization with Python. Retrieved 10/04/2023 from https://matplotlib.org/.

[8] Miniconda — conda documentation. Retrieved 12/04/2023 from https://docs.conda.io/en/latest/miniconda.html.

[9] NumPy Documentation. Retrieved 09/04/2023 from https://numpy.org/doc/.

[10] PyAudio Documentation — PyAudio 0.2.13 documentation. Retreived 09/04/2023 from https://people.csail.mit.edu/hubert/pyaudio/docs/.

[11] Speech to Text. IBM Cloud API Docs. Retrieved 01/04/2023 from https://cloud.ibm.com/apidocs/cloud.ibm.com/apidocs/speech-to-text.

[12] Tone analytics. IBM Cloud Docs. Retrieved 01/04/2023 from https://cloud.ibm.com/docs/cloud.ibm.com/docs/natural-language-understanding.

[13] W3.CSS. Retrieved 12/04/2023 from https://www.w3schools.com/w3css/default.asp.

[14] Web Content Accessibility Guidelines (WCAG) 2.1. Retrieved 09/04/2023 from https://www.w3.org/TR/WCAG21/#resize-text.

[15] WebAIM: Screen Reader User Survey #9 Results. Retrieved 03/12/2022 from https://webaim.org/projects/screenreadersurvey9/.

[16] WebAIM: Web Accessibility for Designers. Retrieved 06/12/2022 from https://webaim.org/resources/designers/.

[17] CadQuery. Retrieved 06/12/2022 from https://github.com/CadQuery/cadquery., Dec. 2022. original-date: 2018-10-28T17:57:18Z.

[18] Al Marouf, A., Hossain, R., Kabir Rasel Sarker, M. R., Pandey, B., and Tanvir Siddiquee, S. M. Recognizing Language and Emotional Tone from Music Lyrics using IBM Watson Tone Analyzer. In *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)* (Feb. 2019), pp. 1–6.

[19] Atasoy, O., and Morewedge, C. K. Digital Goods Are Valued Less Than Physical Goods. *Journal of Consumer Research 44*, 6 (Apr. 2018), 1343–1357.

[20] Bauer, C. M., Hirsch, G. V., Zajac, L., Koo, B.-B., Collignon, O., and Merabet, L. B. Multimodal MR-imaging reveals large-scale structural and functional connectivity changes in profound early blindness. *PLOS ONE 12*, 3 (Mar. 2017), e0173064. Publisher: Public Library of Science.

[21] Blazhenkova, O., and Kumar, M. M. Angular Versus Curved Shapes: Correspondences and Emotional Processing. *Perception 47*, 1 (Jan. 2018), 67–89.

[22] Blei, D. M. Latent Dirichlet Allocation.

[23] Braun, V., and Clarke, V. Using thematic analysis in psychology. *Qualitative Research in Psychology 3*, 2 (Jan. 2006), 77–101. Publisher: Routledge _eprint: https://www.tandfonline.com/doi/pdf/10.1191/1478088706qp063oa.

[24] Castilho, V., Henriques, D., Correia, W., Souza, L., and Melo, S. Embodied Cognition and Tactile Interaction: A Review on How Multi-sensorimotor Experiences Assisted by 3D Printing Can Shape the General Perception of Daily Activities. July 2020, pp. 324–338.

[25] Chen, A., Kim, J., Mankoff, J., Grossman, T., Coros, S., and Hudson, S. Reprise: A Design Tool for Specifying, Generating, and Customizing 3D Printable Adaptations on Everyday Objects. pp. 29–39.

[26] Ebe, Y., and Umemuro, H. Emotion Evoked by Texture and Application to Emotional Communication. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2015), CHI EA '15, Association for Computing Machinery, pp. 1995–2000.

[27] Filippidou, F., and Moussiades, L. α Benchmarking of IBM, Google and Wit Automatic Speech Recognition Systems. In *Artificial Intelligence Applications and Innovations* (Cham, 2020), I. Maglogiannis, L. Iliadis, and E. Pimenidis, Eds., IFIP Advances in Information and Communication Technology, Springer International Publishing, pp. 73–82.

[28] Gordon, I., Voos, A. C., Bennett, R. H., Bolling, D. Z., Pelphrey, K. A., and Kaiser, M. D. Brain mechanisms for processing affective touch. *Human Brain Mapping 34*, 4 (2013), 914–922. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hbm.21480.

[29] Hirst, W., Phelps, E. A., Meksin, R., Vaidya, C. J., Johnson, M. K., Mitchell, K. J., Buckner, R. L., Budson, A. E., Gabrieli, J. D. E., Lustig, C., Mather, M., Ochsner, K. N., Schacter, D., Simons, J. S., Link to external site, t. l. w. o. i. a. n. w., Lyle, K. B., Cuc, A. F., and Olsson, A. A ten-year follow-up of a study of memory for the attack of September 11, 2001: Flashbulb memories and memories for flashbulb events. *Journal of Experimental Psychology: General 144*, 3 (June 2015), 604–623. Num Pages: 604-623 Place: Washington, US Publisher: American Psychological Association (US).

[30] Hutmacher, F., and Kuhbandner, C. Long-Term Memory for Haptically Explored Objects: Fidelity, Durability, Incidental Encoding, and Cross-Modal Transfer. *Psychological Science 29*, 12 (Dec. 2018), 2031–2038. Publisher: SAGE Publications Inc.

[31] Laird, J., Cuniff, M., Sheehan, K., and Shulman, D. Emotion specific effects of facial expressions on memory for life events. *Journal of Social Behavior & Personality 4* (Jan. 1989), 87–98.

[32] Lin, A., Scheller, M., Feng, F., Proulx, M. J., and Metatla, O. Feeling Colours: Crossmodal Correspondences Between Tangible 3D Objects, Colours and Emotions. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, May 2021), CHI '21, Association for Computing Machinery, pp. 1–12.

[33] Musello, C. Studying the Home Mode: An Exploration of Family Photography and Visual Communication. *Studies in Visual Communication 6*, 1 (Mar. 1980), 23–42.

[34] Pakėnaitė, K., Nedelev, P., Kamperou, E., Proulx, M. J., and Hall, P. M. Communicating Photograph Content Through Tactile Images to People With Visual Impairments. *Frontiers in computer science (Lausanne) 3* (2022). Publisher: Frontiers Media S.A.

[35] Phelps, E. A. Human emotion and memory: interactions of the amygdala and hippocampal complex. *Current Opinion in Neurobiology 14*, 2 (Apr. 2004), 198–202.

[36] Ramachandran, V., and Hubbard, E. Hearing Colors, Tasting Shapes. *Scientific American 288* (June 2003), 52–9.

[37] Robert, J. Pydub. Retrieved 09/04/2023 from https://github.com/jiaaro/pydub., Apr. 2023. original-date: 2011-05-02T18:42:38Z.

[38] Révész, G., and Wolff, H. A. Psychology and Art of the Blind. 360.

[39] Salgado Montejo, A., Alvarado, J., Velasco, C., Salgado, C., Hasse, K., and Spence, C. The sweetest thing: the influence of angularity, symmetry, and the number of elements on shape-valence and shape-taste matches. *Frontiers in Psychology 6* (2015).

[40] Samonte, M., Laurente, E., Magno, K., and Perez, C. Braille3D: using haptic and voice feedback for braille recognition and 3D printing for the blind. *IOP Conference Series: Materials Science and Engineering 482* (Mar. 2019), 012027.

[41] Steffens, A. N. V., Langerhuizen, D. W. G., Doornberg, J. N., Ring, D., and Janssen, S. J. Emotional tones in scientific writing: comparison of commercially funded studies and non-commercially funded orthopedic studies. *Acta Orthopaedica 92*, 2 (Apr. 2021), 240–243.

[42] Tang, W., Shu, Y., Bai, S., Peng, Y., Yang, L., and Liu, R. Brain activation related to the tactile perception of touching ridged texture using fingers. *Skin Research and Technology 28*, 2 (2022), 254–264. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/srt.13122.

[43] Turoman, N., Velasco, C., Chen, Y.-C., Huang, P.-C., and Spence, C. Symmetry and its role in the crossmodal correspondence between shape and taste. *Attention, Perception and Psychophysics 80*, 3 (Apr. 2018), 738–751. Num Pages: 738-751 Place: Austin, Netherlands Publisher: Springer Nature B.V.

[44] VanBuskirk, K., and Letson, M. R. An IBM Watson Tone Analysis of Selected Judicial Decisions. *Scribes Journal of Legal Writing 19* (2020), 25–48.

[45] Wang, J., and Hsu, Y. The Relationship of Symmetry, Complexity, and Shape in Mobile Interface Aesthetics, from an Emotional Perspective—A Case Study of the Smartwatch. *Symmetry 12*, 9 (Sept. 2020), 1403. Number: 9 Publisher: Multidisciplinary Digital Publishing Institute.

[46] Wu, W., Zhang, C., and Woodland, P. C. Emotion Recognition by Fusing Time Synchronous and Time Asynchronous Representations. In *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (June 2021), pp. 6269–6273. ISSN: 2379-190X.

[47] Yin, P., Liu, Z., Xu, A., and Nakamura, T. Tone Analyzer for Online Customer Service: An Unsupervised Model with Interfered Training. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (New York, NY, USA, Nov. 2017), CIKM '17, Association for Computing Machinery, pp. 1887–1895.

[48] Zhang, X., Tran, T., Sun, Y., Culhane, I., Jain, S., Fogarty, J., and Mankoff, J. Interactiles: 3D Printed Tactile Interfaces to Enhance Mobile Touchscreen Accessibility. pp. 131–142.