VAMO

WHAT STORY DO *YOU* WANT TO TELL?

ME310 SPRING DOCUMENTATION 2018

ME310

Stanford

Lucas Zhou
Felipe Cabral
Simran Shah

HPI

Winfried Lötzsch
Simon Krogmann
Sebastian Serth

1 Executive Summary

As technology improves it becomes more and more present in our everyday lives. Artificial Intelligence (AI) is a field of research that has existed for the past sixty years but only recently has it become ubiquitous. In the current context we live in, AI is present in smartphones and many applications we use daily from search engines to image processing to recommendation engines. AI is finishing our sentences and learning our taste. Our team, comprised of three graduate students from Stanford University and three from Hasso Plattner Institute (HPI), Potsdam, has been approached by FutureWei Technologies to enhance human communication with Artificial Intelligence.

Throughout our exploration of the problem space, we surveyed technologies and applications that are equipped with Artificial Intelligence and noticed different modes of communication among people as well as between people and the technology equipped with AI. After exploring various user groups and potential applications of Artificial Intelligence such as language translation, self reflection, remote communication and assistance in conversations, we decided to focus on communication within families, particularly between parents and children between the ages of 3 and 5 years.

We found that there was scope for improvement in the communication between parents and children in this age group. Parents with children between 3 and 5 years of age are familiar with the daily struggle of getting their children through routines and the negative home environment that develops as a result of the poor communication. Children of this age are at a crucial stage in terms of development of emotions, language and sense of time. They need time to prepare themselves for transitions between activities and engagement to keep them from being distracted while going through these transitions. The parent on the other hand, is concerned about their child's schedule and getting them from place to place on time, while struggling with several other responsibilities, all while being tired and stressed out after a long day's work. This discrepancy between the child's need to prepare for change and the parent's wish to move from task to task creates a lot of tension, stress and negative communication in the household.

We propose the system VAMO to solve this need. It signals an approaching transition to the child using light queues in the house, an association that the child learns, and provides them with a Buffer Time to prepare for upcoming changes. Once the Buffer Time expires, the system signals the child again and guides them to the desired location. This happens through a story told using Milestones placed around the house on the path between the two activity locations. Artificial Intelligence is used to create personalized stories relevant to the child's life to promote even further engagement by incorporating information about the child. To achieve this, we let the parents enter the following information into a mobile application to control the experience:

- Future events, that the child does not know about, so it can be prepared for them.
- Past experiences, so the child can learn more about something interesting it saw.
- Child personality data, to help the AI select relevant stories.

This information is incorporated into the story for an enriching experience. Thus, the routine stress and negative communication in the family is replaced by positive communication and engagement.

Contents

List of Figures

List of Tables

List of Source Codes

Glossary

Alexa	Amazon's virtual smart home assistant.
Animoji	Animation sent in electronic messages in the iPhone X that maps the user's facial expression onto an animated animal character.
API	Application Programming Interface, a pre-defined set of endpoints to exchange information between two system components.
AI	Otherwise known as Artifical Intelligence, it is the simulation of human intelligence processes by machines, especially computer systems. These processes include learning (the acquisition of information and rules for using the information), reasoning (using the rules to reach approximate or definite conclusions), and self-correction.
Benchmarking	A process of researching and observing to understand the state of the art for a given field or topic.
Buffer Time	A 15 minutes time period that prepares the child for upcoming changes.
BITS	Building Integrated Timing Source, a precise time source used in phone or IT networks.
Bixby	Samsung smartphone voice assistant
Brainstorming	A process by which groups of people generate ideas.
CEP	Otherwise known as a Critical Experience Prototype, this is a prototype built to test the user experience of a concept critical to the problem statement.
CFP	Otherwise known as a Critical Function Prototype, this is a prototype built to get an understanding of the feasibility of a core concept within the problem domain.
Dark Horse	The first prototype in the winter term which focuses on a radical idea that was not focused on in the fall quarter with either the CEP or CFP.
Deep Learning	a form or machine learning in which the machine is able to do unsupervised learning from unlabeled data.
Design Thinking	is a way of finding human needs and creating new solutions using the tools and mindsets of design practitioners.
DSLR	Otherwise known as digital single lens reflex, a professional digital photo camera with a movable mirror
Emoji	Icons used in electronic messaging.
EXPE	Otherwise known as Experience, this is the end of the project and describes the project fair where the final prototypes will be shown.

FaceTime Apple's video conferencing app.

Feature Phone A non-touchscreen phone with hardware keys and limited functionality when compared to a smartphone.

FDM Fused Deposition Modeling.

Filter Bubble a state of intellectual isolation that can result from personalized searches when an AI algorithm selectively guesses what information a user would like to see based on information about the user.

Functional Otherwise known as a Functional System Prototype, this is a prototype showing a whole system and designed to include all functionality identified so far; it is a vision for the final product.

Funky Otherwise known as a Funky System Prototype, this is a prototype built to integrate different aspects of a system from past prototypes in a rough design.

Google Assistant Android smartphone voice assistant.

HCI Human Computer Interaction.

Hunting Plan A Gantt chart to identify how long a part of the project takes, what it depends on and who is responsible for each part.

Ideation Process of generating and forming ideas and concepts.

JSON JavaScript Object Notation, a format to store or exchange with a hierarchy key-value format

Microsoft Cortana Microsoft's personal assistant used in Windows 10.

Milestone The main hardware built in this project. A triangular device that is mounted on the wall and is capable of telling stories and giving various lighting effects.

Needfinding The process of identifying user needs. This is done though two tools: observations and interviews.

NPU Otherwise known as Neural Processing Unit, it is a part of a microchip able to power AI computing.

Raspberry Pi A small single-board computer capable of running Linux

Philips Hue Philips Hue is a line of Wi-Fi-enabled color changing LED lamps and white bulbs created by Philips.

RGB Red, Green, Blue.

Siri Apple's iPhone voice assistant.

SLA Stereolithography.

SWOT Analysis	Stands for Strengths, Weaknesses, Opportunities and Threats Analysis and is a tool to organize and evaluate these four elements of an organization.
VAMO	Name of our final product from the Portuguese word vamos or, colloquially, vamo, meaning let’s go
WeChat	Messaging, social media and payment app.

2 Project Context

The following chapter gives an overview on the project corporate partner, the design challenge and the team working on the project. As well as a short introduction to Design Thinking, the design process used to tackle the project prompt.

2.1 Corporate Partner

FutureWei Technologies, Inc. is one of the four entities of Huawei Technologies Ltd. set up in the United States. The parent company, Huawei Technologies Ltd., engages in the production, sales, marketing as well as research and development of networking and telecommunications equipment. It offers optical networking; BITS clock supply, wireless networking, broadband access, data communications, and video conferencing products. The company provides enterprise network solutions, including video conferencing and networking solutions. Huawei Technologies Ltd. also offers carrier network solutions comprising metro optical, long haul optical, optical networking, mobile network, and broadband. The company was founded in 1987 and is based in Shenzhen, China. The Texas-based company FutureWei Technologies, Inc. was founded in 2001 and focuses on Research and Development in the US.

FutureWei has three main branches: enterprise, carrier, and consumer technology. FutureWei is considered a separate entity from Huawei, but is financially tied to Huawei and has access to Huawei resources. Huawei made a revenue of 75 billion in 2016 and yielded 5.3 billion dollars in profits. It has 180,000 employees globally while FutureWei itself has about 500 to 1000 employees in the United States [11, 9].

2.1.1 SWOT Analysis

We employed a SWOT Analysis (Strengths, Weaknesses, Opportunities and Treats) in order to better understand FutureWei's position in the telecommunication market and what areas it could benefit from exploring.

Strengths (Internal)

- FutureWei's parent company, Huawei is the world's largest telecommunications equipment manufacturer. FutureWei, although independent from Huawei, has access to Huawei resources and both are financially tied.
- There is a large focus on research and development. FutureWei is the USA based R&D branch but Huawei is also has subsidiaries around the world focused on R&D in China, UK, Pakistan, Finland, India, Germany, France, Belgium, Colombia, Sweden, Ireland, Russia, Israel and Turkey.
- As of the beginning of 2010, approximately 80% of the world's top 50 telecommunications companies had worked with Huawei.

- The company has been able to expand to different business sectors offering different kinds of services and expanding within the telecommunications market. The three areas of focus are:
  - Telecom Carrier Networks, building telecommunications networks and services.
  - Enterprise Business, providing equipment, software and services to enterprise customers, e.g. Government Solutions.
  - Devices, manufacturing electronic communications devices.
- FutureWei has been the world's largest vendor of telecommunications devices since 2012.
- Known for selling relatively inexpensive smartphones, which makes them competitive in most markets around the world.
- Strategic partnerships with universities working with an open innovation research model
- World's third biggest and most profitable smartphone manufacturer.

Weaknesses (Internal)

- The company is unable to participate in some major markets due to protectionism for local telecom companies. Countries, like the USA fear that domestic businesses might get driven out of the market.
- Founder's former affiliation with the People's Liberation Army means that there is less trust between the company and foreign markets due to diplomatic concerns regarding the Chinese government.

Opportunities (External)

- Growing market share in Europe.
- On pace with other smartphone companies such as Apple with the launch of their AI chip - Kirin 970 for the new smartphone Mate 10 Pro.
- Market trend to produce phones that are able to do AI computations locally instead of having to use a cloud based server to complete the same tasks.
- The lack of presence in the American market means that this is a large consumer base that the company can still move towards in the future.

Threats (External)

- Not much brand recognition abroad, especially in the USA where the market for smartphones is dominated by Apple and Samsung.
- Fear and accusations that the company uses its devices for espionage and possible ties to the Chinese government.
- Apple's launch of the A11 chip that runs on a neural brain is the culmination of a two year effort to expand their Artificial Intelligence capabilities and research.
- The company has been involved in several intellectual property rights violation cases and has been accused of copying and stealing technologies from its competitors.
- Concerns about treatment of clients and employees including bribes, extortion, overworking and not allowing employees in overseas subsidiaries the same access rights to information, assets and facilities as their Chinese counterparts in the same office.

2.2 Need Statement

FutureWei is partnering with ME310 for the second year. The 2017-2018 FutureWei project aims to promote positive engagement in communication with Artificial Intelligence (AI). The term Artificial intelligence "is frequently applied to the project of developing systems endowed with the intellectual processes characteristic of humans, such as the ability to reason, discover meaning, generalize, or learn from past experience". The increasing availability of computational power has recently made the training of more sophisticated AI models possible. These models can be used to solve more difficult problems like image recognition or translation with higher accuracy. Within our project, Artificial Intelligence can be used in two different ways: On the one hand it can be used to analyze communication between humans to make improvements, but on the other hand, it can also empower a separate autonomous communication partner.

Effective communication between people is largely dependent upon the quantity and quality of contextual information that is taken into consideration in the process. For a long period, synchronous telecommunication technologies (such as the telephone), only allowed users to exchange information with audio. Later, video conferencing technology introduced visual cues, which allowed people to better understand each other with body language, facial expressions, etc. To understand such cues people rely on their understanding of context. Like humans, an AI system could also be able to understand users' backgrounds and psychological profiles. Such a system will be equipped to handle much more contextual information than traditional communication, and therefore offers new potential for taking communication to a new level.

Communication problems between humans are often caused by a different perception of the same situation. Enhancing this communication also means changing the setup for the communication participants to enable a better understanding and less stressful interactions with fewer misunderstandings. Contextual information can be acquired by AI technology and made useful for that purpose.

2.3 Problem Statement

Artificial Intelligence (AI), coupled with Deep Learning, is one of the most vibrant technologies that drives changes in human life. Some notable examples of its spheres of influence include autonomous driving as well as image and voice recognition.

FutureWei is constantly seeking new methods of interaction between human and AI technology and we are prompted to design a system that enhances the user engagement with AI technology. Problems in human communication multi-layered and reach from simple misunderstandings to serious arguments and communication failure. We are looking for reasons why human communication fails and thereby especially concentrate on areas, where communication has a very high impact, like business or family life. We are investigating which factors impact user engagement in communication, and how AI technology can help in the process of human-machine interaction. Our goal is to create a system designed to enhance user engagement by using AI technology.

2.4 What is Design Thinking / ME310

Design Thinking is the latest successor in a long-ranging history of research and development on design methods. According to David and Tom Kelley ”Design Thinking is a way of finding human needs and creating new solutions using the tools and mindsets of design practitioners” [13]. Design Thinkers combine empathy for the context of a problem, creativity in the generation of insights and solutions, and rationality in analyzing and fitting various solutions to the problem context. Although many have tried to define Design Thinking [22], there is still uncertainty whether to describe it as a methodology, a process or a mindset. For the sake of clarity we define Design Thinking as a methodology, attributed by three core elements: multidisciplinary teams, an iterative process and variable space (Figure 2.1).

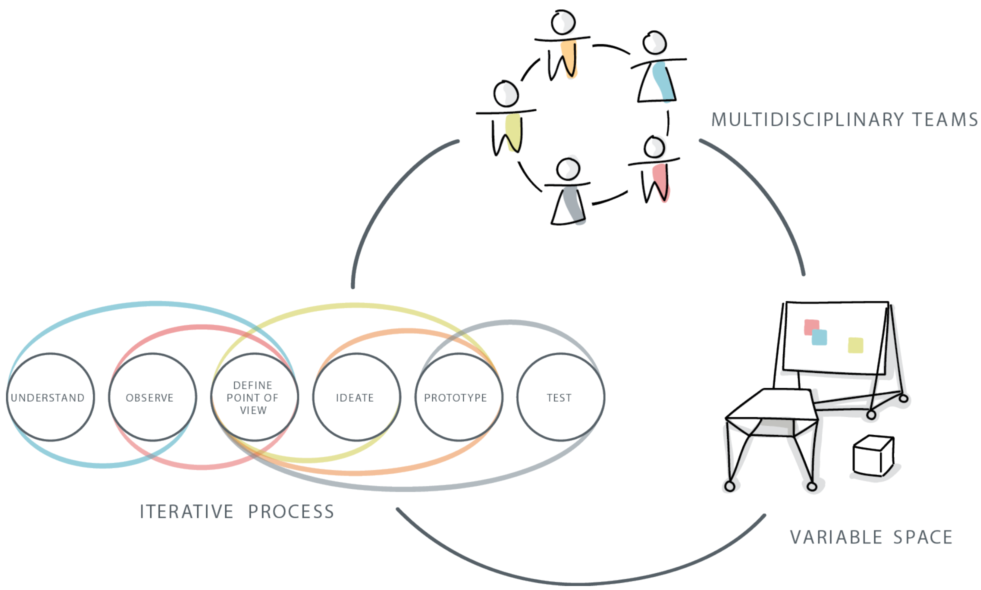


Figure 2.1: The three core principles of Design Thinking

2.4.1 Design Thinking Process

Kumar argues that ”successful innovation can and should be planned and managed like any other organizational function” [15]. Although his definition of the design process is different from Design Thinking, all approaches to design share great similarity in their attributes, e.g. being human-centered, non-linear, iterative and incremental. Brown explains the ”continuum of innovation as a system of overlapping spaces called inspiration, ideation, and implementation” [3], as opposed to a ”predefined series of orderly steps”.

We use this iterative process to find people’s needs and match those with what is technologically feasible and what is viable in terms of business strategy. The key to the process is the strict separation between problem space and solution space. Within the first three steps we only focus on finding, selecting and understanding the right problem, which might be counter-intuitive as humans generally tend to think in terms of solutions. The steps in addressing the solution space include the generation and selection of ideas, as well as prototyping and user testing. This overall process is carried out in iterations, repeating it completely or partially as required.

2.4.2 ME310

The nine-month graduate level engineering course, ME310, has was created over forty years ago at Stanford University and aims to provide engineering students with real-life engineering challenges. The name is derived from the course labelling scheme at Stanford and is short for *Mechanical Engineering 310*.

Today, the course is focused on teaching students the innovation methods and processes required for designers, engineers and project managers of the future [4]. Student teams work on innovation challenges posed by corporate partners over nine months. Throughout the project, students traverse an intense and iterative process of discovering problems and designing and testing various solutions in order to develop a functioning proof-of-concept prototype. The different phases (as shown in Figure 2.2) involve Needfinding, ideation, rapid prototyping and testing. The course starts with benchmarking and needfinding and continues with the development of exploratory prototypes to foster understanding. The first quarter ends with the autumn presentation and documentation. It mostly considers problem understanding and is defined as the *Problem Space*. The next quarter focuses on the exploration of the *Solution Space*, where the first parts of the final concept are developed and tested individually. The last quarter, labeled *Make it real*, serves to integrate the discovered solutions into a complex system prototype.

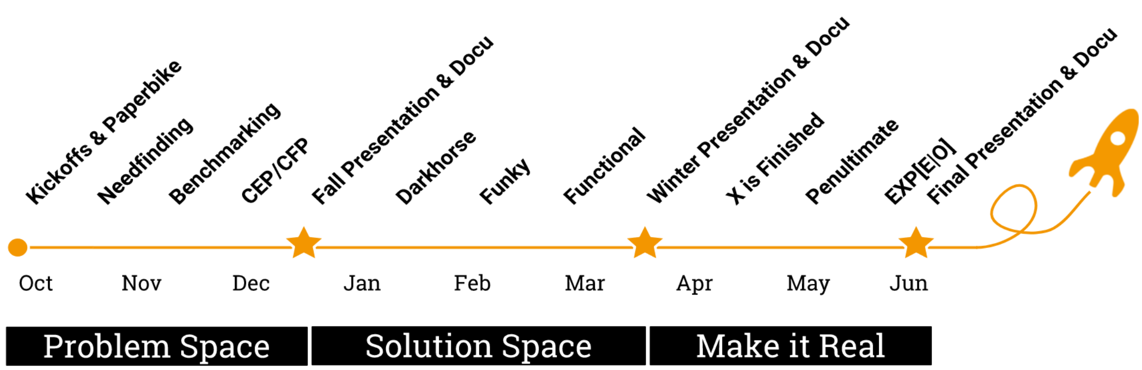


Figure 2.2: Visualization of the three phases the course ME310 is organized in

3 Design Development

During the fall quarter, we focused on the exploration of the design space. During the winter, we moved toward the solution space. We had one last exploration of the design space with a two week exploration of Dark Horse ideas before fully pivoting to the solution space. Our rounds of prototyping and testing components and systems combined with additional Needfinding landed us with a new and definitive user group and need. The last phase of the project development was in the spring. This began with the Stanford team visiting HPI for a week and converging on a solution. The rest of the quarter was dedicated to making the concept real and validating our design choices with user testing and interviews. The project culminated with a final presentation of our solution at EXPE at Stanford.
The following section contains a condensed view of the information from the fall quarter and winter quarter that we used as foundation for our work in the spring. In addition to a detailed description of the solution concept development and realization carried out in the spring. A visual overview is also shown in the Design Space Map (Figure A.1) as shown in the chapter A.

3.1 Benchmarking

We explored many potential research directions that involve existing AI techniques for communicating. The exploration directions include human remote communication, popular apps that make use of AI, personal assistants on smartphones, personal assistants on separate hardware and a comparison of underlying AI hardware systems. For each of those directions, there are a number of existing technologies that utilize AI or have a big potential to use AI in the future. The Benchmarking gives a big picture of our design space and a deeper understanding of commercially available products that are relevant to our project. The following sections give insight on various aspects of communication tools and AI-powered solutions, while Table A.1 provides an overview of AI providers. We start the discussion of different directions for Benchmarking with a recap of the work we did at the start of the course, which is the foundation for the continuing development of the design space during the current quarter. We also continued the research of products that have similar aims to our project direction. Parts of this work are included directly in discussions, where we compare them to our prototypes.

3.1.1 Remote Human Communication

A potential area of focus for our project is remote human communication. This could involve creating a new user experience that is augmented by AI technology. Without telecommunication technology, instant remote communication was almost impossible. Nowadays, a wide range of methods are available, e.g., phone calls, video chat through Skype and instant message using Facebook Messenger, WeChat, WhatsApp. Most of the techniques are based on smartphones and a reliable cellular network.

Almost every person in modern urban society is a user of these technologies. These communication methods make the connection between people easier and faster. People use these tools for different purposes, for example, business negotiations, social activities and emergency notifications.

In general, these tools increase the efficiency of communication and lower the cost of message delivery. As compared to a face-to-face talk, the remote conversation methods have reduced communication quality because of time delay, physical distance and cellular service. In some cases, a response is not guaranteed, for example, a call receiver may decline to answer the call and a text message recipient may ignore the message.

We also identified several extreme users of telecommunication. Air traffic control (ATC) personnel who constantly talk to airplane pilots via radio. Sometimes, because of poor radio quality and incorrect setup, the communication channel between ATC and pilot is not firmly built, therefore causing safety problems. An example of such a failure occurred in 2014 when the tower controller at Lindbergh Field in San Diego instructed a United aircraft to line up and wait for takeoff. After several attempts and no acknowledgement from the crew, the controller discovered that they were still on the Ground frequency: he first reprimanded them for that, only to find out that he was inadvertently attempting to issue the clearance to the ninth plane in the takeoff sequence. Had this plane attempted to takeoff the results could have been catastrophic.

Another example of extreme users are people with phone addiction (Figure 3.5). People with phone addiction spend more time on the smartphone than they realize and mindlessly pass time on a regular basis by staring at the smartphone. With excessive exposure to all kinds of smartphone software, people with phone addiction are a good example of an extreme user group.

Call center employees are another example of extreme users. They are likely to lose patience and get emotional after a long period on the phone [25]. Therefore, most call centers record the phone conversations to ensure service quality. People who run a commercial account on WeChat need to text to different people all day. As an instant message software extreme user, they suffer from finger joint pain and visual fatigue.

Instant message service providers compete to enhance the experience of communication. WeChat incorporates both voice call and video call in its app. WeChat also allows the user to make and send customized Emojis. Apple's iPhone X features Animoji [1] that makes a text message more vivid by mapping the user's emotion to a virtual "face" of an animal. Facebook's Messenger gives the sender a feedback when the recipient has read the message.

From the observation and analysis on these remote communication tools, we found the potential opportunity is to design a system to facilitate or bridge the telecommunication with AI technology.

3.1.2 AI-Powered Apps

In recent years, more and more popular smartphone apps utilize AI technologies to enhance the user experience. Besides the obvious application of AI to search engines, one main type of apps gives smart recommendations based on a user's viewing history, tailoring content to that user's taste. Another type is image processing with AI such as SnapChat, which adds real time filter and stickers to photos and videos. At the same time, other software providers are on the move to implement AI in their apps in order to take an edge on user experience.

Probably the most popular application of AI are web search engines that process a huge amount of data available on the web and order it to make it accessible by humans. Google as the most commonly used search engine is queried 3.5 billion times a day and uses this data to optimize its results and automatically completing search queries (Figure 3.1).

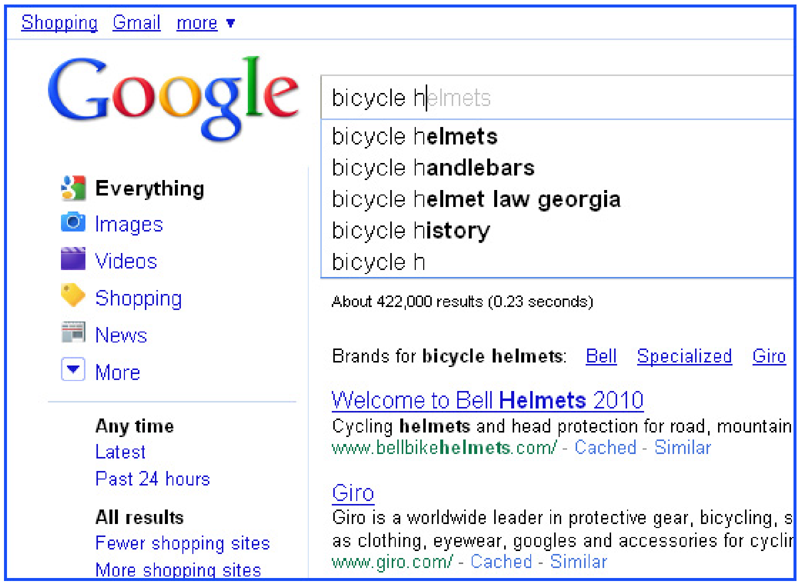


Figure 3.1: Google Search Completion is driven by AI

The recommendation principle applies to various news apps (such as Google News), streaming services for music or video (Netflix, Amazon Prime video, Spotify or YouTube), where a central part of the user experience is a system that filters and recommends new content according to past activity and user votes (Figure 3.2). The goal of these systems is to generate a meaningful personal profile of the user based on this data that makes it possible to suggest not only similar content, but also content from different areas that the user is most likely to appreciate.

Image processing with AI is a very common feature that is used to annotate images with content, for example adding a filter to one's face while taking a selfie (Figure 3.3). This feature enables the user to make fun with friends. However, it was originally designed to make users feel more comfortable sending pictures in situations, when they do not want to show their uncovered face. For example, a SnapChat filter might obscure parts of the users face when taking a selfie. Other apps or services (e.g. let‘s enhance [1]) use AI to simply process images improving the quality or removing artifacts without adding additional content. AI functionality can also be used to automatically adjust the parameters of the camera when taking a photo (as intended by Arsenal [2], an AI-extension for DSLR cameras). Using AI to restore information in an image that has been obscured might have some im-

[1] http://letsenhance.io - website of Let's enhance allowing users to upscale their images
[2] https://witharsenal.com/ - homepage of the Arsenal project

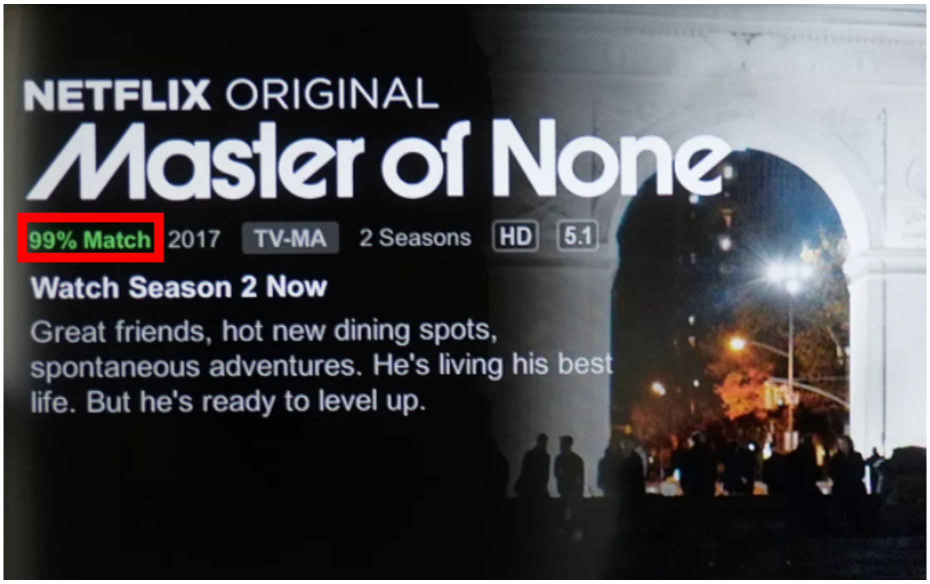


Figure 3.2: Netflix creates a user profile based on relevant data like votes and past activity to estimate scores for unseen movies or TV series

plications on the privacy of users. If parts of images have been obscured willingly to hide personal information, it is not wanted to make it visible again. Most notably, image processing can be used to automatically detect people or their faces, for example, in video surveillance. While this can be used to improve safety it also has important privacy implications and can be easily abused. Hiding your face behind a mask with a Snapchat filter provides another way of removing information from a conversation.

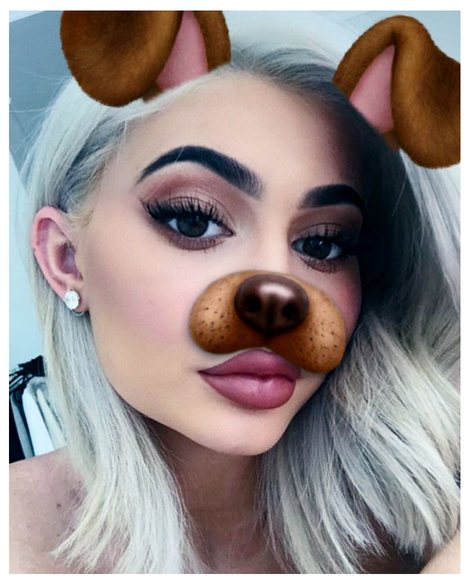


Figure 3.3: Snapchat selfie with filter added based on face feature recognition

AI technology is heavily used to improve the user experience for many applications. For example, YouTube [3] has about 1.3 billion users around the world. That means all these people are benefiting from the AI technology when browsing YouTube. There were about 2.1 billion smartphone users around the world in 2016, indicating that the market for AI empowered apps is gigantic.

In general, the AI provides app users a better experience. However, people also fear *Filter Bubbles*: a state of intellectual isolation that can result from personalized searches when an AI algorithm selectively guesses what information a user would like to see based on information about the user. Gradually, the user will be isolated from the viewpoints or contents they disagree with or prefer less. A good example is the 2016 U.S election. Since

[3]https://www.youtube.com - Website of the video platform YouTube

Facebook feeds are based on past clicks and likes behavior, a liberal New Yorker reports that "a few months ago my Facebook feed was filled with #ImWithHer or #FeelTheBern content in addition to some 'Obama is the greatest' headlines, which I was happy to see. [...] I noticed that the second most popular article shared on social media in the last six months with words "Donald Trump" in the headline, "Why I'm Voting For Donald Trump", had been shared 1.5million times. Yet that story never made it into my Facebook newsfeed." Within the Filter Bubble, this liberal New Yorker was totally not aware of the popularity of Donald Trump until he researched to statistically compare the popularity of Trump and Clinton. [8].

Some AI apps understand the content of texts written by users to provide additional benefits to the users or place adjusted advertisement. The business social network LinkedIn uses AI to inquire information about registered users for example to propose interesting groups or job offers to them. Google uses AI to process mail messages of GMail users to adjust the advertisement accordingly (Figure 3.4).

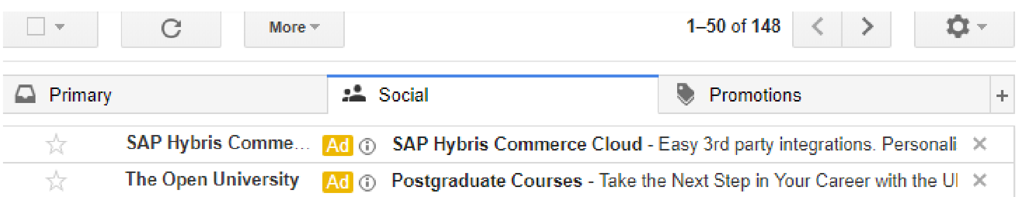


Figure 3.4: Advertisements are automatically placed by GMail

Figure 3.5: Phone addicts are extreme users

A potential design space is to build apps that make users feel more comfortable, feed the users with content they like without creating a "Filter Bubble".

3.1.3 Personal Assistants on smartphones

Most smartphones are equipped with personal assistants to provide the user a new means of interaction with the phone. Apple's iPhones have Siri. Equivalently, Android phones have the Google Assistant. They both can perform certain operations on the phone by listening to the instructions from the user, understanding the human voice and meaning with AI and then execute the tasks.

Any smartphone user can be a user of such a voice recognition and assistance tool. From the needfinding activities that we conducted early in autumn quarter, many users reflect that there are a lot of misunderstandings between the communication of a human and a voice assistant. The voice assistant usually can only perform very simple tasks or answer simple questions. Any task that is beyond its capability will be redirected to a web search, which is not expected by the user, who would prefer a human like answer from the assistant. Also, some smartphone users indicate that they feel uncomfortable to talking to Siri, an imaginary person, especially in public.

The extreme user in this category is likely to be a user appreciating the hands-free advantages, such as a business user. They can use the voice assistants or dictation features if they need to communicate while driving a car so that this time can be used more productive.
We found a potential design space to make the voice assistant tool more human like. That means a human voice is not enough. We can potentially make the voice assistant more human in terms of appearance and other human senses.

3.1.4 Personal Assistants on separate hardware

In addition to the voice assistant on a smartphone, tech giants are extending the developed voice assistant tool to smart home devices and creating a hand-free voice assistant experiences. Amazon's Echo, Microsoft's Invoke Cortana and Google Home are examples of personal assistants built on a separate hardware (see Figure 3.6).

The user group for this category is much smaller than the smartphone personal assistant users, since people purchase such device only if they are certain they want such an omni-present personal assistant at home.

The AI technologies used on these devices are similar to those in the smartphone personal assistants. Their problems and user experience are also similar. People have additional concerns about privacy since these devices are usually placed at home and are able to collect sound all the time.

We found a potential design space to make a new system or experience that assures the users that their privacy is protected when they want the personal assistant to stop listening at home. However, most current users like that their device are available without a touch and use that to get answers for simple questions, to update the shopping list, to control the smart home or simply to control music playback. We think it might be complicated to offer an increased privacy while maintaining the functionality which users like.

Figure 3.6: Google Home in a Kitchen

3.1.5 AI Hardware

In October 2017 Huawei presented the new Kirin 970 processor, which includes a Neural Processing Unit (NPU). The Kirin 970 is mainly used to reduce power consumption, while speeding up computations involving AI up to 25 times. It enables real-time AI processing on the device and thus also enhances privacy, as the sensible data of the user does not need to be sent to the cloud for processing.

In 2012, Huawei established a separate research lab for AI and other similar research areas, called Noah's Ark lab. One goal of the researchers at this lab is, among others, to make AI applicable for everyday uses while applying it to solve problems in data mining, social network analysis, natural language processing and video coding. The impact of this research outside Huawei however is still relatively low, as most publications receive few citations (less than 100). Some patents that are

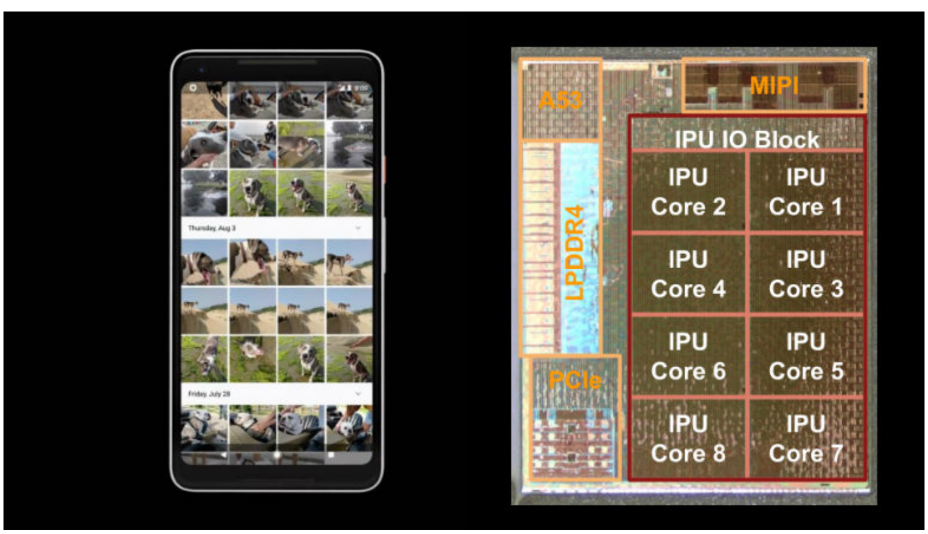


Figure 3.7: Google's Pixel visual core includes eight image processing units.

connected to the field of AI cover the topic of neural networks (SON: self organizing networks), crowdsourcing platforms and a mobile computing infrastructure.

Competitors also presented dedicated AI hardware for smartphones and servers in the last few months. Examples of other mobile computing platforms focusing on AI are the Apple A11 Bionic, the Pixel visual core (Google, Figure 3.7), and the Hololens HPU (Microsoft).

In the area of server computing platforms, the universal high performance chips like the Intel Phi or Nvidia Tesla have been supplemented by dedicated AI hardware like the Tensor Processing Unit (TPU), the Intel Nervana or the NVIDIA DGX-1.

In general, we observed a drift from general purpose processors (CPUs) to specialized hardware setups for tasks like AI. Other companies chose to go the same route as Huawei; starting to incorporate powerful computing units in their respective devices doing more and more computations on board without the need to transfer all data to the cloud. Nevertheless these approaches are all still very new and thus the dedicated hardware still lacks the appropriate software to efficiently make use of it. Today, specialized AI hardware can be found only in very recent devices, where it is applied to image recognition and camera calibration [2]. In addition, this allows developers to use AI computations directly on the smartphone satisfying the demand of users for increased privacy.

3.1.6 Translation applications

Language translation has been a work in progress for almost a decade. Great leaps have been made in this area. Google Translate (c.f. Figure 3.8) made its debut in 2006 and since then has become one of Google's most reliable and popular assets; it serves more than 500 million monthly users in need of 140 billion words per day in a different language. It exists not only as its own stand-alone app but also as an integrated feature within Gmail, Chrome and many other Google products.

Recently, Google Translate got converted to an AI-based system for much of its traffic [17]. The new incarnation, to the pleasant surprise of Google's own engineers, was completed in only nine months. The AI system demonstrated overnight improvements roughly equal to the total gains the old one had accrued over its entire lifetime. The new system, a Deep Learning model known as neural machine translation, effectively trains itself – and reduces translation errors by up to 87% [21].

Google translate also incorporates a feature that detects foreign languages in images and instantaneously overlays the text with its translated form (see Figure 3.9, [21]).

Figure 3.8: Google Translate, Google's translation software that uses AI

The latest development in language translation technology in the form of hardware has been the Google Pixel Buds (c.f. Figure 3.10, [24]). These are headphones that claim to translate up to 40

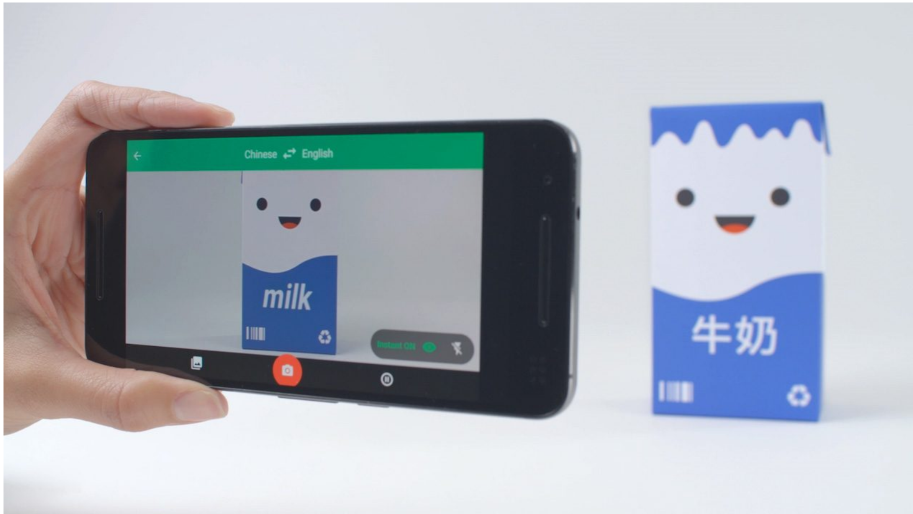


Figure 3.9: Google Translate's feature that translates foreign languages in images.

different languages through the earpieces in real time. But through actual user reviews, this technology does not live up to the hype. The translation system does exist but is extremely slow and full of errors.

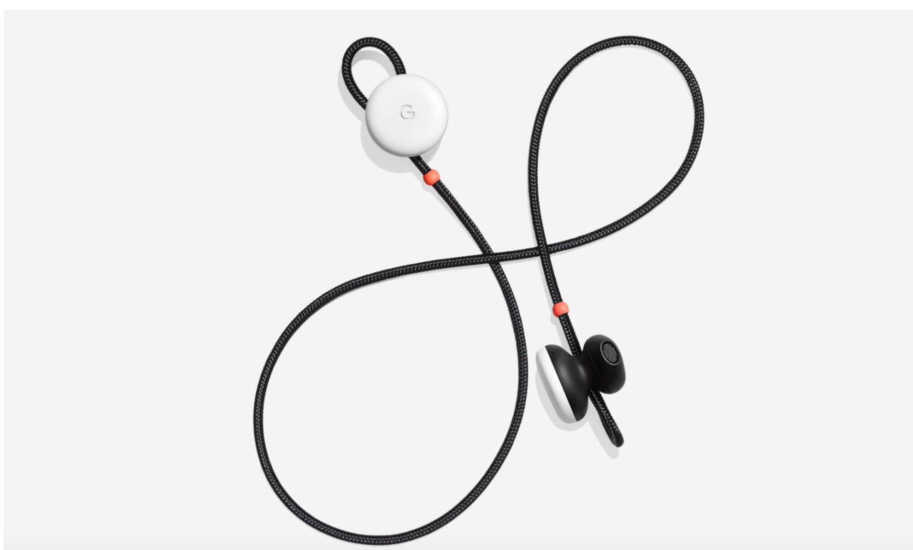


Figure 3.10: Google Pixel Buds, the latest headphones with real time translation

3.2 Initial Exploration

To learn more about potential user groups, we conducted interviews with a wide variety of communication experts, users and especially users in extreme situations, which often face more extreme versions of the same problems. We also built some initial prototypes, to gauge the reaction for new concepts we thought of. In this section we will describe the interviews and prototypes, which helped us reach an initial understanding of the problem domain and what we learned from them.

3.2.1 Kindergarten Child Interview and Observation

Figure 3.11: Ziru Zhou, a kindergarten child from China using a Voice Assist for her web searches.

Ziru (Figure 3.11) is a 6 year-old kindergarten student in Suzhou, China. She is a Mandarin speaker and likes drawing and origami. Usually, she uses the internet for art and origami tutorials. She told us that since she does not know how to type Chinese characters on the keyboard, she has been using Voice Assist to help her make her web searches.

During the FaceTime interview, she showed us using iPhone's Siri to search the methods of making a rabbit origami. Since she was speaking slowly and articulated her questions clearly, Siri caught every word that she said. Siri prompted several websites with shortcut screens. She randomly opened one of the pages and was not satisfied with the content. It is observed that she always used the Home button to return to the main screen and then activated Siri again to refine her search by adding some adjectives like ”beautiful rabbit”. When talking to Siri, she acted like talking to a real person by saying phrases like ”please help me find”. However, on several occasions, Siri did not catch Ziru's words and prompted some irrelevant pages. She always used the Home button to deactivate Siri and then activate Siri again. Not only did Ziru call Siri directly from holding the Home button, but she sometimes also opened the browser directly and used voice input. In this event, she seemed to know that it is not a real person and only spoke keywords for the search, e.g., ”beautiful rabbit origami” rather than ”hi, can you find a beautiful rabbit origami for me?”.

3.2.2 Interview with communication expert

During our Needfinding, we also talked to Eva Bilstein[4], a communication expert (Figure 3.13). We asked questions about her work and the experiences she made coaching many different people, but she also introduced us to some general background knowledge about communication and especially to associated communication problems.

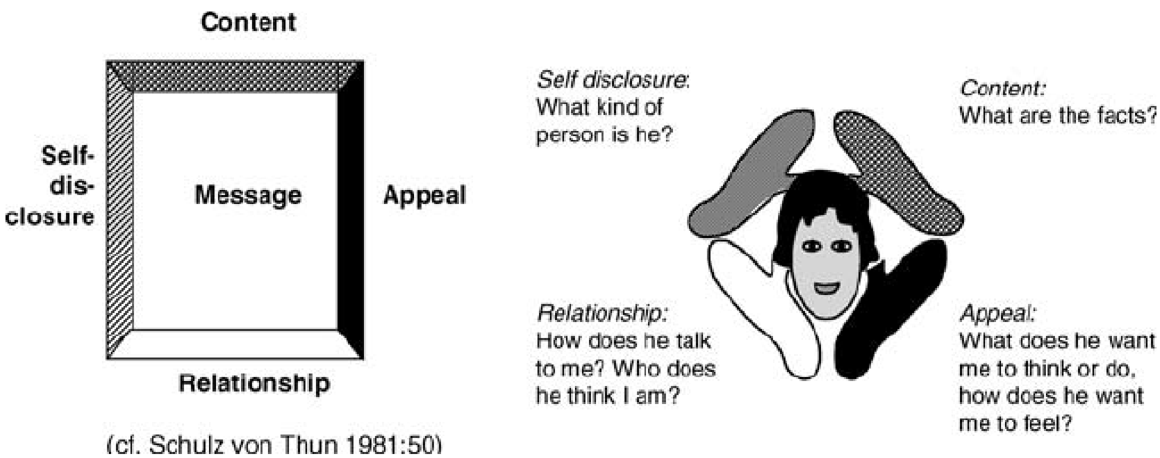


Figure 3.12: Each message in a communication transfers information on multiple layers, according to the communication model of von Thun

During the interview we realized, that communication conflicts are almost always emotionally loaded and carried out on an emotional level, that exists in addition to fact based communication. Following this principle, the most popular communication models organize communication in several layers, the main layers are self-disclosure, relationship, content and appeal (Figure 3.12). When people communicate, all of these layers are involved, even in text based communication.

Communication conflicts played an important role during our interview, because we wanted to explore in the ways we might be able to design an intelligent system to sidestep some of these issues in human communication. Eva Bilstein told us that conflicts are very common and are often not even recognized by everybody who is involved. They occur at any time and mostly result from different needs, that are in some way not satisfied. Being overly polite or reserved can be problematic as well and might lead to misunderstandings. A good way to overcome communication conflicts is to be patient, brave to actively designate existing problems, listen carefully and be open to the other person. It is also possible to learn empathy. From that insight, we generated the idea of an electronic device as a communication trainer, which we did not develop further, but it still provides an interesting future direction. An intelligent app could, for example, train users to listen carefully by telling them a story and asking detailed questions afterward.

Eva Bilstein told us that Emojis are very important for understanding text messages correctly, as plain text leaves a huge space of possible interpretations on all communication layers that need to be considered by the receiver of the message. Smileys are a great help to find the right interpretation.

[4]Website of Eva Bilstein - www.eva-bilstein.de

Figure 3.13: We talked to the communication expert Eva Bilstein, who is a professional coach for conflict and project management. (Photo by Thomas Schaaff, Institut für Aufmerksamkeit)

Text messages in general are a very bad option to settle a conflict and prone to misunderstandings, as facial expressions, voice and gestures cannot be observed and much information gets lost on the way from sender to receiver. As AI is already able to detect emotions reliably in some contexts, we found it interesting to explore the possibility to enhance text based communication with additional information, for example about the mood of the participants.

3.2.3 Smartphone shoppers

During our interviews at a local electronics store (Figure 3.14), we realized that the camera is the most important factor to consider for many people who buy a specific smartphone. People did not know what most of the features of the recent smartphones were for, which matched the information we gained talking to a phone store employee, who told us that most customers do not know what their smartphone can do and they repeatedly come back to ask questions how to use it.

The shoppers we interviewed did not reflect their communication habits very well, while they generally preferred chatting over calling in many everyday situations for convenience. They often wished to talk to others face-to-face when the topic is very sensible or private, and do not even want to make a phone call in this case. Again, these insights led us to the conclusion that annotating text communication might be interesting.

We learned from smartphone shoppers that many of them wished for more contextual information in calendars, chats or navigation apps to organize their daily life better. However, none of them did use any personal assistant like Siri, Google Assistant etc. We were told that existing personal

Figure 3.14: We conducted interviews of smartphone buyers at the Saturn shop in Potsdam

assistants often misunderstand questions and instructions and are thus not very helpful for everyday use (c.f. Figure 3.15a). Another problem is that they have to be explicitly activated, while it sometimes might be more helpful to get additional contextual information while being involved with other tasks like writing text messages or even in a real life conversation.

We also asked people, who bought a Feature Phone (Figure 3.15b), why they decided not to buy a smartphone, and what future smartphone features might be needed for them to change their opinion. We found that most Feature Phone shoppers were old people, who do classical smartphone tasks like writing emails with their computer and who do not send short messages while on the go. They do not care for being connected all time, but want to be able to call for help in case of an emergency. They also find it helpful to coordinate each other, but this is a task they can perfectly solve with a simple text message or a quick phone call, so a smartphone is not necessary. Most of them did not use the Feature Phone for small talk, but some used it to connect to their children and other people of younger age. Other customers we met bought a Feature Phone to hold a second SIM card instead of a dual-SIM smartphone or for outdoor activities, where they did not want to take their smartphone with them (due to size or a potential damage).

3.2.4 Managers

During our research, we realized that business communication is very different from private communication, as business partners need to discuss different topics under different circumstances than friends or family members. In general, business communication logically has a higher demand for efficiency, and should primarily focus on the factual content. Our interview with the communication expert (subsection 3.2.2) showed that business communication also involves emotions. We conducted two additional interviews to further investigate the subject of business communication. Our interviewees were Dr. Matthias Uflacker, the representative of Prof. Plattner's chair at HPI and Harald Fladischer (Figure 3.16), head of sales and marketing at neXenio[5], a Berlin based startup.

[5] https://nexenio.com - Website of neXenio

(a) Siri often misunderstands users and their requests.

(b) Feature Phones are most frequently used by elderly people.

Figure 3.15: We asked smartphone shoppers about Siri and feature phones.

Our interviewees told us that communication problems in the context of business communication are not easily reproducible, as they highly depend on the very specific situation. Each of those problems is unique, which makes it necessary to find an individual solution. There are however some general situations in which communication problems most frequently occur, for example when there are changes in the staff of one team. Many issues and misunderstandings are caused by not communicating enough. If there is an issue and it gets recognized, it can often be easily solved, the hardest part of solving a conflict is to make both parties aware of the conflict.

The managers we interviewed also described that the communication heavily changed in the past few years, blurring the border between private and business contacts. While many customers request information online, they wish to have a personal contact at the company and like to have a direct communication with their sales agent. Therefore it is important to maintain a regular contact to business partners, which may also happen over social networks like LinkedIn or the chat app Slack, according to both managers. Video or phone calls are often problematic, because of technical problems and security restrictions of many companies, but are a good way in private use cases to keep contact to family and friends on a business travel. Meeting in person is important when a new collaboration starts and mostly helpful for getting to know each other, to build additional trust and for networking. This holds true even if a large distance has to be travelled because participants feel uncomfortable in a video chat if they do not know their conversational partner good enough. While the agenda of networking events is not very important, as most of the interesting talks take place away from the main events in a private atmosphere, work meetings can profit from an analog or digital visualization, depending of the meeting type. The visualization may help to capture and agree on important parts of the meeting and is a good take-away for both conversational partners. Especially after negotiations, it is essential for both parties to agree on future goals and results that need to be written down and acknowledged by everybody.

Even though it is important to communicate often to prevent misunderstandings and conflicts, it can be very inefficient to go in a conference without a very concrete plan. As meetings keep the participants from working on their tasks, they should be short and strictly organized. The responsibility of the participants and their respective goals and tasks should be made very clear. It is especially important to be well prepared for meetings with new business partners and also set up an agenda before the meeting. If the meetings occur regularly and the topics are clear to everyone, less preparation can be sufficient. Having as much contextual information as possible in advance helps to understand the other party and to be negotiate more effectively.

Another interesting point during the interviews was the usage of personal assistants. While Dr. Uflacker has a secretary that is involved in every organizational part of his work (e.g. to determine and prepare appointments and to organize all kinds of purchases), Mr. Fladischer is one of the heaviest users of speech recognition and voice assistants we met. He described to use Siri and the iPhone's dictation feature to check for and reply to emails while driving on a highway or to make phone calls with customers or the office. At home, he uses an Amazon Echo to manage a groceries shopping list that is always available to him or his wife on a mobile. With Amazon Echo, he has the same feeling he wants his customers to have: Talking to someone who understands and takes care of a problem makes his life easier and helps to him feel closer to a solution.

Figure 3.16: Harald Fladischer, a sales and marketing manager at neXenio

3.2.5 Interrupting AI (CEP)

During our Needfinding phase we noticed that AI assistants such as Siri were not used by the majority of interviewees, so we thought about increasing usage by creating a more proactive voice assistant. This assistant would not wait for questions to be asked, but instead, provide useful information during a conversation, even before the participants know, that they might need help.

3.2.5.1 Procedure

Figure 3.17: Ongoing conversation with an added AI device on the table (left) and the ”AI” behind the device, which is Felipe listening and talking in another room (middle). The AI device is a phone with the screen facing down (right).

The prototype (Figure 3.17) consists of a voice call between two phones. One of the phones is lying on a table next to two people having a conversation on an arbitrary topic. The other phone is behind the scenes with a human listening to the conversation and only speaking if they have additional information to enhance the conversation. In some cases they also interrupted one of the participants.

For each testing round, we recruited at least two English speaking people and asked them to talk about a topic covering some discussion of schedule or trip planning. Most of the time, the participants decided to plan a short vacation to a foreign city they have never been to. After some time, when the participants felt comfortable with the situation in general, and as soon as we got enough context of the conversation, we joined the conversation and added related information to the conversation through the established phone call.

3.2.5.2 Questions

- If people engage with the AI during the conversation will they recognize it as a person?
- Are the AI’s interruptions in the conversation helpful or desired by the users?
- How do people feel when their conversation is interrupted by AI?
- Are users concerned about their privacy when they know that an AI is actively participating in their conversation?

3.2.5.3 Findings

We noticed that in most cases the conversation was influenced heavily by the AI. While some people continued their conversation after being interrupted, others had conversations centered around asking the AI questions and never returned to the previous topic they had been discussing. Even the participants that continued their conversations asked the AI for help later, often speaking louder and clearer to make it more obvious that the question is directed at the AI. The interruption itself was not seen as annoying, but as helpful instead, if and only if the information was actually useful and not off-topic. For that, the information also had to be provided immediately before any change of topic had occurred.

When asked about long-term usage the testers, who had only participated in a short test, had few concerns about privacy, but either wanted to be able to turn it off for specific conversation or have it off by default and only enable it sometimes. In particular, they thought it might be helpful in planning conversations. Many testers also requested a screen for showing some of the provided information, similar to an Amazon Echo Show[6].

3.2.5.4 Evaluation

In general, the prototype helped with user engagement to virtual assistants, but we have yet to find out if that would continue after the novelty wears off or if the assistant is not as helpful as a human acting as an assistant.

If continuing in this direction a way to present the information from the assistant visually is needed as many testers asked for pictures or just notes of previous information. Perhaps this problem could be solved by sending the information to the user's phone. It would also be interesting to investigate with more detail, in which situations a conversation can be interrupted, e.g., analyzing the user's mood or finding out if natural pauses can be found to provide less urgent information.

3.2.6 Emotional Chat (CFP)

During our interviews and previous CEP, we found out that with text based communication (and even calls) important parts of the exchange get lost (Figure 3.18). Communications expert Eva Bilstein explained that communication happens on at least four layers: Factual information, Appeal, Self-revelation and Relationship. Even text based communication will transport information on all layers, but due to partial information loss on some layers, conflicts occur more frequently than during face-to-face conversations. Therefore, we thought about ways in which we might enhance textual conversation to include some of the lost information. A vital part of this is transmitting emotions, so we tested a way to detect a user's emotions using a webcam.

3.2.6.1 Procedure

We implemented a simple chat service allowing multiple users to communicate via text messages. To keep this simple, we built a website that regularly loaded a list of all messages sent to the server.

[6] An Alexa-enabled device with a screen (similar to a Tablet) with a fixed position on a table

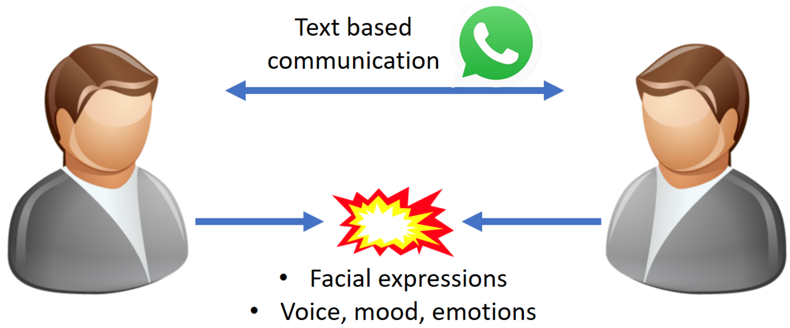


Figure 3.18: When people communicate via text messages, most of the information that is included in person-to-person communication gets lost.

The website was implemented plainly with php and javascript. We used the freely available web host square7[7] for our initial testing. As we uploaded the webpage to an online web host, we could ensure that it is accessible from every computer at any time.

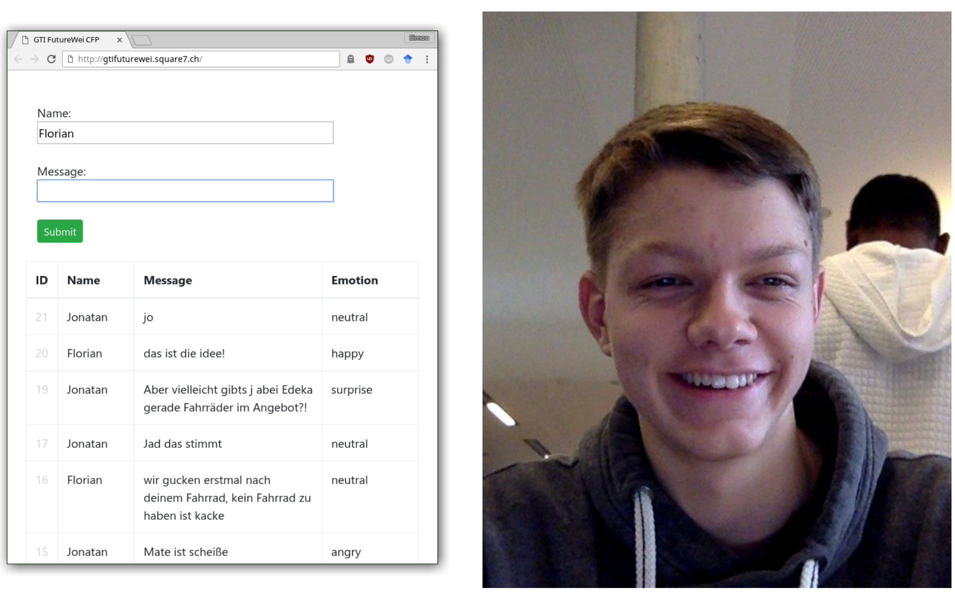


Figure 3.19: Screenshot of the chat service and a photo taken by the webcam while chatting.

In addition to transferring the message to other participants in the chat, a photo of the author is recorded using a webcam and sent to the Microsoft Azure Emotion API[8], which returns a list of possible emotions and their probabilities (see Figure 3.20 for details). The photos and detected emotions are store in an sqlite database for later analysis. The chat messages are then annotated with the selected emotions (Figure 3.19).

[7] https://www.square7.ch - Square7 web host
[8] https://azure.microsoft.com/en-us/services/cognitive-services/emotion/ - Microsoft Azure Emotion API

In total the Azure API can distinguish between eight emotions, e.g. happy, neutral and surprised. We selected the emotion with the highest probability, but added a small bias to only select "neutral" if it was the highest ranked emotion with a probability of at least 70%. Otherwise, we ignored "neutral" and selected the emotion with the second highest probability. With this approach, we ranked small emotions higher compared to using a uniform distribution.
During our tests, we found out that this change was necessary as otherwise most of the communication would have been annotated with the neutral emotion.

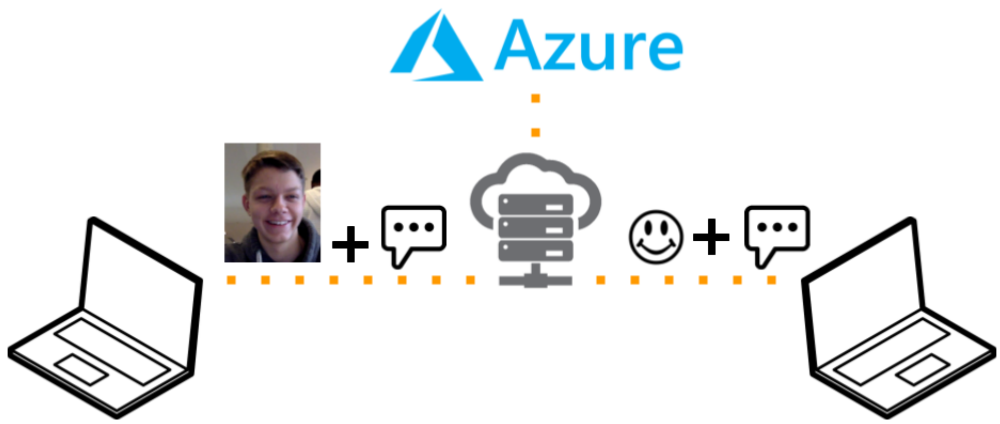


Figure 3.20: The technical architecture of our CFP visualizing the message flow and usage of Microsoft Azure.

We tested the prototype by observing users and comparing the resulting emotions with our subjective impression and the users' self-reported feeling. Additionally, we asked further testers to determine emotions from the webcam pictures and compared these with the emotions detected by the API. We also measured the time necessary to process an facial expression. After the test we talked about the experience with our testers.

3.2.6.2 Questions

With respect to the function of face recognition, we were interested in the following questions:

- Can an AI system recognize facial emotions using a webcam image?
- Can we detect emotions fast enough to annotate text messages in a real-time application, e.g. an instant messenger?
- Are the emotions expressed while sending text messages expressive enough to get recognized by the system?

Since this prototype also had an interesting user experience aspect, we posed some additional questions:

- How do people feel about their emotions being shared?
- Do people trust the emotions transmitted from conversation partners?
- Do people try to fake emotions to manipulate the system?

3.2.6.3 Findings

With an average transmission time of 3 seconds (taken over 10 trials) our annotated chat messages are fast enough for use in real time contexts. Regarding the correctness even our human evaluators differed widely among themselves about what emotions were shown. Our three evaluators only agreed on a single emotion in 8 of 22 images we showed them. Of those 8 images the API returned the same result in 5 cases. Apart from neutral the emotion shown most often was happy, on which the participants often agreed in their self assessment. We also noticed that the testers showed the strongest emotions not while sending the messages, but instead while typing (sometimes multiple emotions for longer messages) or receiving a message.

When their own emotions were actually misclassified, testers became annoyed with the system, but they always trusted the result for their conversation partner anyway, because ”it is the only hint available”, as one tester put it. Another tester suggested to use this system as a quick way to find emojis, but only as suggestion, that they can approve or edit. At some point in their conversation, most testers also tried to invoke a specific result by making the corresponding face.

One big problem with the prototype as a whole was the privacy of users. We had huge problems finding testers, as many people had an issue with having their photos taken by the application. Maybe that would be different when testing outside of Germany, which is notorious for its citizens' privacy concerns.

3.2.6.4 Evaluation

Our CFP showed that detecting emotions for instant messaging is possible, but still not very accurate. Human testers opinions on the other hand were also far from unanimous, so high accuracy might not even be possible. As we tested the CFP with other users, we were also able to draw conclusions about the user experience, therefore the prototype was also incredibly useful as a CEP. We learned for example that working with emotions in textual communication is an interesting field with many open problems.

Despite most users initial concern about privacy, most users found it both, interesting and useful later. They thought that the displayed emotions probably show the real mood of the person, thus unmasking a text with a different intention.

3.2.7 Image Based Translator (CFP)

For this Critical Function Prototype we wanted to explore another extremely prevalent application of AI that promotes positive human communication: Language Translation. Using the example of Google Translate, this application is already well developed. One problem with this form of translation is that it requires the user to know what language will be understood by the receiver. One way to solve that problem would be to have all languages translated to one universal form understood by all: images.

3.2.7.1 Procedure

We selected three different scenarios: a recipe for making tea, a person at a zoo and a person at a restaurant. Regarding the zoo and restaurant scenarios we wrote sentences someone might say or hear at those places in Portuguese and Mandarin respectively. For these two the sentences were unrelated. For the tea recipe we used the whole recipe and broke it up into its chronological steps.

We then prepared strings of images to directly represent each word in the sentences of the various languages. We selected the images from the image search functionality of Google for each word of the sentence in the exact order in which it appeared in the sentences. An example of this can be seen in Figure 3.21.

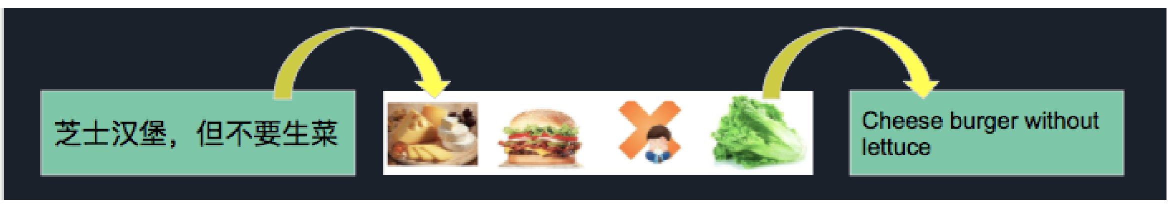


Figure 3.21: Steps in translating a sentence. Translating a sentence into images and then extracting meaning from the images into any given language.

We asked testers to be on the receiving end of the communication to see how accurately the message was conveyed only through images. By identifying moments of failure, we improved on our method of selecting images to convey messages more accurately.

Few assumptions about the level of intelligence of the system:

1. Ability to eliminate some words of the sentence. If helpful images were not available.
2. Ability to make a selection of the best suited image among the first few images suggested by Google.

3.2.7.2 Questions

Through this Critical Function Prototype we wanted to answer the following questions:

1. Is it possible for an intelligent system to facilitate accurate interlingual communication i.e. function as a translator, by translating words to images instead of words to words?
2. To what degree can the intelligent system self learn to improve the accuracy of the translation?

3.2.7.3 Findings and Evaluation

This prototype showed us that it is possible to convey meaning through images into a different language. However, it does have its limitations. Sentences that worked best were ones that were short and objective with one noun and one verb. A challenge we observed was that word order is different in different languages. While English is a subject-verb-object (SVO) language, Hindi is subject-object-verb (SOV). This means that the order the pictures appear will be confusing to an English speaker if translated from Hindi. One possible improvement would be for the AI to figure out what the best order to display the images is depending on the languages and not only what image to choose. Another approach would be to train the AI to convey sentences or phrases instead of individual words.

Another limitation is that of difference in cultural context. Some themes use completely different words in different languages which result in images which cannot be interpreted across cultures to mean the same thing.

In some cases the users also had difficulties in moving beyond the detailed imagery in order to interpret the message the system was trying to convey. This stresses on the need for extremely specific AI training for image selection as any wrong selection of an image can have a negative impact on the communication. In Figure 3.22 an image is used to represent the question posed by a Portuguese speaker in a zoo, ”What time does the penguin eat?”. Due to the details in the image, the user gets confused and fails to understand the question.

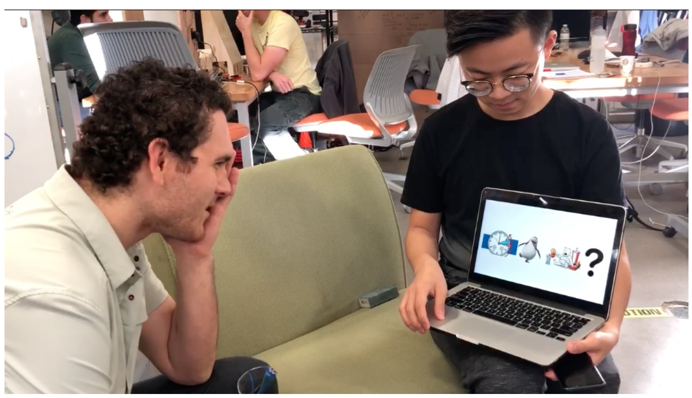


Figure 3.22: The tester has difficulty understanding the image based translation of ”What time does the penguin eat?”

We also recognized the scope for improvement through self learning. The set of images shown below illustrate this. The left part of Figure 3.23 was selected manually as part of the translation of a recipe in Hindi to convey the message to ”Turn the stove off”. When tested with a user, the message was completely misinterpreted as a cassette tape. After the first failure, we learned what an improvement could be and made a correction as shown in the right part of Figure 3.23. This shows that the AI would be capable of increasing its accuracy through self learning – by identifying mistakes by identifying kinks in the flow of the conversation and choosing new images to use in that context.

Figure 3.23: Different visualization of the phrase ”turn the stove off”. The left was misinterpreted, while the right one is more universal.

3.2.8 Dark Horse

In order to explore more risky solution directions and design spaces, areas that were not initially obvious and orthogonal to our previous direction, we dedicated two weeks to prototyping ’dark horse’ ideas. On the one hand, we focused on personal therapy and human communication with machines as a method of self-reflection or therapy. On the other hand, we looked into an AI making decisions for users.

For the first new prototypes, we tried to find the deep underlying assumptions of our previous work and turn them around, in the form of questions like ”What if X were different?”. Some very important assumptions we identified were:

- What if communication via thought was possible?
- What if AI was actually intelligent?
- What if people were OK with sharing everything in every interaction, i.e., never lied and never omitted everything?
- What if AI knew how a recipient would react to a message?
- What if an AI makes decision for a user, like when to initiate communication, or what to write?
- What if AI served as a intermediary in the process of self-reflection or therapy?

We decided to examine the last two assumptions further as they were the most tangible and allowed for a huge range of ideas.

We challenged some of those assumptions listed above by building *dark horse prototypes*, which ignore an important constraint or assumption and through that allow more divergence and make it possible to find insights ignored otherwise.

Some further ideas focused on other questions. For example, we thought about an AI system for business users that could check an email prior to sending it to a conversational partner. This approach would validate the manner and content of the message in the context of the recipient and their cultural background. This could help to minimize conflicts based on the cultural context or to prevent misunderstandings due to lacking information. The system we thought about could either make suggestions or edit the message right away (in a form of control-reversal, see subsection 3.2.9). In addition, it could also be used for incoming emails so that the recipient gets the message in a way, they could work the best with.

Another idea was to create a lie detector that could help in (business) conversations. It would automatically detect lies or parts of a text where crucial information is missing and mark those in a conversation. This could be useful for meetings in order to speed them up and help participants to focus on true and not misleading information. For our testing, we would have been interested in the reception of this idea by testers. In some aspects, one of the main questions is similar to the one we answered with the control-reversal prototype (subsection 3.2.9): What if the AI would be smarter than humans and could control / initiate human actions?

3.2.9 Control Reversal Prototype

With most currently used technology it is always the user who takes control over the device and not the other way around. This is based on the belief, that the users know what they want and what is best for them. With increasing quality of Artificial Intelligence, this might not stay true. Even today, there are already areas like security updates on Windows, where control is taken away from a user. The range of decisions, even in fields relevant to communication, is large and depending on the specifics user acceptance might vary wildly:

- Major life planning, e.g., what job to choose
- When and with whom to initiate a communication
- Communication topics
- Means and style of communication form, e.g., punctuation, level of politeness, or which channel to use

It is important to note, that some corporations already let algorithms make decisions about loans and hiring among others and through that, decisions of normal people are already taken away from them [5]. However, the applicant cannot distinguish between an algorithm and a human inside that company. In our first dark horse prototype we took over the decision of the means of communication, by forcing a call while chatting.

3.2.9.1 Procedure

We gave two participants a remote communication task to accomplish, which we planned to be more easily solvable by a call than by chat. They were required to find three time slots in which they were both free in busy and complicated schedules (Figure 3.25a) we handed them. We separated the participants and each of them was given a specially prepared phone, with a chat application already opened, to discourage finding our intended solution at the start already. We used the app AirDroid[9] to remotely control an Android phone from a web browser (Figure 3.25b). This gave us the options to initiate calls, send SMS and monitor the screen content of the phone, which we used to fake the

[9] https://www.airdroid.com/ – Website of the AirDroid software

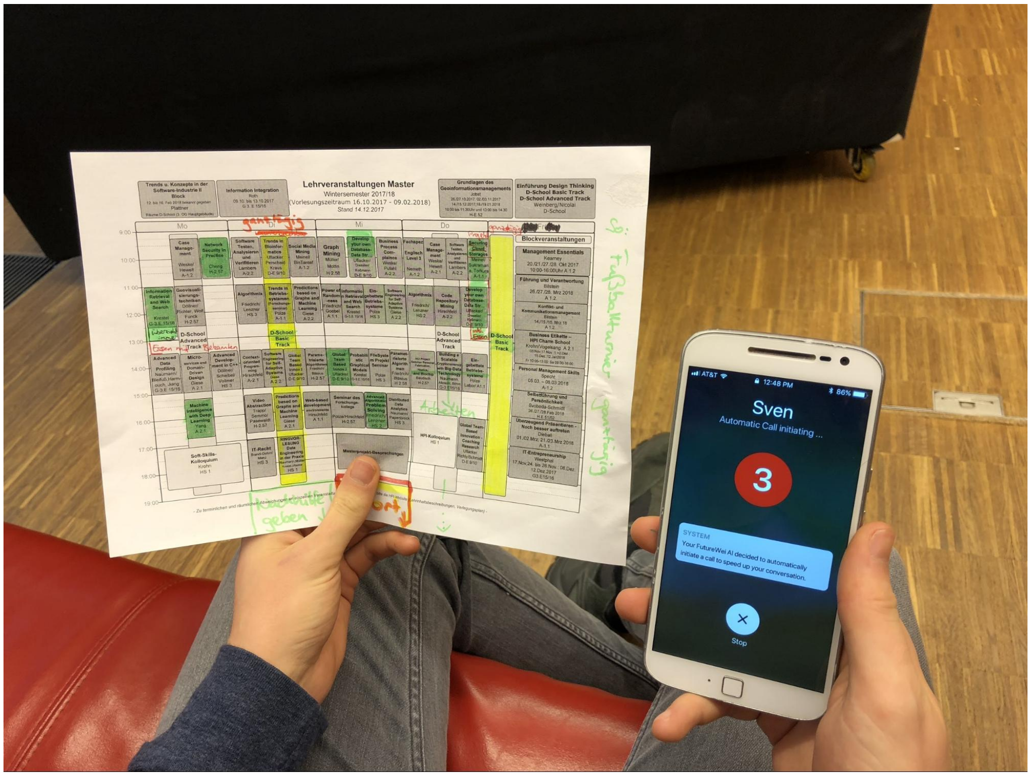
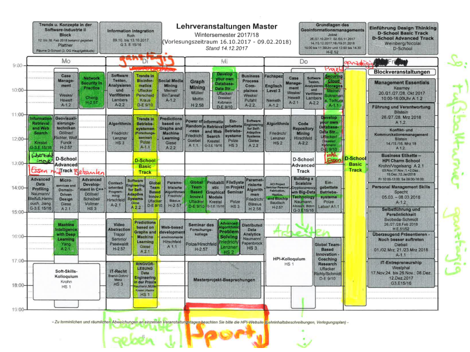
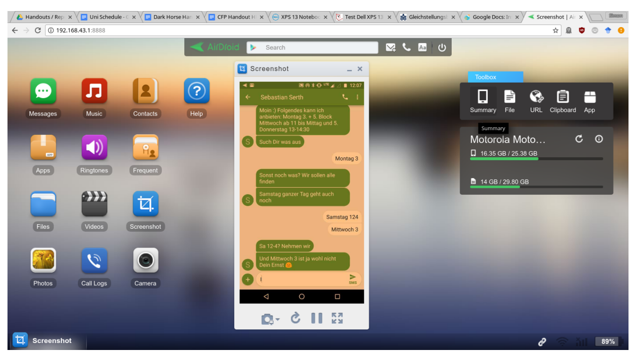


Figure 3.24: A smartphone automatically starting a phone call while a user tries to find a suitable time slot for a meeting.

AI. In addition we used the messenger app Telegram for the chat, as it was easy to insert another account into the conversation, which acted as the System AI giving a countdown prior to starting a call (Figure 3.24).



(a) One of the two schedules in which the participants looked for common time slots.



(b) We remotely controlled the phones from a web browser with AirDroid.

Figure 3.25: The requisites for our control reversal prototype

3.2.10 Questions

- Do users accept that their phones make decisions?
- If not, can we influence them to do so?
- Do areas exist, in which users want decisions to be taken away from them?

3.2.10.1 Insights

We tested this prototype mostly with students at the HPI and in the cafeteria. We realized that the preferences for calling or not calling greatly varied among the testers. A main issue was that testers were not sure about the availability of their conversational partner for a call. Interestingly, the majority of the users accepted to lose control in some situations, especially when they were wasting their time and did not reflect that properly. In other situations, when they do not really need help, the same people would perceive the intrusion of the AI as inappropriate. It is thus very important to design an intelligent system, that is capable of detecting situations very accurately.

3 of the 4 testers told us that they would agree to a call if they knew their conversation partner was available for it. This again supports the finding, that we should be able to analyze the availability of the users for a call before initializing it.

Some users told us, if bad actions are denied, that might lead to learning opportunities, but also to frustration. Denying actions and negative feedback generally does not feel good and will lead to negative connotations with our solution in the heads of the users. We thus should carefully check whether denied actions can be replaced with positive suggestions, that are less rigorous but aim at the same goal of helping the user to behave better.

Another learning was, that users want to be steered in the right direction as early as possible to waste less time on actions that are suboptimal or will be denied. Future solutions might thus preventively encourage good behavior from the very beginning.

A major problem with this solution was that users neither trust their devices nor AI technology in general to function correctly. To function as a proper solution, a system needs to be trusted by the user in order to rely on the suggested actions. This requires a careful and transparent design of the solution to establish trust and reduce barriers.

3.2.10.2 Evaluation

In general, users want to keep control over their devices but they also accept a system, that suggests alternatives or even interrupts them if the reasoning behind the proposed action is comprehensible and transparent. It is also crucial that the proposed action fits the current situation and surroundings of the users. For the given scheduling task, this may be an automatically created Doodle[10], a chat or a call depending on many different parameters. Unfortunately, almost none of the users actually trusted their device to get the action right for their current context. We think, that this might be related to users being trained to distrust devices, by lots of dysfunctional software. This means, that

[10] https://doodle.com – Online Scheduling Tool to find a suitable meeting slot for multiple participants

an AI system would have to rebuild that trust. We may also focus on giving positive suggestion instead of denying user actions, as the latter resulted in frustration.

3.2.11 Message Score Prototype

As we learnt, that forcing an action on a user might not be the best approach, we wanted to try more nuanced versions of making them accept AI decisions. From the feedback we received on the control reversal prototype, we compiled the following options:

- Force decision
- Manipulate user into decision
- Incentivise decision
- Gamify decision making
- Suggest decision

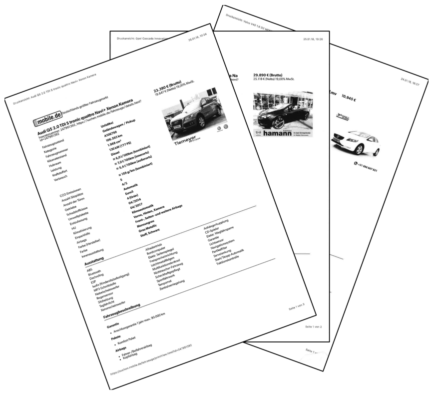
In this prototype, we started working on the aspect of suggestion, by giving the user hints in real time, about whether we considered their communication behavior good. For that we used the topic of rudeness in online gaming:
While playing online games, people are often very rude and offensive to each other. We assume, that there are many similar situations, especially when communicating over the internet with at least partial anonymity, which makes it easier to use bad language and to not respect other participants in the conversation. This bad communication behavior might lead to exclusion of people or even bullying. With our second prototype, we are trying to overcome such problems by using an automatic rating system.

3.2.11.1 Procedure

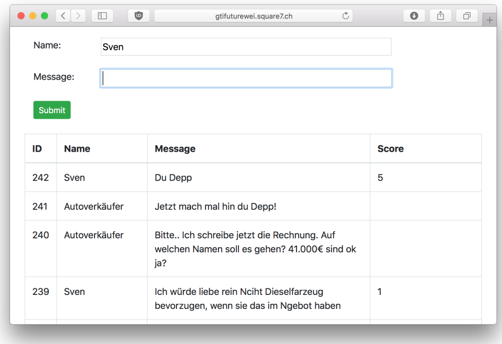
For this prototype we tested a variety of scenarios, always iterating and improving them with the testers' feedback on the prototype itself, because we wanted to provoke a specific kind of communication. All tests were conducted using an online chat app which we developed ourselves. We built a chat app, functioning like a normal chat app, but we scored the messages on another computer using a direct connection to the database and annotated them with the score (from -5 to $+5$) in the app. Each score was only visible to the sender of the message (Figure 3.26b). During all tests we instructed two people to communicate with each other using the chat app and told them, that the scores were generated by an AI using metrics which we varied during the test.

At first, we instructed the testers to negotiate about the price of a car. We gave one tester (the seller) a detailed description of several cars with a minimum price for each car (Figure 3.26a). The other tester played the buyer, whom we gave a budget and some constraints for the new car. We rated all message on their politeness and spelling.

Because testers were somewhat overwhelmed by the many parallel tasks, we switched to a simpler scenario: The two testers only had to negotiate about a fixed amount of money, while one person was told to make a final offer and the other person had to accept the split or not. If they did not



(a) Three of the car offers we gave to testers as an example. They were real offers extracted from mobile.de.



(b) The chat application used in our prototype with annotated scores for the buyer. Here we gave higher scores for impoliteness.

Figure 3.26: The requisites for our message score prototype

agree, they would not receive any money. We also introduced a time limit and let the users play multiple rounds. Additionally, we tried a classical prisoner's dilemma over multiple rounds, where the testers received no (both betraying), one (both cooperating) or two (betraying a cooperating tester) pieces of chocolate, with the same parameters.

Most of the time we did not get the emotionally loaded messages we wanted to achieve, so in our last test we switched back to the car selling scenario, where one of the testers was replaced by one of us. We intentionally wrote offensive messages to provoke the tester into being rude as well.

3.2.12 Questions

With this prototype, we tried to answer the question, whether we can influence the communication behavior of human testers by scoring their messages, when the score is only visible to themselves. Thereby, we found it most interesting to investigate, whether the communication behavior of people can be influenced without any extrinsic motivation, as the score was not coupled to any reward the testers were able to get. As a side effect, we wanted to explore how people feel about being constantly observed and evaluated.

3.2.12.1 Insights

For the message scoring prototype, we iterated several times with slightly changed tasks but the same core concept of our prototype (as described above). If we told the testers about the score at the very beginning of the testing, people recognized maintaining a good score as a primary task for the experiment. Therefore, we hid the score at the beginning and did not introduce it explicitly. This way, testers got their first scores after the first messages, which shifted their focus. As soon as they noted the scores, they started to experiment with the scores and tried to find patterns. If they were not able to understand how the score was created from their messages, they did not worry about getting bad scores and stopped thinking about it. By observing that usage pattern, we understood that the score animates users to find out what it means and only if the understand it, they might change their behaviour to get better scores. Nevertheless, scoring sent messages identified that sometimes this might be too late: Testers, who found the score useful wished they saw it before sending the message in order to edit and improve it.

For the later experiments, where a tester's task was to buy a car, one out of two testers wished to be able to rate the conversational partner with scores. This emphasized that the scores were seen as feedback and with a further inquiry showed that they do not trust the AI very much to give adequate ratings. In these more complex situations (offering or buying a car), the score was less important to tests when compared to the prisoner's dilemma. In most cases (five out of six), the score was interpreted as being related to the task and how the tester performed the given task (even though this was unintended). We learnt how crucial it is to find a good balance between attention to the task and the test scenario. In addition, we were surprised to realize that people like to think that the score is related to their performance in the experiment. Overall people did not care about their conversation being observed or rated in the context of the test. It seems that the content of the messages is not as critical as making photos of testers.

3.2.12.2 Evaluation

In general, we observed that people paid attention to this score regardless of whether we explicitly asked them to maintain a good score or not and even when the score is not visible from the beginning. The human testers showed a high interest in understanding how the score is generated, but if they were unable to reasonably explain the scores, they tended to mistrust the AI. They only try to maintain high scores for the whole experiment, when they thought that a high score would help them achieve another goal like performing well in negotiations. Otherwise they began ignoring the score after a while. We therefore conclude that it is not possible to influence human communication behavior by scoring, unless they can understand how the score is generated and consider it helpful. Another option would be to make these scores publicly visible to give people an incentive to keep a high score. In the future, we want to look into extrinsic motivation or gamification to maintain the score importance over longer time frames.

3.2.13 Psychic AI Prototype

Our first step in exploring the area of self-reflection was building an experience prototype where users would interact with an AI agent that asked them questions that encouraged self-reflection – similar to the role that psychic take with their customers, emulating that relationship.

3.2.13.1 Procedure

We set up an unattended plasma ball with a phone call to Felipe, who performed the AI functionality. The plasma ball called out to passers-by and asked them to have a conversation with it and talk about their worries.

This experience was designed to imitate the relationship people have with psychics or astrologers. People reveal their internal anxieties and stresses to psychics and astrologers and hope for answers that rest their worries.

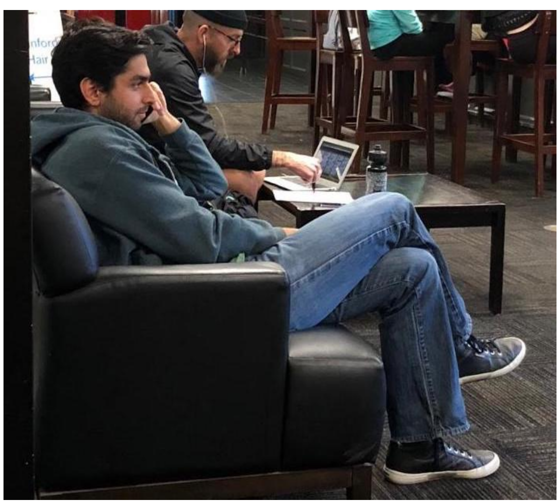
Our vision of a personal therapist that allows for personal reflection, unwinding and giving the required support that one may need on a regular basis.

The experiment (Figure 3.27) itself consisted of two separate parts. The first, the one the user interacts with the orb, which is set on a table with a phone–invisible to the user–underneath it. In addition to this, one of us initiated a conversation with passers-by through a previously set up phone conversation with the hidden phone. We attempted to engage the passers-by in an exchange and get them to share information about themselves–encouraging them to self-reflect.

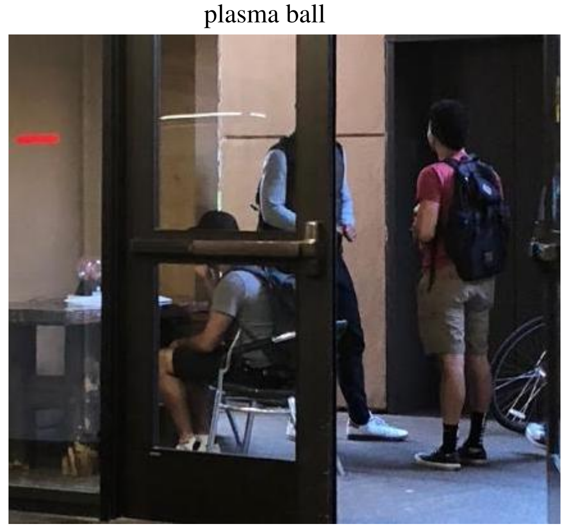
(a) AI psychic experiment setup

(b) Plasma ball with telephonic connection to AI

(c) The phone hidden beneath the plasma ball



(d) Felipe acts as AI presence



(e) One of our customer talking to the ”psychic”

Figure 3.27: Our psychic AI prototype

3.2.13.2 Questions

With this Dark Horse we aimed to find answers to the following questions:

- Can an AI device function as a mode for personal reflection?
- Do people feel comfortable talking to an AI program for self help or reflection?
- Does this concept promote personal therapy?
- Are people concerned about their privacy when talking to AI about personal affairs?
- Do people recognize the AI as a person?

3.2.13.3 Insights

Our prototype testing shed light on assumptions we had made while designing the experience that we had not realized. These made it such that people were reluctant to approach our prototype and engage with it. We found that there was a stigma attached to approaching a psychic. People who were interested in talking to it were the ones in groups. People who were alone decided not to engage even though they were initially intrigued. If users are by themselves they are more vulnerable and so the reluctance to engage increases. When accompanied by a friend, it feels safer so we got engagement from groups.

When we were able to get users to engage with the prototype many of the responses were related to the fascination with the device and with users trying to figure out how it was functioning rather than actually using it for the desired function. This meant that even though we were able to hold a conversation with the user, the goal of helping them with self-reflection was stumped. Therefore we felt that the concept may still work well but the execution needs to be changed to receive the intended user engagement. We realized this since the ”AI” established familiarity over the course of the conversation, the users were more at ease to carry on the interaction.

3.2.13.4 Evaluation

In order to be more successful with the next implementation of this prototype we moved away from explicitly associating the device with the term ”psychic” due to the perceived stigma associated with the term. On top of that we also realized that privacy was a big concern for our users and a reason why they were reluctant to share information, even once an interaction had begun. With that in mind, we would change our prototype to create a more personal testing environment in a private space.

Since most of our interaction with users was limited to very surface-level information exchange and focused on the prototype and how it worked, we would like to allow users to engage with it for a longer time to observe the results beyond initial fascination. To that end, in order to provide a more authentic interaction we would replace ourselves by a trained psychologist to play the AI.

3.2.14 Vent Tent Prototype

Due to the reluctance we observed of users to engage with the first iteration of our prototype, we decided to design a more private interaction. This is how we came up with the Vent Tent — a space where a user can go inside and 'vent', or self-reflect, about their day, or other stressors. This is a private space for reflection set in a public location.

3.2.14.1 Procedure

The prototype was set up to resemble a private space (see Figure 3.28). The tent was set up with a comfortable space to sit with pillows and a blanket and a few stuffed toys. It had signs with instructions (as shown in Figure 3.29) that told users to use the space to vent, de-stress and self reflect.

(a) Vent tent at White Plaza, creating a private environment for people to do self reflection

(b) Interior of the vent tent, with instructions and toys and candies to making an intimate environment

(c) Vent Tent next to an undergraduate dormitory

Figure 3.28: Our vent tent prototype

After having left the tent available to users for a good amount of time, we noticed that people were not open to using it as was desired. We wondered if the observation would be different if the same space was available within a location familiar to the users and if they would perceive it as a safer space if it was within the space they already know to be safe and private, for example, the way students of ME310 perceive the Loft.

To test this hypothesis, we moved The Vent Tent to Slavianskii Dom – an undergraduate dormitory. We allowed the tent to stand there overnight to see if residents of the dorm would warm up to it and approach the space.

Figure 3.29: Instructions for using the vent tent which were display inside of the tent.

3.2.14.2 Questions

With this Dark Horse we aimed to find answers to the following questions:

- Would people be more accepting of a mode of self reflection if it appeared to offer more privacy?
- Would the private space to vent/self reflect be more accepted if it was placed within an environment that was familiar and personal to the user?
- How does the attitude of users towards the service change when the interaction must be initiated by them instead of being approached by the service itself?

3.2.15 Insights

- People are intrigued by the concept and appreciate the idea of it but refrain from actually using it.
- People are constantly concerned about their privacy and security.
- People are less likely to approach a space that lacks any kind of a human touch.

3.2.15.1 Evaluation

The concept of this prototype was a good social experiment that evaluates people's demand of self-reflection. However, this prototype did not have enough users. One reason is that it does not have enough guidance to attract people and lead them to enter the tent. We do have good findings about how people are concerned about privacy and are skeptical about new technology or products that will potentially reduce people's privacy. In order to install a more successful prototype and further explore this design space, we may want to test the prototype as a single device or space belonging to an individual as opposed to a shared device / space. As well as, test a space over a long period of time that exceeds several weeks to see if long term familiarity affects the perception and usage.

Furthermore, we might think of ways to attract people to it and guide them in, which was the biggest issue with this prototype. For that, it might be worth considering to add an explicit introduction to the outside.

3.3 Expanding the Design Space

From our prototypes in the Dark Horse phase and before, we realized that our users and personas were fairly broad and that made it very challenging to find any specific and compelling needs. Therefore, before proceeding in a vague direction without a well defined user to ground our choices, we decided to take a step back and define a user group we would focus on. During the previous prototypes, we noticed to slip off to a single user interacting with a device. Therefore, we intended to look at communication within multiple people, where as family represent a group of many humans of different age groups communicating with each other. We chose the following two closely related user groups:

- Parents with preschool aged children (3 to 5 years old). We thought this would be an interesting group to study because, at this age, children are at a crucial stage of language development and learning how to communicate with their parents and other humans is essential.
- Families with young children (3 to 10 years old) where one parent is travelling a lot for work. This group could be interesting, because young children might struggle disproportionately with a parent being away.

3.3.1 Parents of Preschoolers

Through past prototypes, we found success in the concept relating to visual representation of spoken words. We wanted to revisit that idea while keeping a particular user group in mind and designing for their needs. We chose to pursue the user group: parents of preschoolers. Preschoolers may be defined as children in the age group of 3 to 5 years. By analyzing their needs we decided to design a system that assists the spoken words of the parents with visual representation in the form of animations and pictures to engage the children as well as promote a deeper understanding of the instructions being given to them by the parents.

For a preliminary Needfinding, we conducted several interviews with parents of preschoolers and one with a babysitter of a preschooler. Several common needs emerged. One in particular, that we decided to focus on, was the need for parents to communicate with their children more effectively. They found that children find it difficult to understand instructions and often get distracted or confused and fail to follow through. This leads to parents being frustrated and results in a decline of positive communication.

3.3.2 Interview With a Teacher at a Nursery

Adrienne is a teacher at Bing Nursery with 3 year old twins and a 9 year old son. She works full time at Bing Nursery while raising her three children. She believes that 2nd year of a child's life is an important year in the growth of their language. During this time, they are able to express better, develop a sense of humour and start becoming playmates with each other. Being a teacher who spends a lot of time with children this age, she recognizes that young children aren't good at dealing with change and do not do well under pressure.

While the atmosphere at school is completely different, the home is a place that comes with several obligations in terms of daily tasks. On a daily basis, it might become difficult to get children to go through all required activities include eating, brushing the teeth, taking a bath.

She believes that for children between the ages of 3 and 5 years, it is important for consequences to be tied to their actions closely in order to be effective. She also finds that her children are more likely to stick with a course of action if they have a part in choosing it and do not do as well with things that are forced onto them. This is also her rationale behind letting them guide their own playtime to keep them more engaged.

One of Adrienne's bigger struggles is, getting time to get her chores done. It becomes extremely difficult to get things done while simultaneously having her children engaged in something. She does not believe in eliminating screen time for her children but would prefer limiting it to a certain extent. In some situations, when she needs to keep her children occupied she does take the help of technology even though she does not hope to.

With the complex struggle of balancing her career, raising her children, completing her household duties and all the things that come with them, Adrienne finds herself sleep deprived, extremely emotional and stressed out and sometimes guilty about dealing with situations in less than ideal ways.

3.3.3 Interview With a Working Mother of Two

We interviewed Frances Yang, a working mother with a one year old daughter and a three year old son. She shared her experiences of parenthood with us, shedding light on the differences in raising her first child and raising two children together since the birth of her daughter. Frances works at Stanford University and sends both her children to the Stanford day care from 9am to 3pm daily.

She talked about her struggles with making her children follow instructions and how she deals with these issues. Something that has recently become an issue was his refusal to eat meals and certain types of food. She tries many different approaches to deal with this. She tries using timers to make meals a form of a game to motivate him to complete his meals. On some days she allows him to eat any food of his choice, some days she forces him to eat what she wants him to and on some days she simply gives up and chooses not to care. She emphasized that when the family goes out to dinner, he is not given an option and must eat whatever is provided.

Frances tries to limit the screen time available to her children to 15 minutes but her son always argues for more. She sees a lot of the other children using iPads but wishes to raise her children as individuals who can have a conversation and not be completely dependent on technology for entertainment. She particularly struggles with this problem during outings with other families when the other parents use phones to keep their children occupied and it becomes difficult to exercise control over her children's usage of technology in that situation. In her home, she tries to limit her own personal television time and watches television only during the time that she allows her son to watch.

In terms of content on television, it is a struggle to control some of the things her son witnesses such as scary imagery that might cause nightmares as well as inappropriate advertisements that show up while watching otherwise appropriate family shows. Entertaining her children is especially difficult while travelling.

Frances finds it extremely frustrating to deal with her children's changing moods each day. She finds that 3 years is the age when a child is learning how to understand their own emotions and express themselves, which makes it extremely challenging for a parent. With this daily hassle of getting her children to listen to her instructions, Frances often loses patience to communicate with them and finds herself getting angry and unpleasant after her long day at work.

3.3.4 Interview With a PhD Student

As we found this need of getting children to complete daily tasks without a hassle becoming more common, we spoke to a PhD student Xiao who has two children. She described some of the ways in which she gets her children to brush each day. Xiao has set up a system of using many different toothbrush heads for her son's toothbrush. When he sees a new toothbrush head each day it helps gain his attention and carry out the task with some amount of ease. On days when this does not work, Xiao resorts to letting her son brush wherever he already is to eliminate the hassle of getting him to go the bathroom. This often solves the problem as it is harder to get him to the bathroom than to actually make him brush.

3.3.5 Bing Nursery School Observation

From our Needfinding, we decided to focus on the issue of getting children to transition between different activities. To observe how this might be dealt with in an effective way, we spent some time at Bing Nursery (see Figure 3.30) observing the methods used by the nursery to get the children to transition between tasks. We found that the teachers inform the children in advance about upcoming transitions and give them time to enjoy an activity of their choice for a set amount of time while they prepare themselves for the upcoming transition. This works extremely well as the children are more likely to follow the instructions given to them when they have been given some time to adjust to the idea of it instead of being expected to do something right away. Otherwise, children of this age group will refuse to follow the instructions making it much harder for educators or parents to go through the day in a stressfree way. We learnt that ten to fifteen minutes is ideal for children to prepare for an upcoming transition.

3.3.6 Families with Extreme Parents

In addition to families with young children, we believed to find interesting needs and problems in families with a family member being abroad. In our opinion, this is for the following reasons:

- Family members usually communicate a lot, as they have a close connection to each other and thus will benefit from improvements we make to communication.
- Families engage in emotional conversations, that are especially interesting to us. We explored the direction of emotional communication before, for example with one of our CFPs (Emotional chat CFP).

Figure 3.30: Bing Nursery School Playground

- The field of families and relationships concerns everybody and is thus closely connected to other areas we researched before, like the business area. A commonly occurring situation is one of the parents being a business traveller. We thus were able to reuse many insights we collected during interviews of other user groups and transfer them to the new field.

As the field of families is a very broad one, we further restricted our target user group to focus on the most interesting people, who have rather extreme communication habits. We realized, that extreme situations for parents are especially those in which they have to unite the challenging task of caring for their children and partner with a highly demanding job. We also found the direction of parents, who have to travel a lot for business particularly interesting.

We realized, that we did not know very much about our new target user group from our previous Needfinding and thus decided to conduct additional interviews in February. We identified people with demanding jobs like captains, soldiers or consultants to be member of our target group, but for example also prisoners, people in long distance relationships or lone parents.

Following these ideas, we conducted additional interviews, but also went to places like a playground, where we searched for parents that are not extreme parents themselves, but might have an extreme parent as partner. We thought, it would be also very valuable to look at their perspective.

Among others, during our interviews we aimed at finding answers for these general questions:

- How do extreme parents communicate with their children and partner? What are the technical means? How satisfied are people with those and what are open points or problems?
- How does the communication behavior change when one of the parents is travelling for a long time?
- What misunderstandings occur most often when families communicate?
- What are the most challenging situations in the life of extreme parents, not necessarily related to remote communication?

3.3.7 Interviews with Mothers at the Playground

At different playgrounds in Berlin, we met mothers with their children and asked them if they would like to talk about their family circumstances. We wished to whether clarify they fit to our desired user group of extreme parents – either lone parents or have especially demanding jobs or a partner with a demanding job.

From these interviews, we learned that remote communication happens during a normal day, but also when one of the parents is away for a longer time. We surprisingly did not encounter any complaints about existing communication apps. For one of 4 people privacy was an important concern in her personal communication. This is why she uses the app Signal instead of more popular messenger apps.

Text messages are used in general, and according to our interviewees, they already provide a satisfying level of convenience. Misunderstandings can be settled directly via the text message. After we asked about issues with existing technology, we changed our focus slightly to ask questions about video chats. We did this, because we noticed, that most people do not perceive existing text messengers as problematic. We acknowledged, that there are wishes and issues the people face without telling us when being asked directly. We also searched for other indirect questions about the daily life of the interviewees to tackle these open points.

We also learned, that parents are concerned about the well being of their children, when they cannot be with them. When the parent is at work or travelling, he or she wants to be able to check the status of their kid at home, kindergarten or school. This is currently not possible to the desired extent. Sometimes there is also a need to provide teachers or other people, who take care of the children while the parents are not available, with additional information. This can be for example a number, that should be called in case of an emergency or information about the kid like known health issues or allergies.

We identified, that families under demanding circumstances need to stick to a fixed schedule that has to be agreed upon. Especially when both parents are working, their daily life needs to be well organized and does not allow flexibility. For the lone parent, it is hard to balance her schedule at work with the needs of her child. When the child is sick, this can lead to problems with her employer as she spontaneously needs to stay at home.

3.3.8 Interviews with Business Travellers at the Airport

At the airport, we talked to business travellers and one member of a privacy group. We learned, that travelling for work and family life can be integrated without problems. This however, requires the partner to stay at home and care for the children. All interviewees had an agreement about fixed roles with their partner. In most cases it was the man, who is responsible for earning money and the women stays at home, manages the family and goes shopping. Although we met female business travellers, they had no children of younger age, which supports this finding. If there is a phone call between the family and the travelling person, the time of this call is defined before as well. We concluded from the interviews at the airport, that being an parent with very limited time at home / with the family requires tight organization and does not allow for much flexibility.

The member of the privacy group told us not only about his family life, but also shared some insights his job might provide useful for our project. One of the main points was, that people from Germany care a lot for privacy, but they usually don't have enough technical knowledge to fully understand the implications of all their actions and are careless. They also use voice assistants or other recent technological advances for convenience, but some people also completely refuse to use them solely because of privacy issues. There is a high interest in apps, that care for the privacy of their users like the app Threema. Apps that make their privacy concerns transparent are generally preferred over others.

A general finding that holds true for all of the interviewees was that while they did not consider their work life balance to be problematic and were not complaining, they still would like to spend more time with their families and are disappointed when they miss important events like birthdays.

3.3.9 Phone Interview with a Consultant

We talked via phone to Florian Mann, who is a consultant at McKinsey and also recently became father. He used to travel for work, usually from Monday to Thursday. Since he became a father, he gave up travelling and now permanently stays at the office in Berlin.

From his experience, we learned that prioritization is very important to balance business and private matters. Personal concerns can sometimes be prioritized, even with support of the employer. This is however not possible on a regular basis and normally video chat must be used. The video chat establishes a more intense connection than a phone call from the perspective of the consultant and is thus more appropriate, when he talks to close relatives, friends or his partner. Many situations do, however, require face-to-face communication, which still differs a lot from any form of remote communication.

3.3.10 Interviews with Refugees

During interviews with refugees, we learned that eating together or preparing food and cooking are important parts of a family life and create the feeling of proximity. One refugee told us, that he usually phones his mother in the kitchen, who gives him instructions for cooking. However, this often leads to misunderstandings, because instructions are given at the wrong time or are misunderstood. Other interviewees completely abandoned communication while having anything else to do at the same time. One reason for that is, that they need at least one of their hands to hold the phone and it is therefore harder to do anything else, but it is also hard for them to focus. The remote communication requires more attention than a face-to-face talk and is less convenient. An important difficulty is the limited perception of the surroundings through a video connection. One of the persons sees only a small part of the other person's room or surroundings and often has to guess or ask what the person is doing exactly, if he or she is not just communicating.

Another insight was that people from other countries have a high demand for getting into contact with locals. Our interviewees interestingly used primarily smartphone apps to do this and find other people to meet in the neighborhood.

3.3.11 Interview with a Member of the Army

We interviewed Tobias Luckau, an officer of the German army in the barraks in Golm (which is depicted in Figure 3.31). Officer Luckau himself has no children. He told us, that the communication between the soldiers on a mission in a foreign country and their family at home is ensured by a permanent WiFi supply at the camp of the army. The schedule of the soldiers in addition to this allows daily contact home. Although the soldiers are given the opportunity to regularly stay in contact, problems in relationships occur very frequently during stays in foreign countries and also many break ups. Officer Luckau told us, that he would prefer not to go on a mission in a foreign country, if he had children himself.

Figure 3.31: Barracks of the army in Potsdam, where we met officer Luckau (Photo by Karsten Knuth)

The main communication problem besides missing proximity are the very different circumstances of soldiers and families at home. Everyday issues are not relevant to the soldiers at all, as for everything from cooking to organizing the daily life is already taken care. At the same time, families at home cannot or do not want to understand dangerous situations on a mission or are not even allowed to know details for safety reasons. When being on a mission, the soldiers establish a strong connection to other soldiers and the connection to their families and friends at home gets weaker. The soldiers act as a replacement for the social community at home to some extend.

Communication of soldiers who are parents happens at a fixed times without flexibility. In very rare occasions, like the death of a close relative, visiting home during a mission is possible, this leads however very often to the withdrawal of the soldier from the mission. The communication home is primarily focused on the exchange of information and there are no activities families share over a remote connection like eating or playing a game. Soldiers do consider missions in foreign countries highlights of their career and put emphasis on them.

3.3.12 Interview with a Member of the German Parliament

As a representative of the group of politicians, we had an interview with Alexander Krauß (Figure 3.32), who is a member of the German parliament. He is married and has three children. His family lives in the ore mountains, which are more than 300 kilometers away from the capital city Berlin.

Figure 3.32: We talked to Alexander Krauß, a member of the German parliament. Photo by Sandro Halank (Wikimedia Commons, CC-BY-SA 3.0)

The politician told us about the trouble he faces communicating especially to his younger children. It is hard for them to focus their attention during a phone call and he can most often only ask ”yes or no” questions and talk no longer than 30 to 60 seconds. He calls his family daily at a fixed time, but not during the meals. Alexander Krauß also mentioned several rituals the family normally does together like praying and singing together. These are not carried out remotely and completely fall out when he is not at home.

Mr. Krauß also told us, that he sequentially speaks to the members of his family and never to all at once, because a phone call with a group of people would be hard to coordinate and he will miss some information. This practice however emphasizes his absence and thus makes the communication less efficient for all participants.

3.3.13 Interview with a Partner in a Long Distance Relationship

Klara is one of the ME310 students at HPI during the current program. She is in a long distance relationship across country borders, as her boyfriend is living in Spain.

The couple shares a variety of everyday activities being only remotely connected, for example watching videos online together, reading a book to the other person or cooking together. A common problem is, that they are lacking a sense of space and thus sometimes misunderstand what the other person is doing. They find it particularly hard to synchronize, for example while watching movies together. The most important and recurring problem is that during video calls they are missing important information, because of a wrong focus.

3.3.14 Interview Conclusions and Needs

In general, all our interviewees wish to be able to spend more time with their partners and especially their children. The latter would prevent them from missing important steps in the development of their kids. People think that communication using technical devices is not enough as it is unable to transport closeness and all kinds of emotions. Having a good family life is considered important and thus a sensitive topic. We believe that some people did not tell us about their quarrels for that reason, but also found that while they were really satisfied with existing communication tools like video or phone calls, communication problems occur every day and managing a family over a remote distance is a real challenge. As always, we noticed that people pay attention to how their privacy is handled in their communication, especially within their families. During our interviews, we talked to parents, who are travelling for business or have other highly demanding jobs, but also to persons, who stay at home to validate our findings.

We learned that many communication problems over a distance occur, because family rituals or other common activities cannot be shared anymore. In contrast, it is the case, that almost all our interviewees use remote calls only for the exchange of information and not for sharing parts of their lives and doing activities together. Those who did, however faced some major problems with the existing communication technology. To enable families with extreme parents to continue living their family life over a remote distance and reduce their communication problems, we identified the following needs:

1. Communication should be **hands-free**, because having to hold the phone in your hand during a video or phone call is one of the main reasons, why many people entirely focus on the communication and do not think about doing something else in parallel.

2. The **surroundings** of the persons participating in the call should be better **displayed** to the other end, because that would enable people to follow activities on the other side with less effort and better understand the situation of the remotely connected person.

3. **Communication times** should be as **flexible** as possible to fit well in the demanding every day life of everybody.

4. Conversations should establish **spontaneously**, like when being close together. Little breaks in the daily routine should be used for short chats. The effort to start communicating spontaneously should be as low as possible.

3.3.15 Prototypes

In our family interviews, we learned that communication with a remote family member can be a huge problem. Especially when a parent is travelling and there is a young child (age 3-5) at home, it is hard for them to form a connection. But even if for family members in the same location communication problems arise: Young parents sometimes have extreme difficulties with getting their children to understand and follow their instructions. Therefore we developed prototype ideas to ease the communication:

- An RC car to guide children
- A connected home, which figured out, when participants on both sides had some free time for a call
- Two toys connected via internet
- A robot arm which automatically controls what can be seen during a video call

3.3.16 User testing at Bing Nursery

With the CEP, we wanted to test how children perform in transition of tasks if given a buffer time. We conducted a critical experience prototype test at Bing Nursery school with Sky who is three and a half years old. We brought a toy car and first had a conversation with Sky (see Figure 3.33). We expected him to prepare for the upcoming transition and to be ready to move to another activity (in our example to go to the water fountain), which is not as much fun as his current activity.

3.3.16.1 Approach:

1. We told him that in about 15 minutes, the toy car will be shining lights and will guide him to the water fountain.
2. We kept chatting with him.
3. After 15 minutes, we turned the lights on the car on and moved it to the water fountain.

Seeing that the car was moving, Sky just stood up and followed the car to the designated place without any hesitation.

3.3.16.2 Summary of Lesson Learned

- Children of this age group are interested in play and their attention is easily attracted by lighting effects and a moving car.
- Children consider their toys as friends, therefore it is critical to give the toy personality. After Sky returned to his classroom, he immediately mentioned to his friends that he just made a new friend, which was the toy car. Giving the toy a personality will help children develop friendship with it.
- Children of this age group are in the process of developing sense of time. Giving a buffer time helps the child adapt and prepare for upcoming changes. This mainly allows children to end their current activity and become ready for the transition.

Figure 3.33: CEP testing in Bing Nursery School with Sky

- During task transition, the difficult part is to get the children to stop their current activity and move to the next place.

3.3.17 RC Car (Functional System Prototype)

For our Functional system prototype we took on the problem faced by parents of preschoolers in trying to communicate with their children in order to make them complete essential daily tasks such as brushing, eating etc. Through our Needfinding we found that several parents struggled with this issue on a daily basis and complained of increasingly negative communication with their children in the limited time they spend with them each day.

We decided to tackle the problem by designing a vision for an AI system that gives parents the provision of giving their children a buffer time before having to perform a task and assist them with a guiding toy that leads the children to the space where the parents need them to be.

3.3.17.1 Design Requirements

1. Automated guided toy. For the Functional system prototype, we built a car that follows a line because this is one of the easiest ways to guide a robot car between two locations. For the final product, we need to have a toy that is self-guided without a solid or visible guiding line. One of the benchmark products is an Automated Guided Vehicle (AGV) as depicted in Figure 3.34 in a smart warehouse. It uses either invisible guiding tape or laser target navigation.
2. Safety. Since our product is designed for children of 3 to 5 years old, safety is one of the main concerns. We need to get rid of sharp objects or components that will potentially hurt the users.

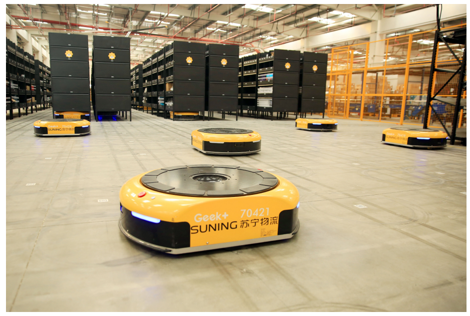


Figure 3.34: AGV in Sunning warehouse

3. Positive engagement between parents and children. This product is served to assist parents to take care of their children, rather than taking over the role of parents. We need to make the children understand the command is from their parents, not from the toy.

4. AI learning improves performance and adapts to the child's preferences and moods. From the project prompt, we are also tasked to utilize Artificial intelligence (AI) in the product. We need to use advanced algorithm in AI to predict the mood and preference of the children users and adapt to better attract and interact with children. The potential input parameters include sound level and child movement. The sound level gives an indication whether the child is screaming or in an uncomfortable situation. The movement detection can tell whether the user is being attracted to the toy and following it. Figure 3.35 summarizes possible in- and outputs for an AI.

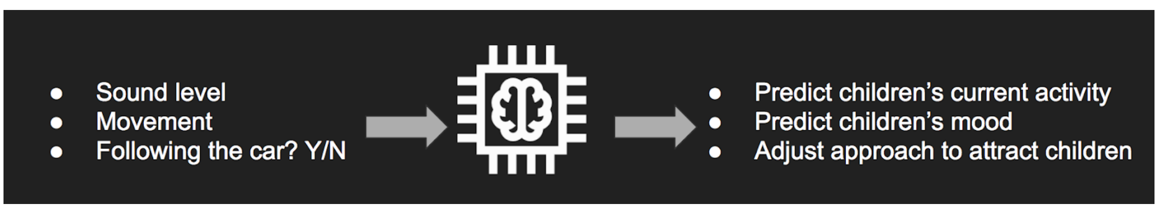


Figure 3.35: Future AI inputs and outputs

5. A refined form factor and interface of the guiding robot will be more appealing to children. Adding lighting and sound effects is identified to be a key feature of the final product.

3.3.17.2 What we Built

We built a system prototype to test out the functions of the app and the guiding toy (see Figure 3.37). The toy in the form of a line following robot car has some additional LEDs. The system intends to first receive a input for a desired time for the activity from the parent and the desired buffer period. It first sends an alarm to the parent to initiate the first communication with the child and simultaneously start the LEDs on the car. Our code is attached as Source Code A.1 in the chapter A. The second alarm initiates the motion of the car guiding the child to the bathroom (illustrated in Figure 3.37).

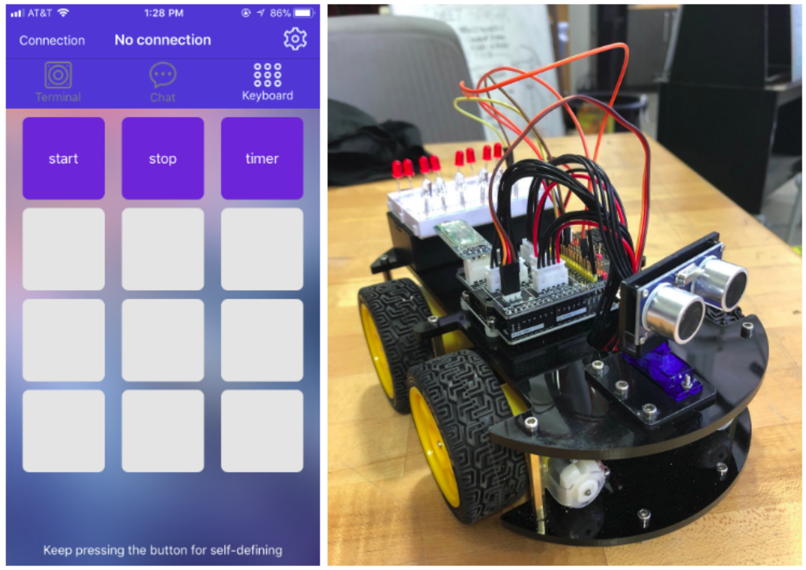


Figure 3.36: Workaround App (left) Arduino Car as a moving object to lead and guide the children (right);

Figure 3.37: Guiding toy in a home environment

3.3.17.3 Results

We were able to successfully build the car along with the assistive lights and a bluetooth module (Figure 3.36). The app we built was able to successfully connect with the bluetooth module, however it failed to communicate the commands. We expect to solve this problem by changing the bluetooth module on the Arduino. Currently, we are using a workaround for this problem by using the existing app (see Figure A.4) and changing the commands according to our design requirement.

Once we achieve this, we will be able to test the system together. The results we received from testing the sub systems individually have been promising and show that our concept is capable of achieving the intended purpose. Introducing the buffer time followed by the guidance within the home works well as our tests showed.

3.3.18 Self - Opening Video Window Prototype

This prototype sets up a short interaction/call, when both people are available or doing things that do not need their full attention. For example, the system could be installed in a kitchen (see Figure 3.38) to use spare time while waiting for the coffee machine to make a coffee.

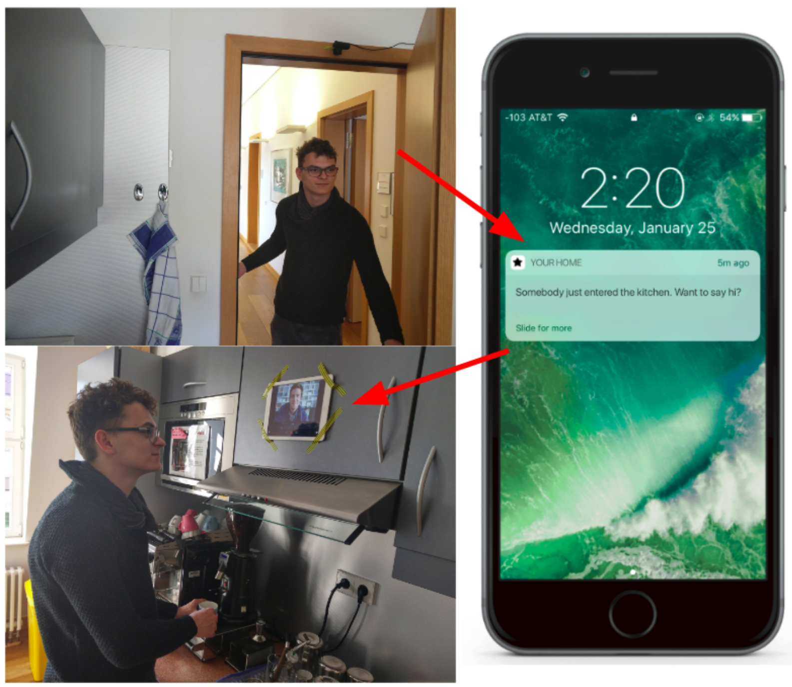


Figure 3.38: A camera notices, that the user enters the kitchen in his home to make coffee. His partner gets notified, that someone entered the kitchen. If both are available, it starts a call.

Most families reported only having regular scheduled calls with the family member abroad, without having any interactions at random times. A specific time to talk reduces the chance to talk about normal things that happen everyday: The communication might seem staged. People with different environments do consider very different things as important and a busy schedule on both sides makes it hard to find a common time to talk. Talking for a longer time is often considered as inefficient and unpleasant for both sides, but they still want to share what happened in their lives. However, many times people are afraid to just call, because they think it might not be an appropriate time for their partner, and therefore short breaks go unused and calls are less frequent.

But in a normal household, family members would meet each other by chance, e.g., when entering the kitchen, or when someone arrives home, and could then talk for some minutes. To recreate this, we want to find those times when both sides are available and then set up a quick call. This may lead to feeling more like being in the same place. For that, we focused on long distance relationships first, since both partners need to organize their daily schedule, which often leads to problems.

During other research, we found that a long distance between the partners does not necessarily harm a relationship [12]. It is however crucial to communicate very well [20] and the psychological study has shown that missing or wrong communication is the reason for break ups [12]. Fixed communication times and missing flexibility influence communication behavior and the success of communication negatively [12].

3.3.18.1 Procedure

The prototype involves a detection when a person enters a room. This can later be expanded to the detection of other events and generalized. The following example shows how the prototype works:

Person A is the person at home, person B is the business traveler. The shared room at home is equipped with a special device, that has a screen and microphone/camera (Figure 3.38). This device can be used for other tasks like maintaining a family calendar/taking notes/searching for recipes on the web etc. Person B has to install a specialized app on his or her smartphone or tablet. Person A enters a shared room (could be for instance the kitchen) and is available. The availability is detected by the special device and its camera. Person B enters a private room (for instance the hotel room), looks on his or her mobile phone and sees that person A is available. When person B starts to interact with the mobile app (virtually ”enters” the shared room), a video and audio connection is automatically established to the shared room at home and both persons can talk. When person A leaves the shared room, he or she is marked as not available.

3.3.18.2 Questions

- Will the frequency of calls increase through the notification?
- How does the content and type of the communication change?
- Do the users feel more like they are in the same location?

3.3.18.3 Benchmarking

Existing solutions include smart video surveillance cameras that could be installed in the exterior or interior of a home. In most cases, these cameras are targeted as a home security device to prevent burglaries or help in enlightening criminal offences. Many devices can be connected with the internet and thus allow remote access to the live feed. Some advanced products, such as the Nest Cam IQ[11], offer an object detection and trigger an alarm, recording or notification, if a human enters the recording area.

In the area of messaging apps, the availability status of a conversational partner is very common. In most cases, these apps (such as WhatsApp, Slack or Skype) differentiate between a user being online or offline. Some allow setting a custom status to transport a more specific type of status

[11] https://nest.com/cameras/nest-cam-iq-indoor/overview/ - Product page of the smart webcam Nest Cam IQ

(e.g. *commuting*) or automatically update their status to show the user's activity (e.g. a song he is listening or that he is currently making a phone call).

Another, similar solution is the video window that is installed in some locations of the Design Factory Global Network. This video window consists of a 55" screen installed at a wall in the kitchen of participating facilities. Above the screen, a movable camera is installed that records video and audio. This setup is always connected to a second video window in another location, where the stream is shown live. The remote party has the ability to remote control the movement of the camera to look around in the opponents kitchen and so select the scene that it would like to see. Local users might disable the recording or playback if not wanted.

Our solution has several advantages compared to other devices. First (and similar to the video window in the Design Factory Global Network), it is always connected and ready to use. In contrast to the permanent video connection of the Design Factory Global Network, we do not want to establish a permanent video connection, which would not be possible, because the travelling person does not have a separate device and has to start the call from his or her smartphone. It is also very likely the case, that people at home would be annoyed, if the video connection would be active for too long. We however do want to automate the process of establishing the connection, if both communication partners are free, as convenient as possible. Still, every user is aware of the location of the device and what's its designated purpose is. If someone does not want to interact with the device, he does not have to or can leave the room. In addition, our solution allows users to do more than just talking and support other actions (such as cooking) in a natural way while the device is active. This enables remote family members to follow various aspects of everyday life in rooms that are equipped with this device as long as his or her time schedule allows that. As we identified the need of the travelling person to follow the activities at home, this device would allow him or her to follow every aspect of the everyday life in that room as far as his or her own time schedule allows.

3.3.18.4 Collected Feedback

We also asked visitors of our booth in Karlsruhe to give feedback for our future prototypes. The idea of this prototype to mount the device at a fixed place in a room was considered a benefit, as the users are not required to hold in in their hands anymore and it has an appealing interpretation, being perceived as a window to another room. One visitor stated, that she would like the device to have more flexibility, for example by making it move through the room. All visitors agreed, that they would not like this form of establishing a connection, because it is too invasive. Both people, the one entering the kitchen, but also the other person, that gets notified would feel a high social pressure to initiate a communication, even if it might not fit very well. Refusing to communicate would be considered an insult on both sides, if a video call is suggested and even weird behavior like avoiding to enter the kitchen might establish.

The visitors of our booth however suggested another idea that follows a similar direction like this prototype and overcomes this weakness. Both people would ideally be able to state for themselves whether they are available for a call or not. Writing a text message in every such case is however not very convenient and the availability for communication should only be visible to the other person if he or she is willing to communicate as well. The idea of a simple device like a touchable lamp for both communication partners came up. Touching the lamp once might mean, that the status should be set to available and touching it a second time resets it to not available. The availability could be displayed in a way that is not invasive, for example by changing the color of the lamp on the other side. We found this idea very interesting, as it solves the need for convenience and flexibility, but also gives both sides the freedom to choose and does not build up any pressure.

In addition, we got in contact with two users of the video window, which is installed in different locations of the Design Factory Global Network, e.g. in Aalto (Finland) and in Porto (Portugal). Markku, a ME310 Program Manager from Aalto University, said they rarely use the video window, if at all. While he mentioned that it is better than Skype, especially with respect to the sound quality, he usually finds the audio connection is muted to suppress unintended sounds, such as ones from plates and dishes being handled in the kitchen, to be remotely streamed. Markku also compared the video window to telepresence robots and found the video window more useful and easier to be remote-controlled. However, from his perspective, the video window is only suitable for casual meetings and not for work meetings and thus the casual meetings in front of the video window are not planned at all. Madalena, a student from the Porto Design Factory we met in Karlsruhe, Germany, confirmed many insights we got from Markku (such as the casual meeting style without appointments) but described that she and her team uses the video windows regularly. From her perspective, it only works between people that already know each other, but for these it's just fun and a possibility to do some smalltalk without extensive technical setup.

3.3.19 Remote Controlled Teddy Bear Prototype

Travellers find it difficult to stay in touch with their young children (age 2-6) while being away from home. The phone is not a good communication tool as only voice is transmitted, video calls are better but still not that interactive. Abroad parents often feel that they have to compete with the television and cartoons. While children like to play with physical toys and dolls, only local family members can participate. This might be solved by having connected toys, with which the parent and the child can play at the same time or even asynchronously (see Figure 3.39). Actions from the toy and voice could be transmitted to the other side to be played back.

3.3.19.1 Procedure

The doll has multiple motors, at least one per joint. In addition, it is connected to the internet, e.g., through Wi-Fi or a paired bluetooth smartphone and paired with a second doll. By moving the doll, the paired doll also moves the corresponding part (each action is mirrored). Each doll might feature a microphone and a speaker to allow recording and playing voice messages. If a game session is not active, each interaction could be recorded and replayed together with the input from the microphone at a later time, to handle problems with time zones or complicated working schedules. We want to build two devices that look very similar: Like a teddy bear, that is not too big (max. 20cm tall), they should look very cute and be easy to cuddle. The parent at work can move one arm up and down, and this is replicated on the kids side. The parent has the option to record a voice message, which is played, when the kid hugs the teddy or moves it around. When the kid plays with the teddy, the parent teddy vibrates softly and thus allows the parent to interact with the child. The doll‘s movement should be classified in different categories to prevent unintended vibrations of the parents teddy and to give the parent an impression of the child‘s mood.

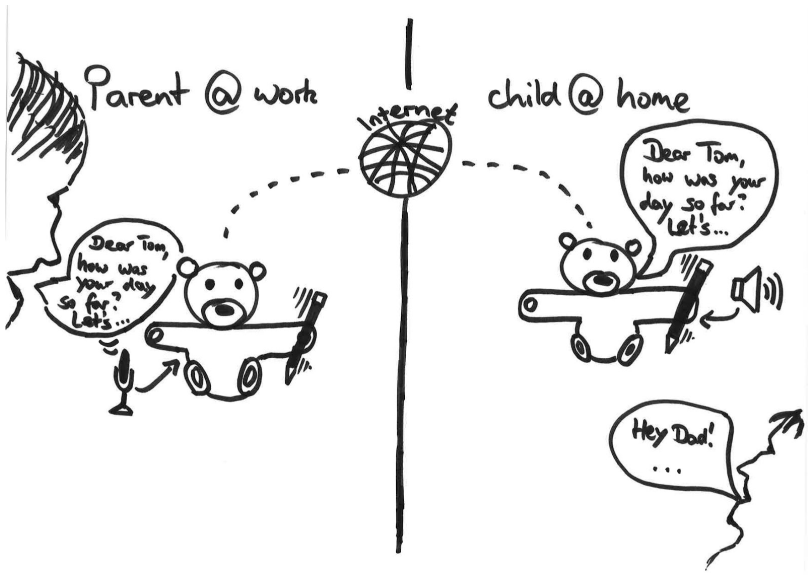


Figure 3.39: Remote teddy bear

3.3.19.2 Questions

- Do children understand, that their parent is responsible for the actions of the doll?
- Is the bond between parent and child increased?
- Is the toy used long term or discarded after a couple of days?

3.3.19.3 Benchmarking

Existing products cover some aspects of this idea but are, in general, not dedicated to allow any physical remote interaction. Available products, such as ”my friend Cayla” [12] are more similar to a voice assistant in form of a doll. This product lacks any form of motors and thus does not visualize an interactivity with moving parts. Due to privacy concerns with the microphone and the online-transmission of recorded speech input, ”my friend Cayla” was abandoned in Germany by the authorities. The last category of products, such as ”my Magical Mermaid” (a swimming doll by Zuru [13])include movable parts but no voice interactivity. All dolls available in the market lack any form of remote connectivity, e.g. for remote working parents.

The idea of our doll is that the toy should be interactive and allow new ways of interaction between a remote working parent and their child. It tries to mimic the local presence of the parent for the play, thus giving both (the parent and the child) a closer feeling of each other. It also allows the child to get in contact with the parent without dialing the number. The parent could put his mirrored toy on their desk (at work, in a hotel room, . . .) to get notified about interactions. The communication is end-to-end encrypted to prevent any abuse and side attacks. The parent might be annoyed by

[12] https://www.myfriendcayla.com/ - Product page of ”My friend Cayla”
[13] http://www.zuru.com - Zuru, the vendor of dolls such as My Magical Mermaid

another app on his or her smartphone, while we think that a teddy creates positive emotions and make the parent remember nice moments.

3.3.19.4 Collected Feedback

We presented this prototype idea among others at our booth in Karlsruhe, Germany. We found a large number of visitors concerned by the safety and privacy of this device. Many people highly value their private information, especially when talking to their children. They told us, that they would want to be sure, that this device will never be hacked. Another concern we heard was, that as with many toys, the kids might get bored fast by this one and not use it anymore. Most of the people agreed that it would be hard to explain the connection between the remote toy and the toy at home to a younger child, even if there is a video call involved in parallel. They feared that the kid might be afraid of the automatically moving toy and not amused.

3.3.20 Robot Arm Prototype

We learned from our interviews, that video calls are often preferred over phone calls, especially in the case of families. The video improves the quality of the communication, but also requires the participants to use their hands for holding the phone properly. It is thus hard for the people at home to do anything else in parallel. Even if they do, people have very little awareness of what is happening on the other side of a video call, especially if doing something together (e.g. cooking or eating), or when multiple people are involved. It is hard to connect to the person in a video, if there is no additional technical support and the remote person will most likely fail to perceive the surroundings at home.

Figure 3.40: Our prototype with a smartphone being mounted to a movable robot arm. The Arm automatically turns the smartphone to reflect where a human would pay attention to.

Our interviews showed, that in most situations, video calls happen at a fixed time without flexibility. During this time, the people at home focus entirely on the call and do nothing else. We also learned, that the business traveller often talks to one person after the other and not to the whole family at the same time. We realized that this behavior emerges as it is hard to focus for on things that are important if the situation gets more complex. Therefore we built a prototype to remove the burden of controlling the phone during a video call, while simultaneously making it easier to speak with multiple people. The prototype is mainly built for families (children 8-10 years old) having dinner together, while one parent is on business trip. During our interviews, we noticed that the age group of 8-10 years is the most interesting to tackle. When children are younger, parents find it more difficult to interact with them over a long distance while older children can be reached individually using their own smartphone.

3.3.20.1 Procedure

The smartphone is mounted on a robot arm, so that it can be moved around and look in multiple directions (Figure 3.40). The arm will be steered by an AI, which decides what to see based on the following five parameters:

- Who is speaking?
- What is the remote user looking at on his screen?
- What are the conversation partners looking at?
- Is someone pointing at something or wants to show something?
- Randomly look around to give a feeling of the location.

For the robot arm, we used a uArm Swift Pro[14] and mounted a smartphone on it (c.f. Figure 3.41). Instead of it being steered automatically, we use a remote control app on another smartphone connected with Bluetooth. We plan to use a freely available Python API for steering and we will also add another camera and microphone as input for the steering algorithm. A 360 degree camera might be useful, to detect off screen movement.

3.3.20.2 Questions

- Can a remote family member become a more equal participant in the conversation?
- Is being moved around by an AI disorienting or does it help with awareness of the room?
- Will the robot arm be perceived as creepy or could it make the remote participant more human by emulating his movements?
- Should the display and camera be fixed together, to show what the remote participant is looking at, or move independently?

[14] http://www.ufactory.cc/en/uarmswift - Website of the uArm Swift

Figure 3.41: The robot arm on which the phone is mounted automatically moves it, so the caller can see the user who is talking.

3.3.20.3 Benchmarking

Existing video chat apps usually allow a direct call between two or more participants. Similar to Skype, they are based on a call, so that one person initiates the call and all others get a form of invitation for this specific video call (Figure 3.42 a). The houseparty app[15], which is designed to connect friends remotely while doing something else (Figure 3.42 b), uses some kind of virtual rooms to allow friends to meet in their spare time without previous agreements. Kubi Telepresence[16] is a tablet stand, that can rotate and change its angle (Figure 3.42 c). It is remotely controlled by the caller. Another approach are remote controlled telepresence robots, which are mounted on wheels and thus allow the caller to actually move around within the space (Figure 3.42 d). An example is the Double 2 telepresence stand[17], which is equipped with an iPad as well. "However, pilot users reported that it was difficult to focus on the social interaction while moving around [...]", a survey of telepresence robot states [14].

Cisco offers the Cisco TelePresence MX Series [18], a business video conferencing system which allows automatic tracking of a speaker walking around the room. Making such a system more affordable and smaller to ready it for consumer use might be a game changer for connecting a travelling family member. The startup Piccolo Labs[19] provides camera technology with gesture recognition for smart home applications. Having a closer look at this might be useful for creating the robot arm control algorithm.

[15] https://houseparty.com - Website of the houseparty app
[16] https://www.revolverobotics.com - Website of Kubi Telepresence
[17] https://www.doublerobotics.com - The double 2 telepresence stand
[18] https://www.cisco.com/c/en/us/products/collaboration-endpoints/telepresence-mx-series/ - Cisco Telepresence
[19] https://www.piccololabs.com/ - Piccolo camera

Figure 3.42: From left to right: a) A video chat between two participants, where both are displayed full-screen. b) The group video chat houseparty, which allows friends to join a chat to allow short conversations from anywhere and connects multiple persons at the same time. c) The Kubi Telepresence stand, that can move around with an attached smartphone or tablet. d) The Double 2 telepresence robot which also allows inter-room movements.

Telepresence robots, however, are very expensive, difficult to setup and to move around. Manually moving the camera on a remote robot is very tedious and can be hard to do, because the caller does not know where everything is or where sound comes from. It is also not possible to steer a camera or robot while, e.g., eating, because you have something else to do with your hands.

Steering the robot manually is annoying over a longer time (has been confirmed by users! Steering is for instance much harder than walking) and with many different interesting things to focus on. Ideally, the AI would focus important things faster than a human would be able to do and could control the robot more precisely.

Focusing on important things is also much more useful for both sides as just showing the whole room to the caller, which might be possible with multiple cameras or a 360 degree camera. This is because looking in a specific direction is much more natural and less overwhelming. Our prototype also enables the people at home to directly see where the caller is looking.

A recent technology that can be used to track human motion and gestures is the piccolo camera (c.f. Figure 3.43). This camera is specifically designed to work under diverse lighting circumstances and will also recognize the position of objects and persons in the room. The detected gestures can be used to control other devices like switching the lights on or off, adjusting the volume of a radio or TV and so on.
Benefits for the caller: The caller is able to perceive the surroundings and the focus will at the same time automatically change to things that are most important at each moment. To summarize, the caller has increased freedom with less dependence of the people at home (e.g. no need to wait or ask them to show something).

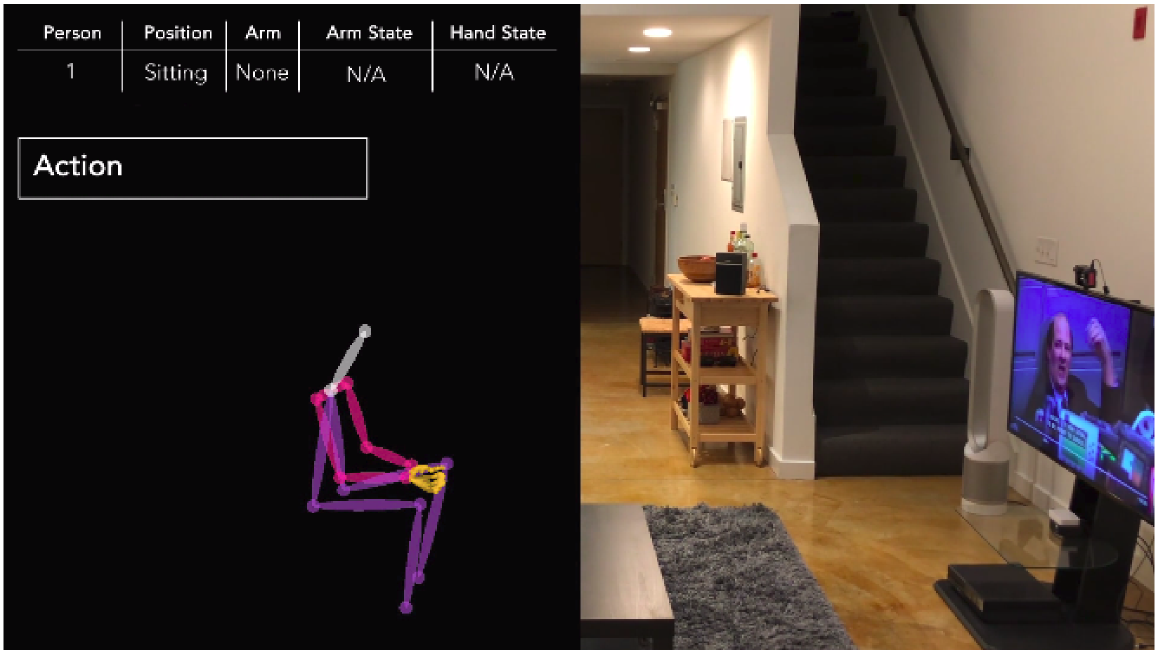


Figure 3.43: The piccolo camera detects a human sitting in front of the TV and analyzes his gestures

Benefits for the people at home: The family at home benefits from hands free communication at all times. They can do other things in parallel, while they are still able to see what the caller currently focused on. It is also not necessary for them to be in the view of the camera all the time. The whole conversation will feel more natural for them, and the caller will integrate as an equal participant into the group at home.
Benefits for both: Both the family at home and the caller benefit from flexible communication times. More flexibility allows longer call and easier chats on different occasions.

3.3.20.4 Collected Feedback

In our limited tests, the robot arm could be steered sufficiently well using the app, although it was sometimes a bit slow. The next steps will be tests with volunteers outside the target group, to optimize the test setup and then start testing with families.
We still have to implement the algorithm that automatically controls the robot. At a later stage, manual control could be added for the remote user as a possible backup besides the automatic algorithm. We will try to identify the needs of the remote connected person with future user testing. Until now, we only included one robot arm for the conversation (for the people at home). In future, we might think of adding a robot arm for each participant of the conversation to bring the advantages to all conversational partners. As we currently focus on business travelling parents, we postpone this to a later step.

At our Winter presentation in Karlsruhe, we already gathered the following feedback: The device should be portable, for use in multiple places at home, and also affordable, since families may not want to spend large amounts of money just for video calling. It might be useful to rent this device for when a family member is travelling for only a limited time, e.g., studying abroad for only a

semester. It was seen as positive, that it uses a smartphone as the central piece, because users would already have that available.
As additional inputs for the algorithm, eye tracking and speech analysis were suggested and for movement of the camera and display there were also lots of alternatives:

- A gimbal or steadicam for cheaper camera movement
- Phone vibration motors, which are already used to take 360 degree pictures
- Prisms or lenses, so the camera does not need to move
- Adaptive Lighting to highlight focus areas of the remote participant
- A drone, to enable even more movement. Possible problems with noise and battery life must be investigated for this idea.

The visitors in our booth also wanted to use this system in long distance relationships, at concerts and for video conferences. Customizability of the algorithm and system was emphasized, so that use cases, we did not think of, could be supported.

3.4 Concept Development

After having explored a huge design space, it was time to finally focus on a single user group and need, so that we could work towards our final product. After that, we needed to develop and iterate this final product. We write about these processes in the following sections.

3.4.1 Convergence

At the end of winter quarter, the two teams at Stanford University and HPI had arrived at prototypes trying to solve the issues related to communication within families. During the week of March 25th, the two teams met in Potsdam, Germany at Hasso Plattner Institute to converge on one combined vision for the final product to be realized in June. We followed a structured process to explore all the possible ideas for a good product vision and work out the details required to proceed with the realization during Spring Quarter.

The process began with reviewing all the various directions and prototypes considered from the beginning of the project until the end of Winter Quarter. It was important to us to discuss all the user groups we had thought of in the past and worked on briefly to discuss if we wanted to revisit any of those directions again and more importantly, to emphasize the lessons we learned from all the prototypes which could be useful for future design decisions.

Through the process of reviewing all the directions we had pursued so far, we were able to identify the key user groups that our past prototypes had explored. We classified the user groups or types of communication as follows: tourists, professional communication (either group meetings or direct communication between individuals), personal communication (either in person or through technology), communication between parents and children (either within the home or for parents who spend a lot of time away from home), and lastly personal thoughts and mental health.

Based on the observations of these various forms of communication, identified several needs that should be addressed and that were not exclusive to any particular user group. Many of the needs were common to more than one form of communication and a few needs were found to be prevalent across all the use cases that we had identified, which made them extremely important to pay attention to. These were the need to have control over the technology as opposed to the technology having control over the user, the need for privacy and the increasing amount of discomfort people feel about losing control over their privacy, and lastly, the need to learn and promote positive communication practices.

After discussing the different user groups we had identified at length, we began the process of narrowing down the selection. The direction of tackling communication problems among tourists was eliminated as we found that there were several hurdles in terms of cultural differences around the world that would have to be addressed. The research as well as the level of artificial intelligence required for this problem would be far more advanced than what was possible for us as a team within the duration of the project in order for it to be successful. The issue of mental health was another such area that required a large amount of expertise in psychology and long term user testing that was beyond the scope of our project. We were drawn towards the two use cases of business communication and communication between parents and children as these were two cases that we had conducted observations and interviews for and had accumulated a fair amount of data that gave us a strong understanding of the needs and ideas for possible solutions.

Upon further discussion, we identified some of the potential problems that could arise in pursuing the direction of business communication. The issues that needed to be addressed were loaded with problems that rely heavily on individual personalities as well as cultural context that would probably be difficult to solve using the technology that we currently have access to. Thus we decided to focus on the communication between parents and children and finalized our user group after having explored all the possible directions.

Having finalized our user group as parents and children, we dived deeper into their needs and the various directions within this use case that could be pursued. We explored ideas for both use cases - parents at home and parents away from home. This time we approached the process of Ideation using a brainstorming session to generate a large number of ideas for potential solutions for the needs of these users. The ideas that came up could be grouped into three main ideas for intentions of the solutions: ways to keep a child safe when alone through engagement, interactive ways to get children to complete tasks and ways to guide children using objects or environmental cues.

We had yet another brainstorming session to come up with product ideas around these needs. Based on these ideas, a final product vision was developed using inspiration from several different ideas. The vision was similar to the prototype developed by the Stanford Team at the end of Winter Quarter and modified to fulfill the needs better and to incorporate the lessons learned from the previous prototype. The vision was formulated as follows: An ambient, home system designed to reduce stress and enable positive communication between parents and their children when tasked with transitioning between locations and activities using lights, sounds and interactive storytelling. The system uses artificial intelligence to determine the best choice of environmental interaction with the child on a given day in order to maximize the possibility of a positive outcome.

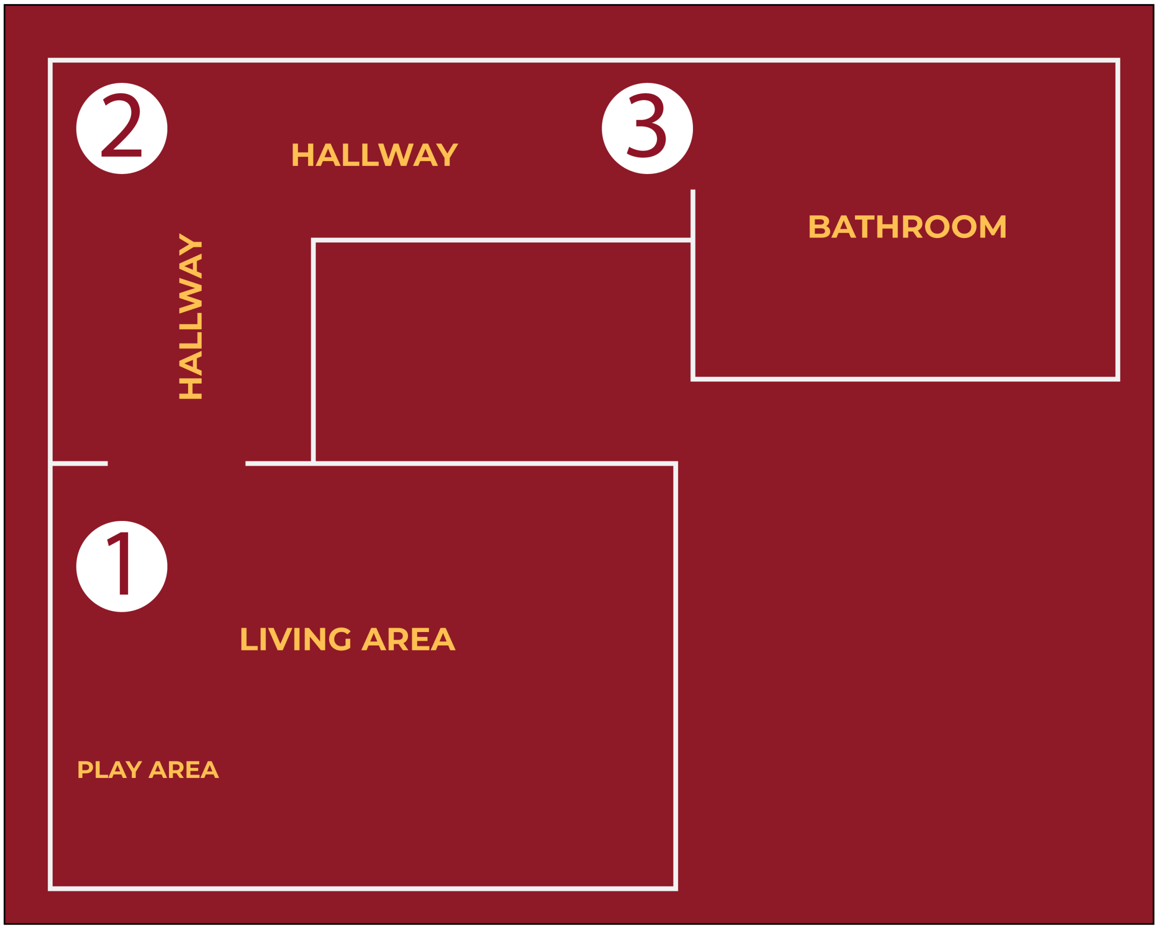


Figure 3.44: System Flow Diagram: Milestones are placed along the way to the bathroom and each plays another part of the same story.

A system flow diagram was created as shown in Figure 3.44. At this stage of the vision, the system would have 'Milestones' using bluetooth beacons placed along the walls of the house. The child would use a toy to move along the house for certain activities and as the toy would come close to a beacon, the child's proximity to it would be revealed thus triggering the activation of lights, sounds or stories in the beacon. This would keep the child engaged while moving in the house during transitions between different activities. In addition, a smartphone application would be available to the parent to input activities that they needed the child to complete, which would give the parent notifications to give instructions to the child in order to provide them with a certain amount of time before an activity to prepare themselves for the upcoming change. This time was named 'Buffer Time'. The same smartphone application would be used to implement the artificial intelligence that would use inputs from the parent about the child's mood and feedback at the end of the system interaction to make intelligent choices and facilitate self learning for improvement of future choices.

The critical components of the system and decisions left to be made about the system were identified as follows:

3.4.1.1 App

In order to control the whole system, we were sure to need a smartphone application. However, we were not sure about whether we should use a native iOS or Android app or build a mobile web app. The app should save paths in the home and distribute different story parts accordingly. These story parts should than be played when the toy is close to the milestone. In order to allow the parent to interrupt the playback of the story, we thought about a wireless communication between the app and the toy but were not sure about the concrete technology (e.g. WiFi or Bluetooth). Once the story is the over, we wanted to collect feedback from the parents within the app about their experience. For that, and for the initial emotion input, we thought about using different Emojis. The advantage we saw was that this would allow parents to enter information with almost no effort. Otherwise, we were concerned that parents would not enter this information. To remind parents about the Buffer Time or to remind them about the feedback, the app should also be able to send a notifications. Last but not least, we thought that the app should contain all AI features associated with our solution. This is due to the processing power modern smartphones have, especially through the integration of the AI co-processors.

3.4.1.2 Milestones

For the Milestones, we thought about a small hardware device that allows the child to listen to parts of a story. Therefore, each should contain a speaker and a Raspberry Pi to control the sound output. In addition, the Raspberry Pi should communicate with both, the smartphone app and the toy. Main purpose of the communication with the toy is to identify when to play the desired part of the story. That should be done automatically as soon as the child (and the toy) reach the Milestone. At that point in time, we were not sure about which technology to use but thought of either Bluetooth Low Energy, Infrared, WiFi or an (optical) motion sensor. We intended to figure out what to use through Benchmarking or during user testing. For the testing, we did not only plan to test with one Milestone and the toy (especially for location sensing), but also to test with multiple to get a better understanding of the whole system. Other hardware requirements of the Milestone we thought of were some LEDs to indicate a status (e.g. for app communication or for the story) and a button to increase the interactivity of stories. The device was intended to be battery-powered to allow free placement within the home.

3.4.1.3 Toy

Prior to deciding a form factor for the toy, we planned to get in touch with a child psychologist in order to understand the possible impact that this decision might have. In addition to that, we thought about testing it ourselves and defining all components that should fit inside. The electronics in the toy should be rechargeable and completely hidden from the child for safety reasons. If we decided to require interactivity for the toy, it should also contain a speaker, some light and a button. All of these should be controlled by a microprocessor, probably in form of another Raspberry Pi. This could then also take care of the required connectivity to the Milestones and fit their technology. Another use case of the toy was to collect data from the child for the AI part of the app, for example to sense the activity level or the mood of the child. This might imply to add special sensors to collect the desired data.

3.4.1.4 AI

An important part of the product is the AI functionality to generate and select the right story. Besides the answers to questions asked parents about their experience with the child, we planned to include the sensor data from the toy and milestone or the history of played stories. In addition, we intended to use the feedback provided by parents after listing to a story with their child, even though not all parents might provide that data reliably.

3.4.1.5 Other

Finally, we thought about integrating other smart devices in the home to our product, especially smart lighting to indicate the Buffer Time. In order to test the product and to show the usage, we know to require different stories. Either we thought about searching for free existing stories, about buying some or about creating and recording our own. As for the entire process, we especially identified the stories as a matter to check with a child psychologist. Besides the user experience itself, we also started to work on the business aspect of the product to get an idea of the packaging, unique selling point and the user-generation of content. This also lead us to the required preparation for the EXPE including the presentation, poster or brochure.
After identifying the critical components and questions that still needed to be addressed, we developed a Spring Hunting Plan to clearly distribute the responsibilities between the two teams in order to work smoothly while at Stanford and Potsdam and have a plan for the final integration in time for EXPE. An official version of the Spring Hunting Plan has been attached as an chapter A.

3.4.2 Final interviews

After focusing on the final vision of the product at the end of March, we still had some open questions to tackle. Especially the HPI team decided to conduct some further interviews to get an own understanding of the situation and to get in touch with families with a 3-5 year old child for the upcoming user testing. During these final interviews, we mainly focused on how parents deal with transitions they want their children to do. We also asked whether their children like stories, which form of stories they prefer and what these stories are about. At the end of each interview, we presented our idea to the parents in order to get some feedback before we actually spent time building the prototype.

Starting with the sense of time, parents told us that their children find a time span of five or ten minutes very abstract and do not have a gut feeling yet. Therefore, one family bought a working analog clock and a non-functional wooden one. If they want their daughter to stop with her current activity soon, they put the desired time on the wooden clock and put the working one next to it. Then, their daughter may compare the big hand on the clocks to see how much time is left. This ritual works very well for that family and other families use similar approaches to prepare their children for an upcoming transition. In general, reoccurring rituals and rules work pretty well with children and it is a bigger issue for parents to interrupt an existing ritual. Nevertheless, reoccurring daily activities, such as brushing the teeth or preparing to leave the home might lead to struggle over and over again. In these situations, parents describe that their argumentation does not work at all and they have to come up with something unusual such as brushing the teeth with music playing in the background or using another bathroom if available. For non-daily activities, such as going to the dentist or buying shoes for the very first time, the most important thing is to prepare children for these events. If parents do so, their children will not do much trouble but only if each step of the upcoming process was included in the preparation. Sometimes, even the order of different activities is important to the child and deviance might lead to problems. In order to prepare children for these events, parents tell their children what to expect either directly or using an existing story.

For all parents we interviewed, bed-time stories were an important part of the bed-time routine. It is a fun part children look forward to and parents use as an instrument to speed up previous activities. Children like to look at and listen to picture books their parents read. Usually, parents allow their children to choose a book they want and for many weeks their kids decide to listen to the same story over and over again. In some occasions, parents select a book, either if the child can't decide or if the parents want to prepare their child for an upcoming event. In these cases, it works best if their kids are familiar with the main character in a book series and experiences different adventures. A common series all parents we interviewed in Germany know (and use) is *Meine Freundin Conni*[20] (engl. *My friend Conni*) that is available as picture book, video and audio book. Parents describe the series as being drawn from life and as a favorite for their children. As with all stories for children, this series does not contain an arc of suspense from an adult perspective.

As we intended to tell stories through speakers, we were especially interested in experiences parents have with audio books. We observed that for audio books, parents allow their children to listen to whatever story they want to listen to without their control. While half of the parents we interviewed bought or rebuilt a Toniebox, others allowed their children to use a streaming service. The Toniebox[21] is an audio player designed for children that plays a story associated with a NFC-enabled physical character as soon as it is put on top of the Toniebox. Popular streaming services include Spotify[22], YouTube Kids[23] or Amazon Music[24], the latter especially through easy use with the Alexa voice assistant. Spotify and Amazon Music offer audio books from publishers and thus parents are sure that no inappropriate content will be played back. For YouTube, the YouTube Kids app only allows access to a limited subset of YouTube content and includes parental control features.

When it comes to digital devices, the parents we interviewed (80% of them had a strong background in a field related to IT) try to limit the time children spent with it. They describe that tablets might produce a firework of emotions which they do not want to happen regularly. Instead, they try to introduce their kids to technology slowly. This also spans to the data they actively share with services about their children and definitely do not want to upload photos to the services. While parents described themselves as very restrictive to the usage, we observed that they are willing to share some personal data if the service they get is helpful for them (e.g. allowing their children to use Alexa for music playback where the voice is sent to Amazon's servers). For our concept, they were willing to provide past experiences and future events as the corresponding countervalue is the preparation of their child. For privacy reasons and simplicity, parents expressed the wish to maintain this data separately and not use another digital calender with existing or more events.

After telling parents more about our concept and the usage of the Milestones, they were curious

[20] Website of the Conni series – https://www.conni.de
[21] Toniebox website and specifications – https://tonies.de/toniebox/
[22] Stream service Spotify – https://www.spotify.com
[23] https://www.youtube.com/yt/kids/
[24] Amazon's Music streaming service – https://music.amazon.com

about it and interested in testing it. They raised three main concerns, which we tried to tackle. First, they were unsure about using a toy as their children use different stuffed animals on a weekly or monthly basis beside one favorite one. This beloved toy might be hard to replace and none of the parents understood what the child likes about their favorite stuffed animal. A second concern was the price of the overall system and the third one the optical aesthetics of the milestones as they shall be permanently set up in the house. While this raised the concern about the look and feel, parents on the other hand liked having the Milestones ready for daily usage without any preparation.

Finally, we also got in touch with a child psychologist to understand more about the impact of our product and the stages of development. When parents have problems getting their children to do what they want to, he always suggests to give clear instructions in simple sentences and repeat these only once. If the kid still does not want to follow the instructions, parents should introduce possible consequences. His suggestion is based on the *Positive Parenting Program* [23], a well-researched program to help parents "confidently manage their children's behavior and prevent problems developing"[25]. In order to prepare children for an upcoming transition, he recommends to visualize the remaining time to children, either explicit using some kind of progress bar or implicit through one remaining song the child is allowed to listen to. Audio also works if parents need their child to go to another room in the house by starting music at the destination so that the child needs to go there to listen to it. If everything works fine several times, the child should be rewarded with a trip, e.g. to the zoo. The child psychologist thought that the concept we presented to him could actually work as it includes small rewards along the way at every Milestone. It is not harmful in any way to children. The only concern he expressed was that stories that are only told through audio might not work as good as a picture book for younger children. However, in his opinion, the device should not include a screen at all (which confirms the impression we got through our interviews) so that it relies in the parent's decision to accompany the children throughout the way and to draw the children's attention to the story.

3.4.3 User Persona: Denise and Taylor

Denise is a single mother having a packed schedule starting in the early morning, getting ready for work and getting Taylor, her child ready for school (Figure 3.45). She works a full day and picks up Taylor on her way back home, where she spends an hour preparing the family's meal. Once dinner is over, she has to clean up the kitchen. Taylor plays with his toys while Denise is busy in the kitchen.

Figure 3.45: Denise and Taylor

When it is time to get Taylor ready for bed, Denise tells him to go brush his teeth. Taylor is reluctant to stop playing. He does not want to brush his teeth and ignores his mother. Denise is tired from her packed routine and exhausting day. A defiant 4-year-old is the last thing she wants at

[25] Cite taken from the Triple P website – https://www.triplep.net/glo-en/home/

this time of night. Taylor continues to play and ignores his mother. Finally, Denise loses her temper and yells at Taylor. The two are now fighting. What started off with a simple instruction to go brush teeth has become an event -– an extra bucket of stress poured on Denise's head. Now that Taylor is agitated, going to bed will be even more difficult. Denise realizes the bucket is starting to look more like a hose and the constant stream of stress piles on. The already long day feels endless.

3.4.4 Final concept

The decisions we made during the convergence phase pointed in the direction of a home guiding system. Additional interviews helped us to become more specific about the development of a final system and with these findings we decided to implement VAMO (from the Portuguese word vamos or, colloquially, vamo, meaning let's go), an intelligent system that is installed in the house.

We found that children between 3 and 5 years of age are at a crucial stage in terms of development of emotions and language and therefore do not deal well with sudden changes and need time to prepare themselves for upcoming activities. We call this 'Buffer Time'. Our final system addresses the needs of Taylor to be prepared for an upcoming transition to another activity in the house. VAMO connects to smart lights in the proximity of the kid and changes it slightly to display the progressing Buffer Time in an intuitive way, that is easily understandable also by younger kids.

VAMO also addresses the need of our persona Taylor to be engaged during a transition to an activity. VAMO includes the interactive Milestone devices, which are mounted along the wall in the house. They are activated according to the desired movement in the house. When the buffer time is over, they will be activated and attract the kids attention. The kid will then move through the house interacting with the Milestones and listening to a story until it arrives at the target.

As the process of transitioning between activities becomes easier and less stressful, the need of Denise to spend more quality time with her child and have a less exhausting time at home is solved.

The hardware components of VAMO are interactive devices, that are mounted on the wall. We call those devices Milestones. VAMO also includes a smart phone app, that is used to control the whole system and intelligently generate stories. While smart lights are not a part of VAMO, they are still a requirement for the full experience. VAMO's app connects to smart lights and controls them to support the home guiding system accordingly. The following sections will give more detail on how our final system VAMO solves the needs we identified being most important.

3.4.4.1 Buffer time

We implemented the Buffer Time concept by connecting to smart lights in the house. The parent controls when transitions should happen by using VAMO's app interface. It is possible to start transitions to activities like brushing the teeth or going to bed immediately, if this is necessary. If there is a regular schedule for these transitions like a bed routine that happens every evening at the same time, parents can also enter this recurring information to the app. The length of the Buffer Time itself can also be specified by the parent and thus adapts to the needs of the kid or special situations.

Figure 3.46: Smart Light Color at the Begin of the Buffer Time

Figure 3.47: Smart Light Color at the End of the Buffer Time

The app will then notify the parent before the planned Buffer Time for the kid starts. An option is given to delay the transitions for situations when the daily routine changes. Otherwise, the parent should inform the kid about the upcoming transition at that moment to give them enough time to prepare. VAMO connects at the same time to the smart lights that are located near the kid and starts to change their color gradually. The color changes from white to a light red as shown in Figures 3.46 and 3.47.

As soon as the Buffer Time is over, the parent gets another notification on the smart phone. The parent should now go to the room where the kid currently is and take it to the next activity. This beginning of this transition is indicated to the child by blinking smart light, shortly before the parent enters the room. The kid and the parent are now guided throughout the whole transition by VAMO and the interactive Milestone devices that will be explained next.

Figure 3.48 illustrates the buffer time concept and how VAMO solves it as a whole.

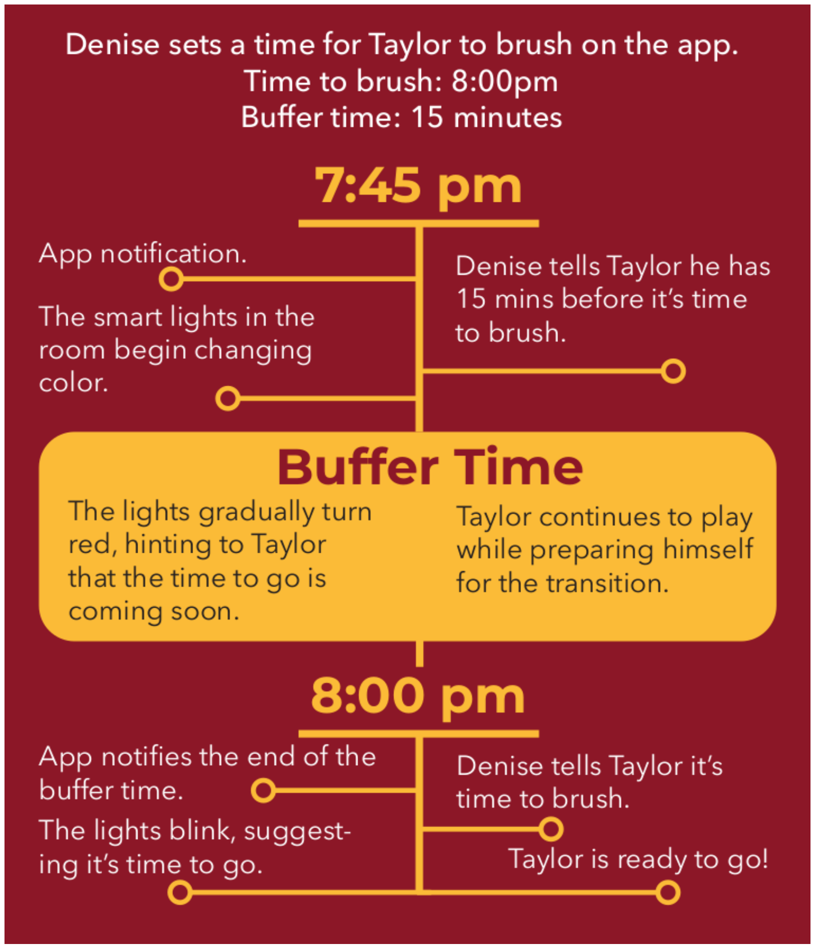


Figure 3.48: The Buffer Time Concept

3.4.4.2 Milestones

The Milestones are interactive devices and the central component of VAMO. Their main interactive part is a button, that can be pressed by the kid to start or continue a story that is played via audio on attached speakers. The Milestones are connected to each other and the smart phone app. They are designed to play the interactive story on the way through the house, that our personas are going to take.

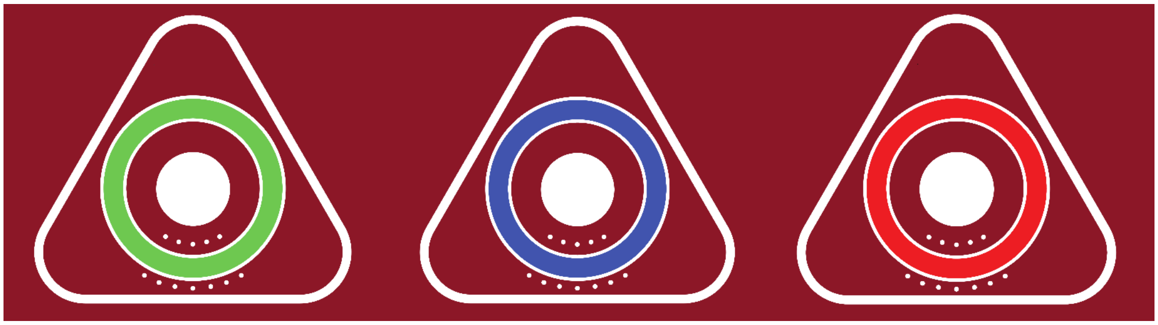


Figure 3.49: Milestones that show different colors indicating the three different states

Each Milestone can be in different states, that are indicated by the light of an LED ring. The LED ring is mounted on the top of each milestone. The three different states, which are illustrated in Figure 3.49, are the following:

- If the Milestone has a GREEN blinking light, it is active. Active Milestones invite the child to press the button to get a part of the story started.

- When the button is pressed on one Milestone and the story begins, a BLUE light appears instead of the green light, indicating that the Milestone is currently playing a part of the story.

- While a part of the story is in progress on one Milestone, all the Milestones that come after it, wait in a stage of inactivity by shining a RED light. If the kid presses the button on one of these Milestones, nothing will happen.

As soon as the provided Buffer Time is over and the start of the story has been signaled to the kid by the smart light and to the parent by the app, the first Milestone will become active and light up green. Once the kid pressed the button on the first Milestone and completed listening to the first part of the story, the second Milestone will become active. The blinking green light attracts the attention of the kid and it will move through the house to the next Milestone and press the button. The story then continues, until part two is over and the third milestone becomes active and so on.

The VAMO system is designed to be used with a variable number of Milestones to adapt to different home environments, the number and size of rooms or the parents and kids preferences. We implemented our final prototype that resembles VAMO in its first revision with three Milestones.

The Milestones along the way solve the need of the kid to keep engaged and reduce distraction by other stimuli along the way. The story is designed to attract the kids attention and keep it excited about how the story might continue.

The use of the button to indicate the transition between the Milestones replaces the need for a toy. The final version therefore does not use the previously described concept and instead simplifies the system to the collection of interactive Milestones.

3.4.4.3 AI & Storytelling

Because the needs solved by VAMO are long term needs, that occur every day over a long period of time, VAMO also needs to be designed for the recurrence of these needs. The system is designed to be used by the same kid over a time interval of 3 years, as our target group are children of the age 3 to 5. Especially the need of our persona, Taylor, to keep engaged has to be considered on a long time span. If the kid becomes bored of the system after a few days or even weeks, it would not be valuable to install and mount the system in the house. Most importantly, our solution intends to help parents and children have a less stressful time over a longer time span and thus improve the amount quality time kids and parents spend with each other.

We identified the story being played by VAMO as the integral part, that determines whether the kid will enjoy using the system over a long time span or not. This is because most of the other components do not change very much and nearly always behave the same. The variety of the experience must thus come from different stories.

The idea behind VAMO is making stories attractive by personalizing them to the kid. If the story is about something that is already on the kid's mind, it will most likely attentively follow the story.

The information we use for personalizing of the stories is entered by the parent into the smart phone app of VAMO. An intelligent algorithm powered by AI technology then aggregates this information and uses it to generate personalized stories, that will be played on the Milestones. Figure 3.50 summarizes the information that flows into the AI system on the left side. The system supports three main types of data, that can be entered:

- **Personality of the kid:** The parent has the ability to enter the character traits of the kid to the smart phone app. The information will be saved permanently and needs to be entered only once. The personality of the kid plays an important role regarding which stories they will most like or dislike.
- **Recent experiences:** Experiences that the kid just made will very likely be on their mind. They should therefore be an integral part of the generated stories. Whether it is a visit at the zoo, a day playing with a dog at the park or something completely different, the kid will be eager to learn more about the things they experienced recently. Parents will enter these experiences to the VAMO app once the kid made them. The app records the time, when the experience was added and older experiences will decrease in importance.
- **Upcoming events:** Important events in the future like the birth of a sibling, going to the dentist for the first time or even visiting the grandparents play an important role in the kid's life. While the kids cannot know about these event themselves, parents will tell them about these events in advance to prepare them. Thus, upcoming events will be on the kids mind, also because of the fact that kids understand that these things are important to them. VAMO helps parents to prepare their kids for upcoming events by integrating this information into the generated stories. Again, this also helps to keep the kid engaged over a long time. Parents are able to keep track of all important events with a calendar function within the VAMO app.

The AI system of VAMO as depicted in figure 3.50 combines this information and aims to generate stories. As we realized, that it is technically not yet feasible to generate full stories with AI tech-

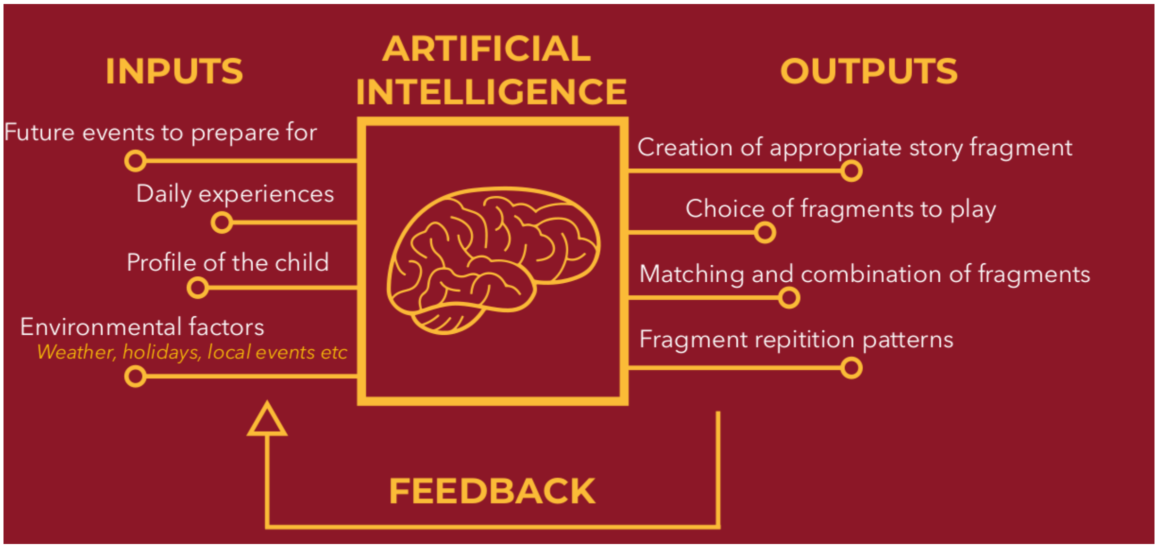


Figure 3.50: AI system of VAMO. All input options besides the environmental factors have been implemented. VAMO's AI currently chooses fragments and combines them without generating a transition.

nology, we use a simpler approach. It is, however, the long term vision for VAMO to generate full stories from scratch, as it allows the highest flexibility to adapt to the data that was entered to the app and also a maximum of variability to keep the kid engaged.

The simpler approach, takes a range of story fragments and generates a story by combining these fragments. VAMO right now combines two fragments at one time to obtain one story. There are two types of fragments: a larger main story and a shorter intermediate story. The main story contains a gap, where a short intermediate fragment can be added. Every generated story consists of one main fragment and one intermediate fragment. The main fragment will provide an introduction, followed by the complete intermediate fragment. After that, the main fragment continues to the end (See figure 3.51). All fragments of both types are annotated by tags, that are used to ensure that two matched fragments fit together. By using this technique, we ensure that the stories generated by our system have a common theme and make sense overall.

The tags have another important purpose, which is selecting what fragments should be part of the story played. The AI system assigns a score to each of the fragments by comparing it's tags to the information the parent provided in the app. The more matches the algorithm can find, the higher the score of the fragment gets. The two fragments with the highest additive score will be combined to form the story that is finally played. A vision of VAMO's AI system that can be implemented sooner in the future will be to have a larger number of fragments combined to one story and automatically generate short transition sentences between the fragments to make them fit better.

Instead of simply matching the tags of the story fragments to the input data from the app, a statistical model will be used in the future to make reasonable predictions of story fragments depending on past experiences with the system from the same parent and kid and also other families. Af-

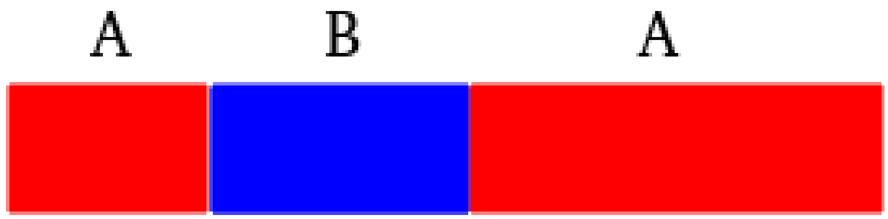


Figure 3.51: Main stories (A) are combined with intermediate stories (B) in the current implementation of VAMO's AI.

ter VAMO has finished playing a story, the parent will be able to rate the whole experience and give feedback, whether the kid liked the story or not. This feedback will be used to improve the intelligent algorithm.

3.4.5 User Testing

In addition to the testing of earlier prototypes we already did during our Needfinding, we also conducted more tests to validate our product.

3.4.5.1 Initial testing in Potsdam

We tested our final system three times in Potsdam, but used only one Milestone (depicted in Figure 3.52), as we were not able to fabricate three in a short time. To simulate the functionality of the whole system, once a part of the story was finished, one of us carried the milestone to the next location and the kid followed. While this does not resemble the exact functionality of the final system, it contains all important steps, especially moving from one point to the next, while approaching the Milestone. We went to the HCI lab at the HPI campus and were able to print one Milestone on an ultimaker 3 3D printer. We tested with two different kids, for one kid we repeated the test after several days and another story.

Figure 3.52: The milestone printed in low quality, that we used for user testing in Potsdam

Both kids were highly attracted by the milestone device as we hoped. We realized, that they were most attracted by the visual cue provided by the pixel ring. We thus implemented a light flushing animation in the final system only for the active Milestone waiting to be pressed and weaker light for the inactive milestones to draw the attention of the child to the correct Milestone.

We also observed, that the kids interacted in a very intuitive and natural way with the Milestone, although we did not give many specific instructions in advance. The idea, that the button has to be pressed to continue the story, was clearly understood by the children.

Both children were initially shy. One kid repeatedly refused to press the button herself, but was still attracted by the system and understood how it was supposed to work. The kid and her mother, who was in the room the whole time, pressed the button together. The other kid was less shy, but still showed initial hesitation. We talked to the parents to analyze this behavior and they told us both that in a known environment like the home, the children feel much more comfortable interacting with new things. We therefore affiliate the hesitation of our testers to the testing environment, which was our ME310 lab, and not the system itself.

3.4.5.2 Testing at the EXPE

During the EXPE we were able to test the full setup with three Milestones, smart lighting and app integration. One kid, that was present at EXPE, interacted with our system and was able to perform the correct actions without any instruction. This means, the kid pressed the buttons on the Milestones in the right order and transitioned accordingly while listening to the story. The child said they had fun going through it and later in the day, when talking to their mother, said that they wanted to go through the experience again.

3.4.6 Hardware Design Iterations

The hardware changed significantly during the design process. In the following, we describe those changes and the reasoning behind them.

3.4.6.1 Electronic Components

The key components we have in the Milestone includes a computing unit, a light unit, a power unit, a push button and an audio unit. The key criteria in selecting components is to minimize the size of each.

For the computing unit, we started with an ESP8266 Module. It is a compact development board with WiFi module and micro-processor programmable in C. We used it to control the LED ring lights and the speaker unit successfully. However, we found Python is a more suitable language for our project which involves mass computation and complex communication between devices. We then moved to use a Raspberry Pi Zero W (see Figure 3.53) which is a micro computer with built-in WiFi and is programmable in Python. It is also compact and have a decent computing power.

Lighting is a key function of our product, therefore a lighting unit that is capable of showing beautiful light effects is needed. We chose to use a NeoPixel Ring as shown in Figure 3.54. 24 ultra bright smart RGB LED NeoPixels are arranged in a circle with 2.6" (66mm) outer diameter and can be programmed to show diverse lighting effects.

The Raspberry Pi Zero W works with a 5V power supply. Therefore, a power source of more than 5V and a voltage regulator are needed. We considered a 4-AA battery pack, a 4-AAA battery pack

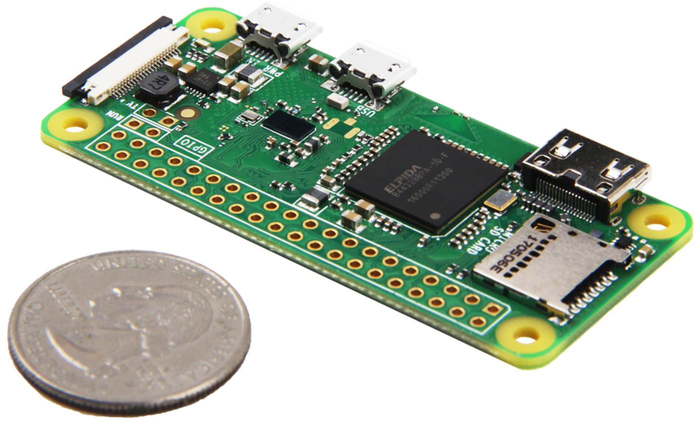


Figure 3.53: Raspberry Pi Zero W

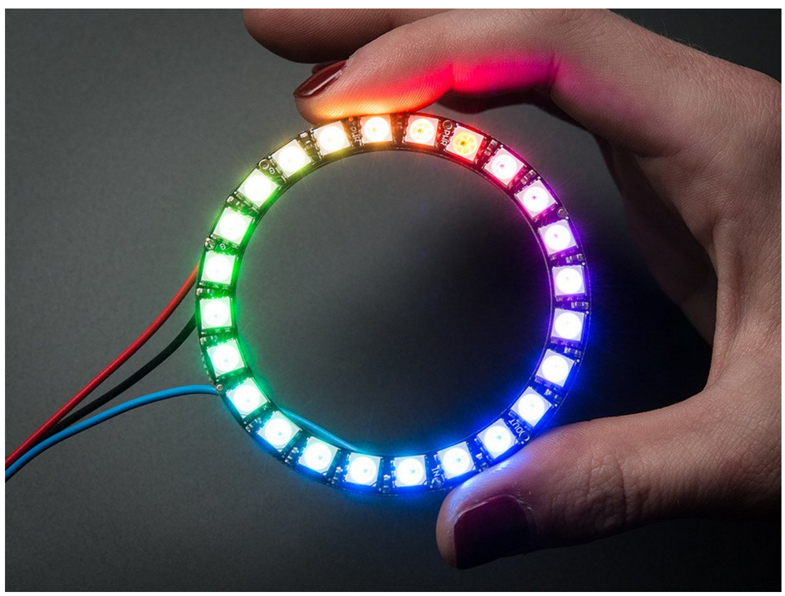


Figure 3.54: NeoPixel 24 LED Ring

and 9V standard battery. A 4-AA battery pack occupied a significant amount of space. A 9V standard battery only has an average capacity of 300mAh. Finally, a 4-AAA battery pack was selected.

We used 4 Alkaline AAA batteries for initial testing. However, we found that the power drained very rapidly and the batteries heated up significantly. Through our research, we learned that our device draws a current of around 600mAh. Given the properties of Alkaline battery, as current goes up, the internal resistance increases significantly, the battery itself consumes a huge amount of energy. We then switched to using Lithium battery, which is a good candidate for large current device.

A voltage regulator is needed to adjust input power to a constant 5V supply for the static performance of a Raspberry Pi Zero W. We purchased 4 types of regulators from Pololu, an electronics manufacturer and online retailer serving education, maker, and professional engineering industries.

Model	Min Input Volt	Max Input Volt	Output Volt	Max Current	Dimensions(in)
S7V8F3	2.7V	11.8V	3.3V	1000mA	0.45×0.65×0.1
S7V8F5	2.7V	11.8V	5V	1000mA	0.45×0.65×0.1
S9V11F5	2V	16V	5V	1500mA	0.3×0.45×0.15
S18V20F5	2.9V	32V	5V	2000mA	0.825×1.7×0.38

Table 3.1: Comparison of four Pololu Voltage Regulators

A full 4-AAA battery pack gives a voltage of around 6.4V and all these three regulators work in our application. S9V11F5 can supply a max current of 1.5A and our device draws about 600mA. S9V11F5 was selected because it supplies 5V output and gives a safety factor of 2.5 for current draw.

We wanted to have a large push button that can give the user an interactive experience. We found a 30mm diameter push button with LED that is perfect for our device. It is shown in Figure 3.55.

Figure 3.55: Push Button with LED

For the audio unit, we started from Adafruit Audio FX Mini Sound Board and a Adafruit Mono 2.5W Class D Audio Amplifier. These components were used mainly because we were initially using C for the software development. We switched to use a speaker pHAT from Pimoroni (see Figure 3.56) as we changed to the Raspberry Pi. The speaker pHAT is an extension board designed for Pi family with built-in sound board and 3W amplifier.

Figure 3.56: Speaker pHAT overlays on the Pi Zero

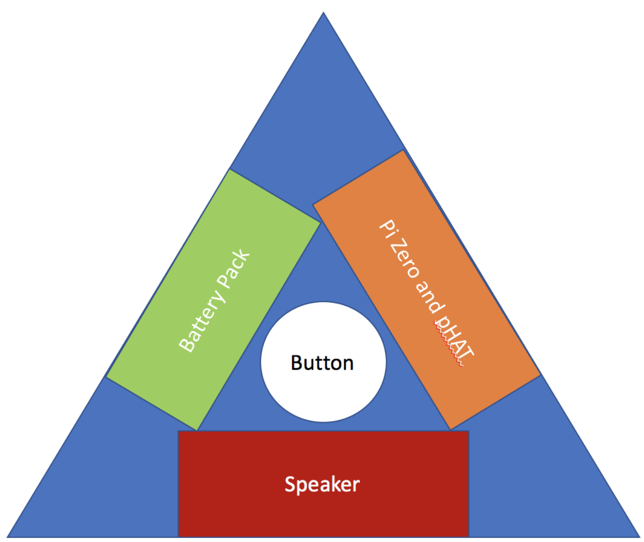


Figure 3.57: Part Arrangement in a Milestone

3.4.6.2 Device Housing

With all the electronic components in hand, we started to design a housing that is user-friendly and designed for manufacturing. We found that the most efficient arrangement was an equilateral triangle. With the push button and LED ring in the center of the triangle, each of the speaker, battery pack and pi take an edge.

The device housing consists of four pieces: main housing, LED ring cover, battery chamber door and back lid. Each of them went through several iterations of design and modifications. In this section, the development process for each of the component is listed here.

Main Housing
Starting from the general concept of how to arrange all components (Figure 3.57), we created a basic triangle shell with a hole in the center for the push button. As shown in Figure 3.58, a circular cavity is created for the LED ring.

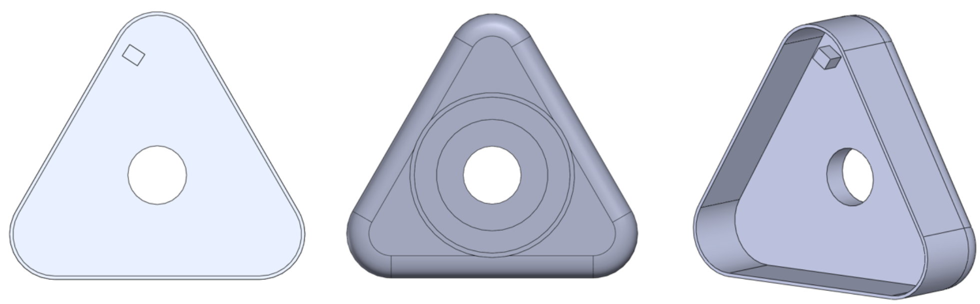


Figure 3.58: Initial housing design on May 07, 2018

After several iterations of modifications to the model, the mid-phase design is shown in Figure 3.59. In this design, many features have been added, including battery chamber, slot for LED ring wires, supporting structures for Raspberry Pi and back lid. A series of small holes are designed in front of the speaker for better sound quality.

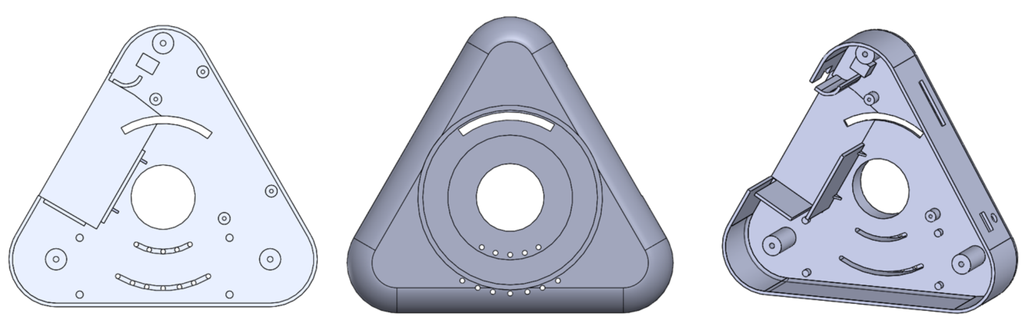


Figure 3.59: Updated housing design on May 11, 2018

The final CAD design for main housing is presented in Figure 3.60. Compared with the previous iteration, some reinforcement supporting structures are added to prevent unexpected structure failures. The cavity for speaker holes are optimized and an opening for mini-HDMI cable is created as well.

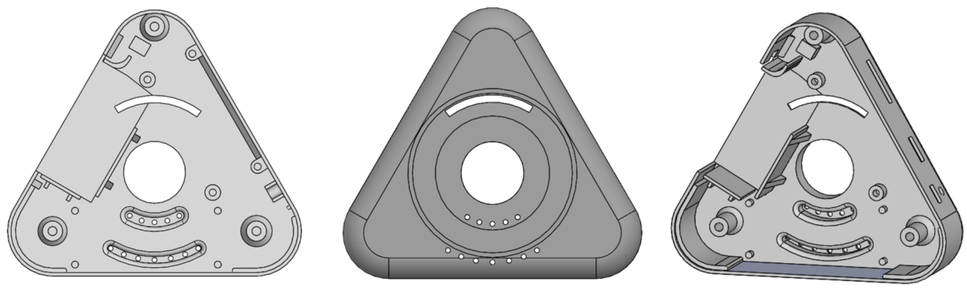


Figure 3.60: Final housing design on Jun 04, 2018

LED Ring Cover
The initial design of the ring cover is shown in Figure 3.61.

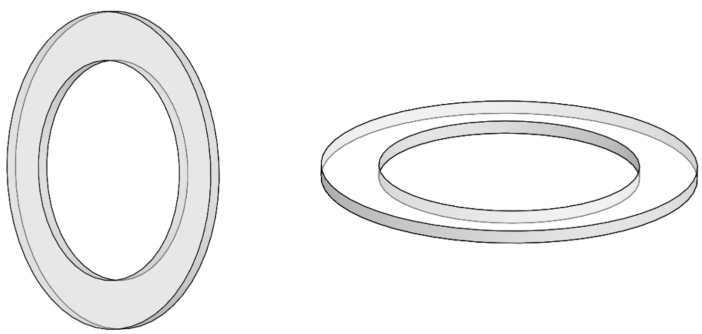


Figure 3.61: Initial LED ring cover design on May 11, 2018

The new design has arched surface (Figure 3.62) for better light diffusion.

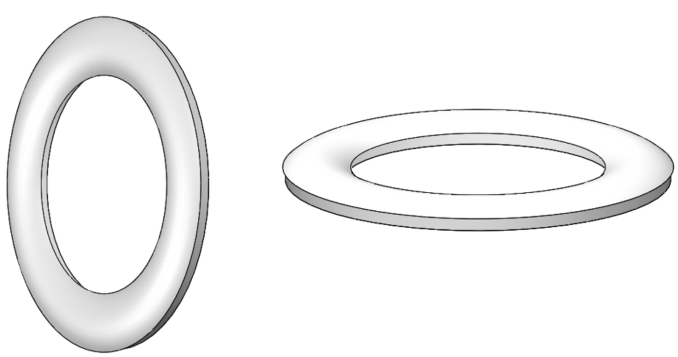


Figure 3.62: Final LED ring cover design on Jun 04, 2018

Battery Door

The first version of the battery chamber door is shown in Figure 3.63.

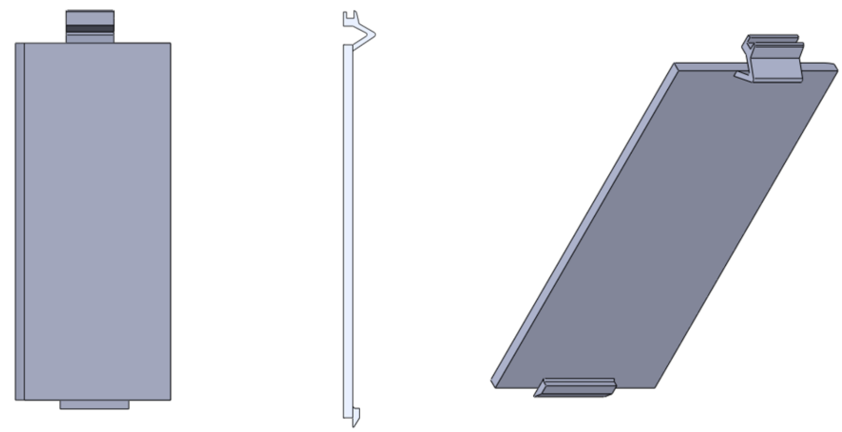


Figure 3.63: Initial battery chamber door design on May 11, 2018

The final version has a wider tooth that prevents the door from sliding in the slot. (Figure 3.64)

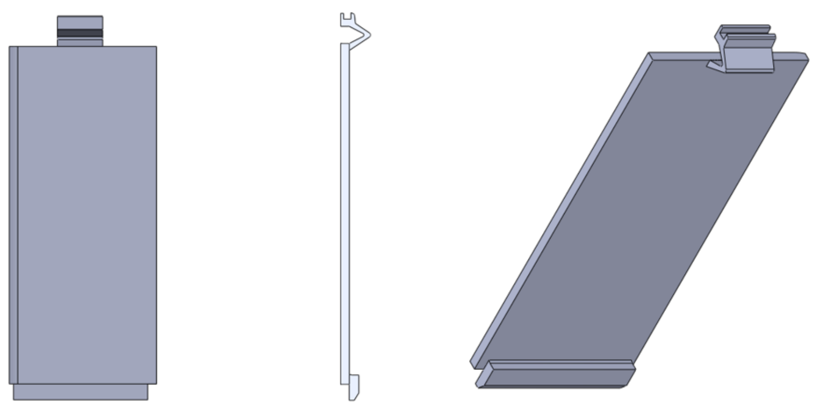


Figure 3.64: Final battery chamber door design on Jun 04, 2018

Back Lid
The initial design of the back lid is in Figure 3.65

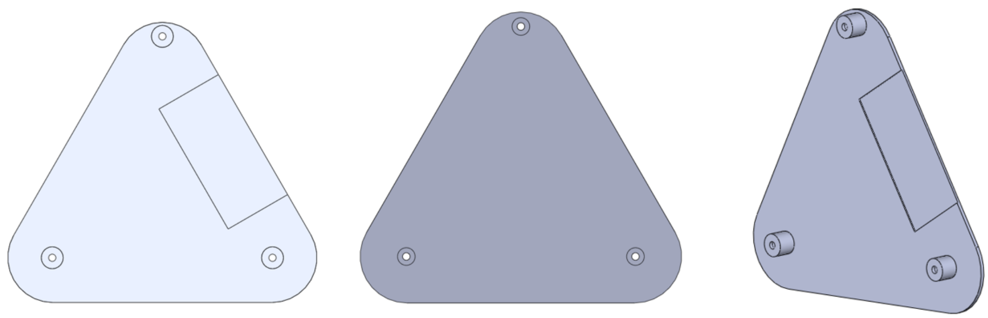


Figure 3.65: Initial back lid design on May 11, 2018

The updated final version has some reinforcement features are added. Since the back lid only has a thickness of 1.5mm, these features significantly enhance the robustness of the structure. (Figure 3.66)

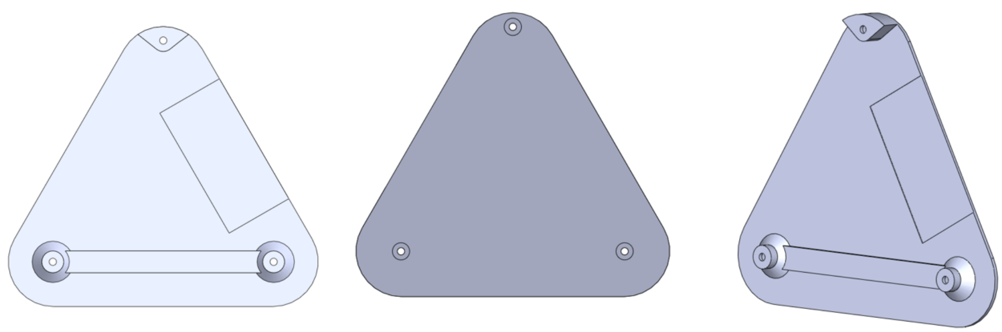


Figure 3.66: Final back lid design on Jun 04, 2018

4 Design Requirements

Our final product will help parents and children to have a less stressful time by using ambient lights, sounds and interactive stories around transitions in everyday activities like brushing teeth.

The aim is to make it easier for children to follow instructions without hassle and stressful communication with their parents; replacing quarrelling with positive, and engaging communication between the parent and their children.

4.1 Overall Objectives

The following aims were most important in building our product and ensure its success.

4.1.1 Provide Buffer Time

Studies have shown that giving children a short buffer period between giving them instructions and expecting them to follow the instructions improves their likelihood of following instructions to perform a certain task without a hassle. The Buffer Time provided allows the child's mind to adjust to the idea of an upcoming transition making them accept the change while simultaneously detaching themselves from the activity they were involved in, before the instruction was given.

4.1.2 Engage the Child

The system involves the use of push buttons to give the child an interactive storytelling experience. While the main aim is to get the child to move to a location for an activity that their parents need them to do, the interactive and personalized storytelling will aid in the transition by keeping the child engaged. The lights on the Milestone display different patterns which provoke the child to press the button to get the different parts of the story started. The stories may be of the general bedtime storytelling theme or be directly related to the child's life: They can relate to their previous experiences or events in the future the child needs to be prepared for. An AI will be used to provide personalized content in the stories based on information provided by the parents through the mobile application.

4.1.3 Develop Positive Habits for Daily Routines

The system will use a distinct color change in the ambient lights to signal the time for a particular activity that the child needs to do. Studies show that the child will soon associate the change in color with that particular activity and be willing to participate in the routine activity more readily. This will develop positive habits for daily routines and significantly reduce the stress and negative communication usually associated with these routines.

4.1.4 Promote Positive Communication Between Parent and Child

The system will focus on creating a shared experience between parent and child, because our main goal is to provide quality time between parent and children. This time will be filled with a story that could be used to start a meaningful and educational conversation.

4.2 Functional Requirements

Requirement	**Method of Realization**	**Rationale**
The Milestone must have no sharp edges.	Curves that define the edges of the piece are continuous.	The use case involves contact with children, therefore the product needs to be safe for use.
The Milestone must not have any loose parts small enough to be swallowed.	Comply with the US Consumer Products Safety Commission small parts regulation.	The use case involves contact with children, therefore the product cannot provide a choking hazard.
The lights on the Milestone must be bright enough to be seen (threshold).	Light must be distinguishable from the background in a well lit room by a 3 year old from a distance of 10m.	The child must be able to see the light patterns in order to know which button needs to be pressed.
The batteries must be accessible for replacement when the Milestone is mounted on the wall.	Adults can change the battery without detaching the Milestone from the wall in less than one minute.	If the batteries cannot be changed rapidly, the parents may stop using the system altogether, when it loses power.
The Milestone must be light enough and the adhesive strong enough for the adhesive tape to hold the weight when mounted.	The milestone sticks to the wall and does not fall off when bumped and is able to remain there for three years.	If the milestone keeps falling off, remounting might become to a big of hassle.
The button must be large enough for the child to press easily.	The button must be larger than 1.5 inches.	Children in the target age group might not have the full motor skills necessary for precise interactions.
The parents need to be able to control the experience in an easy and convenient way.	A smartphone app communicates with the Milestones.	If the method of entering information is not always at hand, the parents will not enter sufficient information.

4.3 Constraints

The application of the Artificial Intelligence is constrained due to the lack of time for acquiring sufficient training data. The AI system used in this prototype will be a basic framework for its application that can be improved in the future through more training data collected over time. The lengths of stories and the buffer time as well as fragment selection and repetition timing might benefit with more training data especially.

4.4 Opportunities

The AI system could also be used for other educational purposes or to start conversations that parents would like to have with the child but face difficulty in having otherwise. The conversation topic could simply be introduced through a story.

4.5 Assumptions

The design of the prototype assumes a family setting involving a single mother and one child between the ages of 3 and 5 years. The design of the prototype assumes the floor plan of the house where the system will be installed to be similar to that shown in Figure 4.1.

Figure 4.1: An example of a floor plan of a house, in which VAMO could be installed in.

Figure 5.1: The Milestone device mounted to a wall

5 Design Specifications

This chapter describes our hard- and software design of the final Milestone as shown in Figure 5.1, of the Python web server running on the Milestones and of the Android App to control the experience.

5.1 CAD

All computer aided designs are finished in SolidWorks 2017. The final CAD design consists of four parts: main housing (Figure 3.60), battery chamber door (Figure 3.64), back lid (Figure 3.66) and LED ring cover (Figure 3.62). The final assembly view from various orientations can be found in Figure 5.2, Figure 5.3, Figure 5.4 and Figure 5.5. An exploded view of the milestone assembly is also presented in Figure 5.6 to describe the alignment of each component in the device.

5.2 Drawings

Engineering drawings were exported from SolidWorks. Since the manufacturing method we took is 3D printing, drawings were not used in the manufacturing process. Drawings are listed below for reference only. The drawing for the main housing is in Figure 5.7. The drawing for the back side lid is in Figure 5.8. The drawing for battery chamber door is in Figure 5.9, and the drawing for LED ring cover is in Figure 5.10.

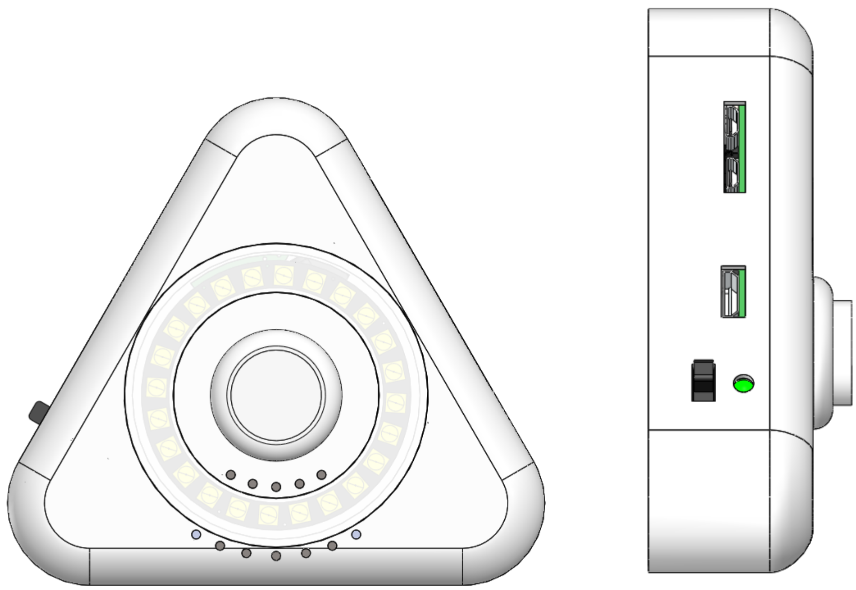


Figure 5.2: Milestone assembly front view (left) and left view (right)

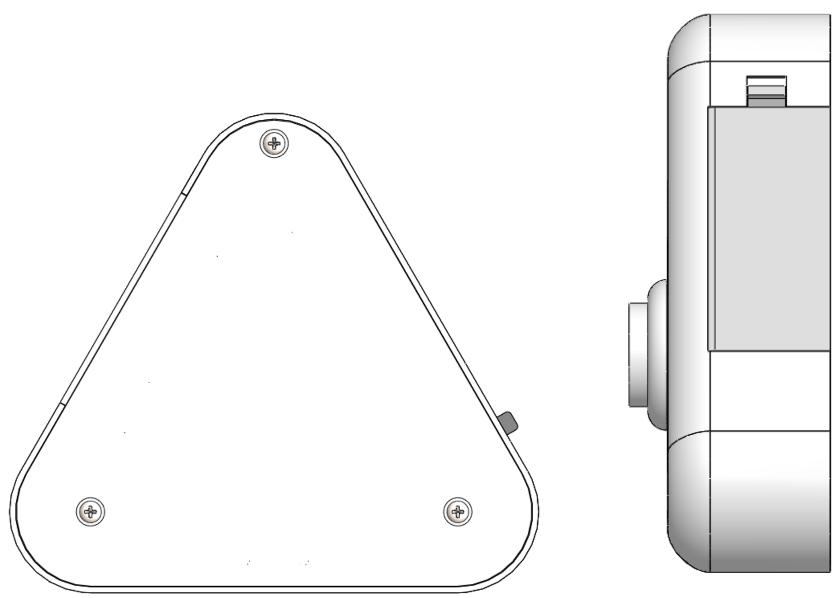


Figure 5.3: Milestone assembly back view (left) and right view (right)

Figure 5.4: Milestone assembly right tilt view (left) and left tilt view (right)

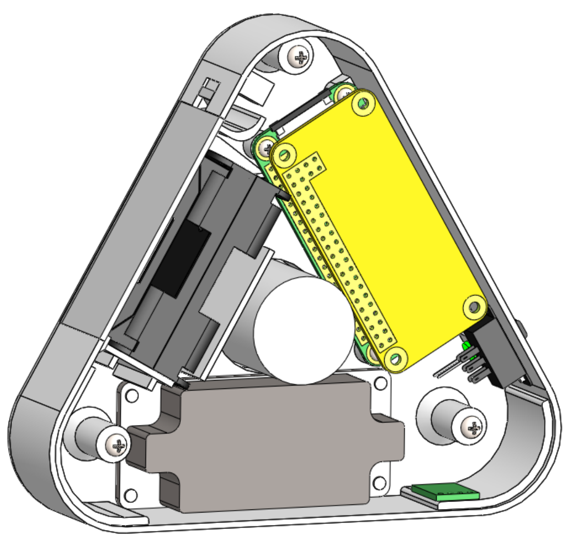


Figure 5.5: Milestone assembly backside view

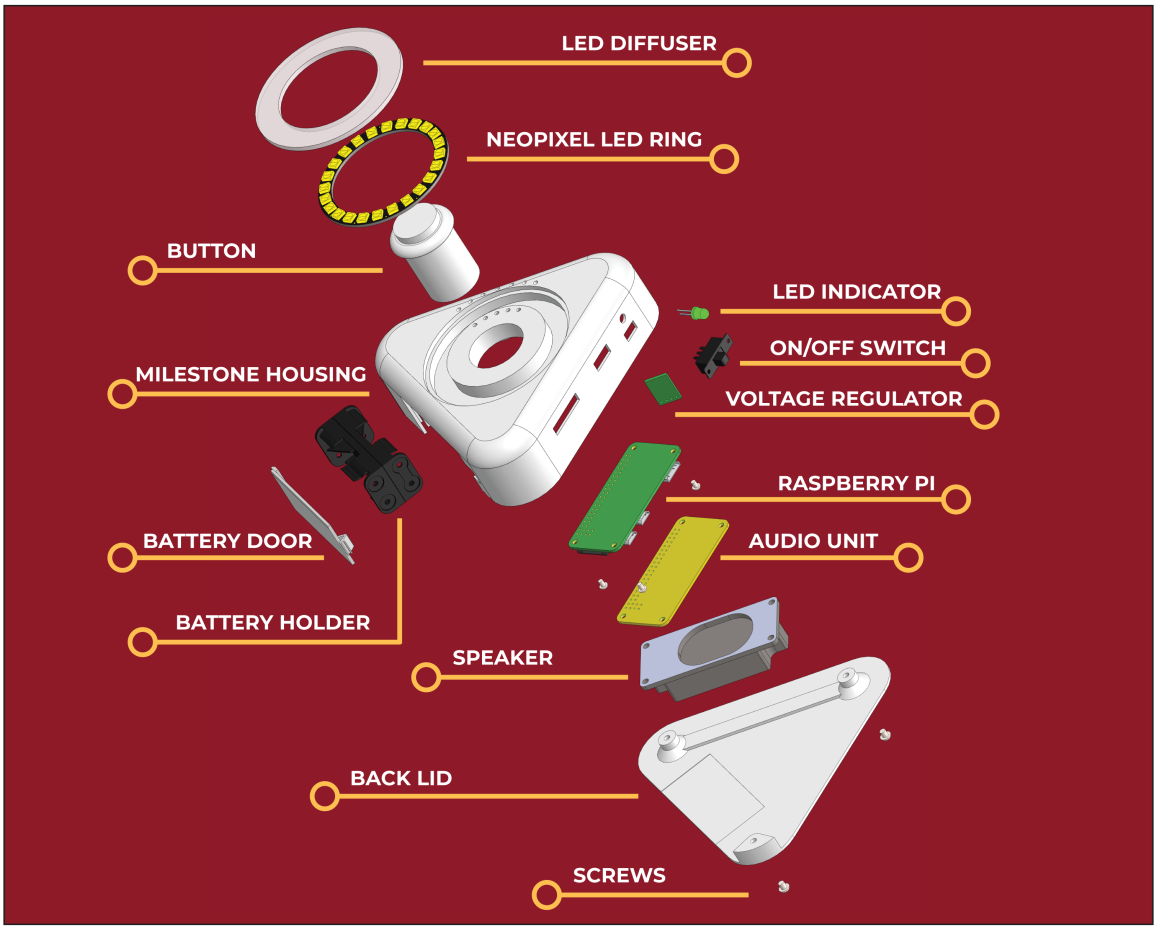


Figure 5.6: Milestone exploded view

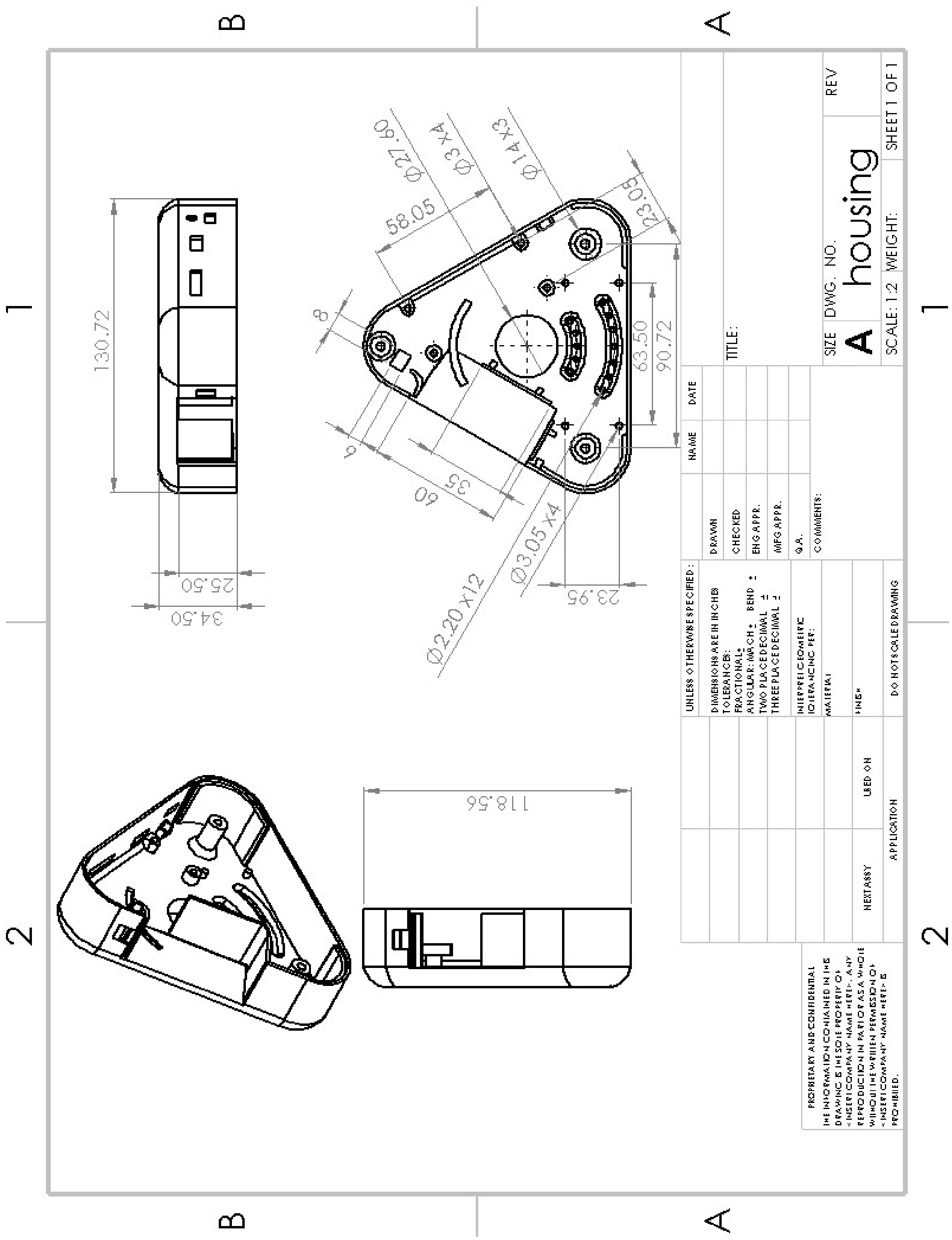


Figure 5.7: Main housing drawing

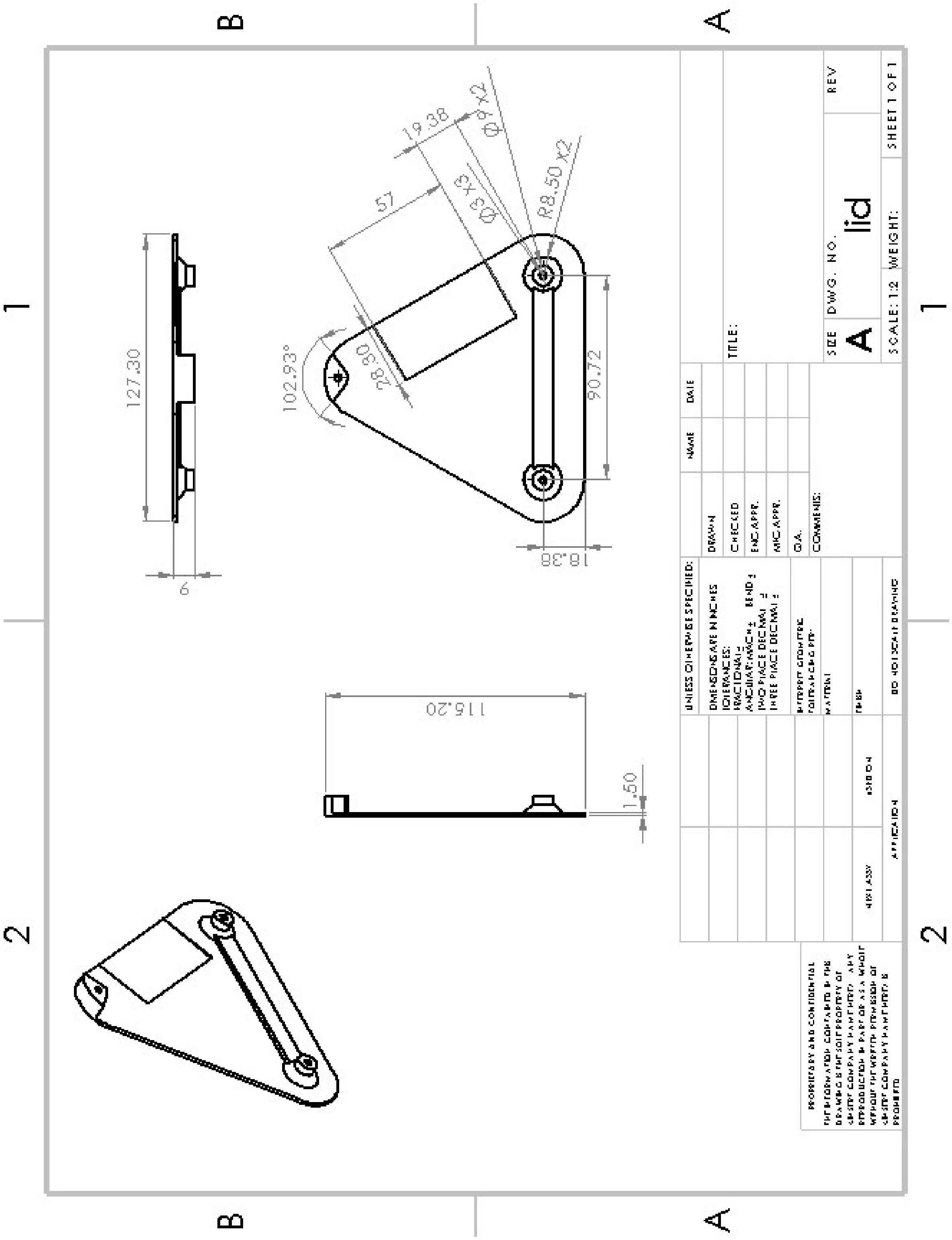


Figure 5.8: Back lid drawing

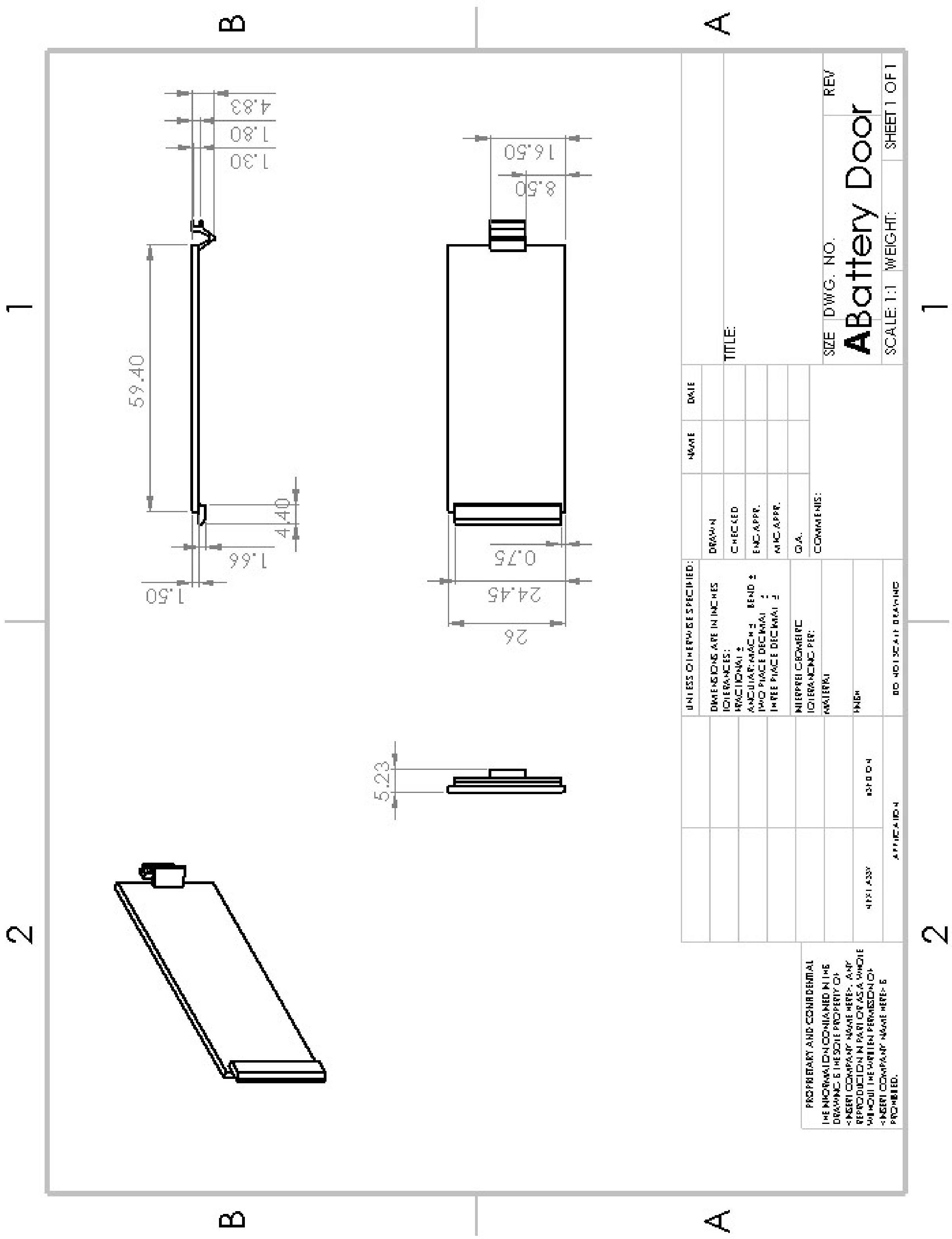


Figure 5.9: Battery door drawing

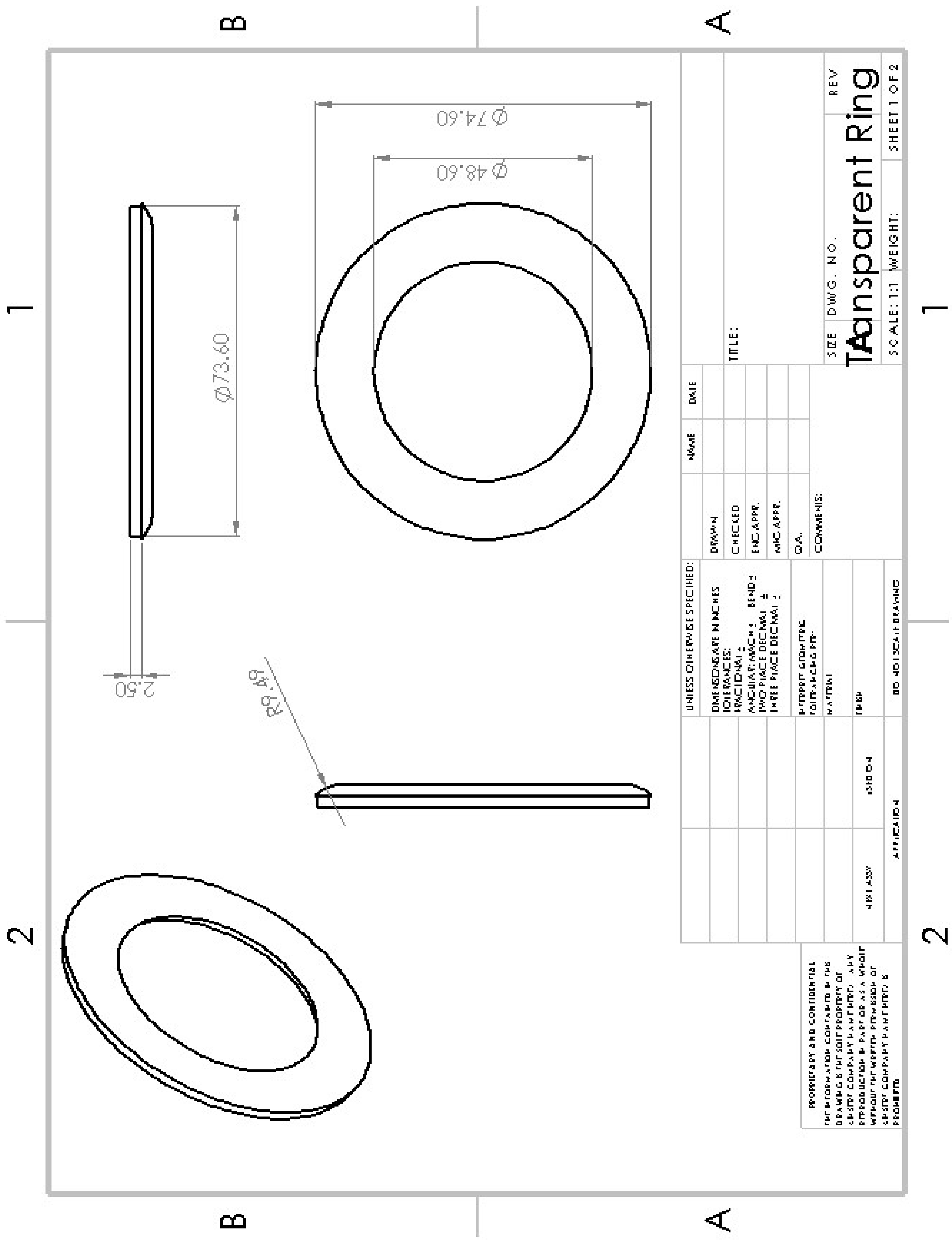


Figure 5.10: Ring cover drawing

5.3 Wiring Diagram

A wiring diagram is shown in Figure 5.11 to represent the connectivity of each electronic component in the Milestone. The power source is connected to the switch and then to the voltage regulator. Therefore, the switch controls all possible current draw from the power source. From the voltage regulator, wires are connected to the LED indicator and Raspberry Pi in parallel. Both speaker pHAT and NeoPixel 24 Ring draw current from the Pi.

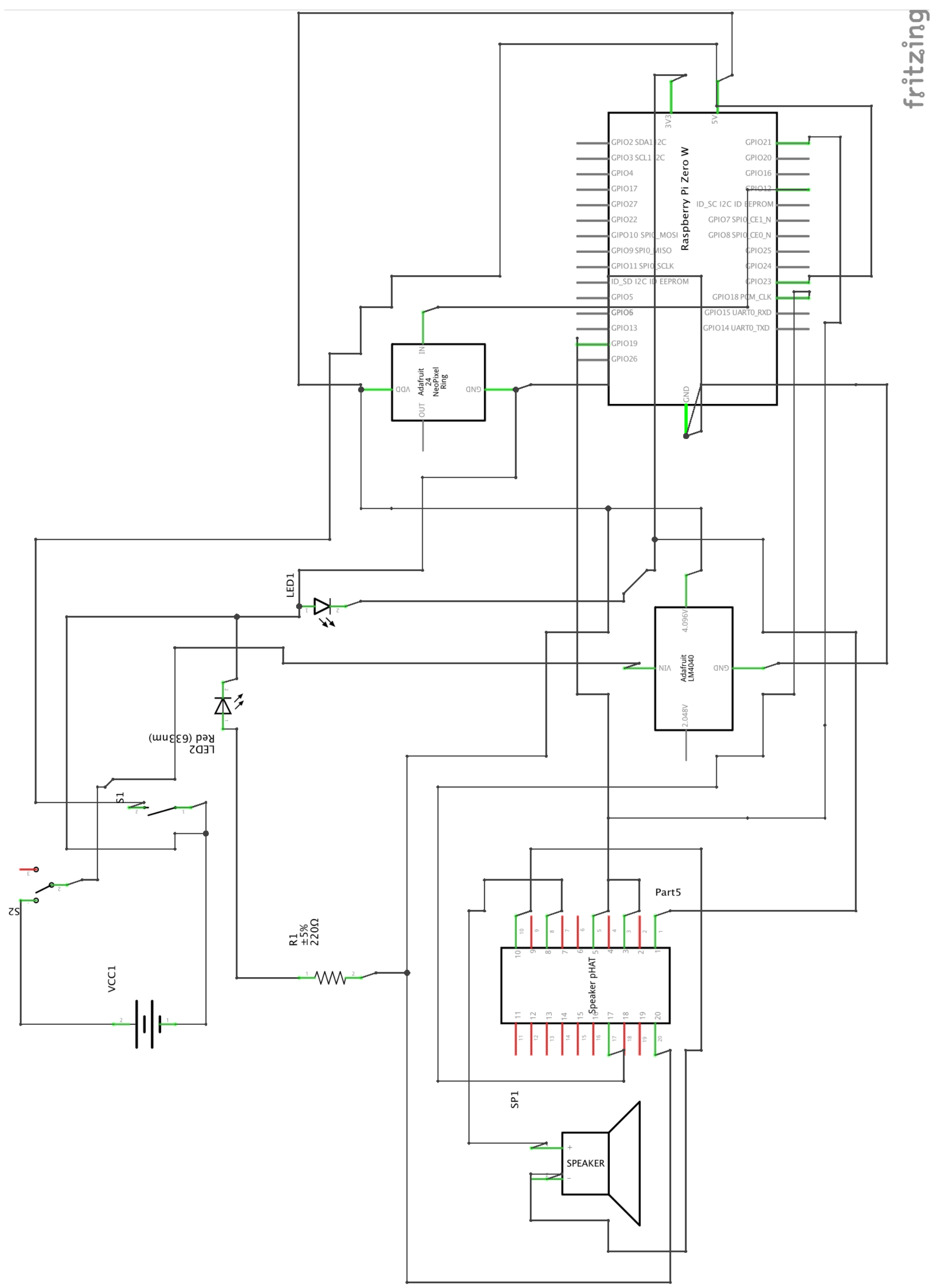


Figure 5.11: Milestone wiring diagram

5.3.1 Bill of Materials - BOM

The following Bill of Materials (BOM) in Table 5.1 lists all the components that are used in the milestone, including the purchased items and 3D printed parts. The total cost is 92.88 USD.

No	Qty	Description	Price Per Item	Link	Subtotal
01	1	Adafruit Stereo Enclosed Speaker Set - 3W 4 Ohm	5.41	http://amzn.com/B00QSJ1HA4	5.41
02	1	Pimoroni Speaker pHAT for Raspberry Pi Zero	12.49	https://www.adafruit.com/product/3401	12.49
03	1	EG Starts 5x 30mm Arcade LED Lights Push Button Built-in Switch 5V Illuminated Buttons	2.00	http://amzn.com/B01N11BDX9	2.00
04	1	Adafruit 24 RGB LED Neopixel Ring	16.95	https://www.adafruit.com/product/1586	16.95
05	1	Raspberry Pi Zero W	10.00	https://www.adafruit.com/product/3400	10.00
06	4	2-56 3/16” Steel Pan Head Phillips Screw	0.03	https://www.mcmaster.com/#standard-rounded-head-screws/=1d78rio	0.12
07	4	2-56 Brass Heat-Set Threaded Inserts for Plastics	0.10	https://www.mcmaster.com/#catalog/124/3395/=1d78min	0.40
08	3	4-40 3/16” Steel Pan Head Phillips Screw	0.03	https://www.mcmaster.com/#standard-rounded-head-screws/=1d78txi	0.09
09	3	4-40 Brass Heat-Set Threaded Inserts for Plastics	0.10	https://www.mcmaster.com/#catalog/124/3395/=1d78t38	0.30
10	4	AAA Battery 1.5V	0.33	http://amzn.com/B0002DUQDQ	1.32
11	1	3M Adhesive Tape RP25, 2” Diameter Circles	2.27	http://amzn.com/B00HV6ZGOG	2.27
12	1	Switch	0.73		0.73
13	1	LED	0.22	https://www.mouser.com/ProductDetail/Cree-Inc/C512A-WNN-CZ0B0151	0.22
14	1	4*AAA Battery Holder	1.58		1.58
15	1	5V Step-Up/Down Voltage Regulator S9V11F5	5.95	https://www.pololu.com/product/2836	5.95
16	1	Lid	8.00		8.00
17	1	Housing	20.00		20.00
18	1	Battery Door	2.00		2.00
19	1	Ring	3.00		3.00
20	1	Electrical Wire	0.05		0.05

Table 5.1: Bill of Materials, total price: $92.88

5.4 Software for Milestones

This section explains the software specifications for the Milestones and specifies their connection to Smart Light systems in the home, specifically the Phillips Hue as well as a breakdown of the Android app and how the user interacts with it and what the app does in the backend.

5.4.1 Python Flask web server

The Milestones are equipped with a Raspberry Pi running Raspbian[1]. Thus, they can run a variety of software available for Linux. As we knew that the Milestones should be contacted by the app, it was obvious that the Milestones should run a web server. For the sake of simplicity and existing knowledge within our team, we decided to use the Python-based web server Flask. Python[2] is an interpreted high-level programming language having a big community support with many different packages. We used the newer version Python 3 for our prototype. Flask[3] is one of these packages providing an extensible web server. In addition to Python, our software relied on `ffmpeg`[4] to edit sound files and on `mpg321`[5] to play these. Except for uploading the audio file, all actions are performed with a `GET` request.

5.4.1.1 Story States within the Python server

```python
class StoryState(Enum):
    Waiting_for_preparation = 1
    Preparing = 2
    Sending = 3
    File_Received = 4
    Waiting_for_start = 5
    Waiting_for_button = 6
    Waiting_for_previous = 7
    Playing = 8
    Done = 9
```

Source Code 5.1: All possible states of a Milestone

Each Milestone has an internal state (out of nine possible) defining whether the story has been prepared, distributed, or is being played back or finished completely. All states and their order are show in Source Code 5.1. The first one (`Waiting_for_preparation` is entered after the device boots, after the Hue Bridge has been found (see subsection 5.4.2) and the service has started to get announced (as described in subsubsection 5.4.1.2. After that, the Android app performs a `GET` request to prepare the story indicating which experience and event to use and how long the buffer time should be: `prepare_story?exp=money&event=city&buffer=15`. The Milestone than enters the `Preparing` state indicating that the story is being generated (c.f. subsection 5.4.3).

[1] Raspbian is a version of the Debian OS optimized for the Raspberry Pi – http://raspbian.org
[2] Website of the Python language – http://python.org
[3] Homepage of the flask microframework – http://flask.pocoo.org
[4] https://www.ffmpeg.org
[5] http://mpg321.sourceforge.net

In addition, the smart lights will start indicating the buffer time. All other Milestones enter the second state without preparing the story to ease the app to get the current status. Once the story files are generated on one device, they are distributed to all Milestones using a `POST` request. The Milestones reflect their progress either using the `Sending` or `File_Received` state. After finishing the file distribution, all Milestones enter the `Waiting_for_start` state. From now, the user might end the buffer time and start the story through the app using another `GET` request. Depending on the order, each Milestone will either enter the `Waiting_for_button` state to start the playback or the `Waiting_for_previous` state if one or more previous Milestones have not finished playing their part of the story. This will also disable the smart light and turn it off while simultaneous the LED pixel ring on each Milestone lights up (see subsubsection 3.4.4.2). If the Milestone is in the `Waiting_for_button` state and the button gets pressed, it will start playing the first part of the story in using `mpg321` in a new thread and indicate that through the `Playing` state. As soon as the playback finishes, the Milestone enters the `Done` state and notify the next Milestone to change its state from `Waiting_for_previous` to `Waiting_for_button`.

5.4.1.2 Milestone Discovery

In order for the Milestones to be accessed by the app (e.g. to start a story) or other Milestones (in order to listen to the next part of a story), they need to advertise themselves. For this, we used a combination of multicast DNS (mDNS) [7] and DNS-Service Discovery (DNS-SD) [6]. The first one describes how to resolve DNS queries in a network without a unicast DNS server, while the second one introduces a way to query hostnames that offer a desired service. We will refer to this technology as Zeroconf[6] which is the superordinate term for both. We used the HTTP Service `._http._tcp.local.` to announce the web server and the weight to identify the order of the different Milestones. The higher the weight, the later the Milestone should be in the story telling. So far, this weight has to be configured during setup in a configuration file and is assumed not to be found twice in the same network. In addition, it has to be ensured that the hostnames of all Milestones are unique. Each Milestone keeps a list of other discovered Milestones and updates it on changes. The app uses the same technology to discover the Milestones and sends requests to any of the available Milestones. Using the `zeroconf` Python package, the service discovery works parallel to the Flask web server.

5.4.1.3 Connecting hardware: LED pixel ring and button

The LED pixelring we used from Adafruit (c.f. Table 5.1) was not mainly designed to be used with a Raspberry Pi but instead to work with an Arduino. This is due to controlling the color of the LEDs using Pulse-width modulation (PWM). PWM uses one signal and timings to encode information. As the Raspberry Pi does not run a real-time operating system, the exact timings cannot be ensured. We used the Python module of the open-source library `rpi_ws281x` [7] to steer the LEDs. This worked satisfying, even though we sometimes experienced a crash of the Python interpreter when switching the color of the LEDs. Unfortunately, we were unable to identify the root cause for this issue but think it is related to timings or a race-condition within the library. All requests made to the API of the library are blocking and thus prevent the one-threaded Flask web server from answering other requests. For the green and blue light of the LED pixel ring, we wanted to show a pulsing light

[6] Website regarding Zero Configuration Networking – http://zeroconf.org
[7] https://github.com/jgarff/rpi_ws281x

(meaning that it fades from a high brightness to a low brightness and back in a loop). To implement this effect, we added a thread running in the background regularly checking for a condition to stop. Besides the LEDs, we connected an illuminated push button to the Raspberry Pi. Instead of polling the state of the button, we registered a callback to get notified once the button has been pressed. For that, we used the Python module `RPi.GPIO`, an internal pull up resistor of the Raspberry Pi and one second of software debouncing. This combination ensures that the callback is only called once reliably.

5.4.2 Connecting to Smart Lights

The python code on the first Milestone connects to an available smart lighting solution in the house to signal the progressing of the Buffer Time. We decided against implementing the control of the smart light within the smart phone app, as we cannot ensure, that this app will be active during the whole Buffer Time. The app however sends an HTTP signal to the first Milestone to start the Buffer Time and for how long. We used the Philips Hue smart lights for our final product. The first Milestone connects to the Philips Hue Bridge, which in turn controls the different light sources. Connecting to the smart lights requires two essential steps for the Milestone:

1. The Milestone is able to discover a Philips Hue Bridge in the same WiFi network, without any additional setup. When the Milestone starts it issues a broadcast request to the local network and thereby scans the MAC addresses of all local devices. The Philips Hue Bridge has a known pattern in the MAC address (starting with "00:17:88"), which is used to identify this device.

2. Once the Milestone receives the signal to start the Buffer Time, it will periodically signal the first smart light bulb connected to the Philips Hue Bridge to change its color gradually from white to red. Updates are send every 0.2 seconds. After the Buffer Time is over, the Milestone sends commands to the Hue Bridge to make the light bulb blink 11 times with increasing frequency and after that 4 times with the highest frequency.

All communication between the Milestone and the Philips Hue Bridge is based on the HTTP interface of the Philips Hue Bridge, which allows us to query its status and send instructions in a JSON based format. Before any command can be send, an API key needs to be generated on the Hue Bridge, that is later used in every request. The API key can be generated by calling an HTTP function on the Hue Bridge. Listing 5.2 shows a sample request, that is used to change the color of the smart light.

```
http://{hue-address}/api/{api-key}/lights/2/state
message body (set color to white): {"hue":46920.0}
```

Source Code 5.2: POST Request issued to the Philips Hue Bridge to change the color of the light.

5.4.3 Story Generation Algorithm

The vision of our final product includes an intelligent story telling algorithm, that runs on the smart phone to make use of dedicated AI hardware. For the prototype of the final system, that we implemented, we decided to use a simpler algorithm that combines fragments of two kinds. (3.4.4). The simple intelligent algorithm we implemented combines story fragments stored directly on the first Milestone.

The story fragments have to be provided as mp3 files in a specified folder in the Milestone for the algorithm to detect them. Using mp3 files was the most flexible solution, as our system prototype can support both stories generated by text-to-speech from text files and human-recorded stories. For the final demonstration, we decided to use two stories, that were both generated by the Google text-to-speech system in advance and stored as mp3 files on all Milestones.

The provided story fragments need to be annotated to provide additional information to the system. We chose to use JSON files to provide this meta information, which the python script on the Milestone reads in before the algorithm starts to combine fragments. An example of such a file can be seen in listing 5.3 and listing 5.4.

```
{
  "type": "mainstory",
  "connectorStart": ["street", "walking", "grandpa", "dog"],
  "connectorEnd": [],
  "tags": ["icecream", "toy shop", "traffic lights", "traffic", "car", "city"],
  "storyFile": "city_trip.txt",
  "introFile": "city_trip_intro.txt",
  "audio": "city_trip.mp3",
  "introAudio": "city_trip_intro.mp3",
  "splitFileAt": "59000"
}
```

Source Code 5.3: Main Fragment example configuration in JSON file format.

```
{
  "type": "sidestory",
  "connectorStart": [],
  "connectorEnd": ["street", "walking", "grandpa"],
  "tags": ["money", "bank", "shoes", "bank account"],
  "storyFile": "money.txt",
  "audio": "money.mp3"
}
```

Source Code 5.4: Intermediate Fragment example configuration in JSON file format.

The two fragment types require different kinds of information to be provided. As described in section 3.4.4, there are short intermediate fragments and longer main fragments. The whole story is intended to be played on three Milestones. The first Milestone plays the introduction of the main story, which is provided as a separate mp3 file and the whole intermediate fragment. The second Milestone plays the first part of the main fragment. The last Milestone plays the rest of the main fragment. The main fragment thus has to be split in two parts again. The parameter splitFileAt determines, where the algorithm should split that file.

The story generation algorithm works by finding matches of tags, that are provided for each fragment, and information that has been entered in the smart phone app. We added an additional constraint, that ensures a smooth transition between main and intermediate fragments. We do this by providing an array of connector tags for each fragment. The connector tags specify which persons are involved at the end or start of the fragment and in which situation they are. Only if all connector tags of the intermediate fragment can be found in the main fragment, the fragments are allowed to be combined. This technique ensures, that there are no abrupt changes of environment or activity in the final generated stories and that no acting persons or animals are abruptly missing.

The story generation algorithm will take this constraint into consideration and thus selects the first combination of main fragment and intermediate fragment, that does not violate the connection constraint and has the highest matching score over all possible combinations, that means it has the most tags, that match with the information about experiences and events entered to the app.

Finally, after the story generation algorithm finished, three mp3 files (one for each Milestone) are created on the first Milestone. The second and third file are transferred to the corresponding Milestones using a POST request. Each file is stored under the predefinied name `sound_part.mp3` to remove the need to keep track of the most recent file.

5.5 Android App

For the parents to control the whole experience, we created a smartphone app. Since most of our personal devices used Android and Huawei sells Android phones, we decided to target Android. The app has four main fragments for everyday interaction and a notification mechanism, but does not include any functionality for setup (c. f. section 7.1). We will describe the functionality of the app in the following sections.

5.5.1 Home Fragment

The home fragment (Figure 5.12) includes a quick way to add experiences and events, to make the process as fast and hassle-free as possible. For both, an entry will pop up, with which all available experiences and events can be searched. After an event is selected, a date picker will let the user choose the

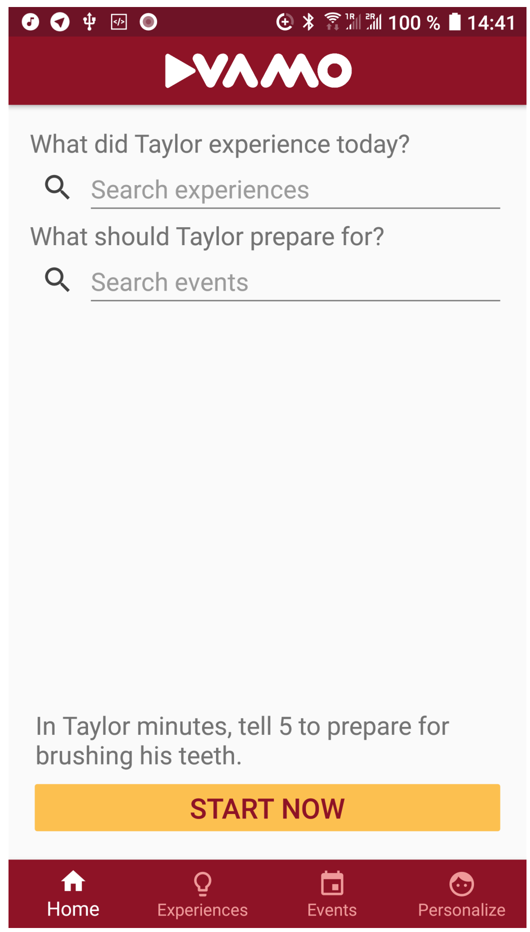


Figure 5.12: The Home fragment of the Android app provides a status overview and widgets for adding experiences and events.

date of the event, while for the experiences the current date is taken. There is also a status display on the fragment, which show when the Buffer Time/story will start or if it has already started.

5.5.2 Experience Fragment and Events Fragment

These fragments (Figure 5.13) include the same widgets for adding experiences/events as the home fragment as well as a list of the ones already added. Those can be removed via buttons.

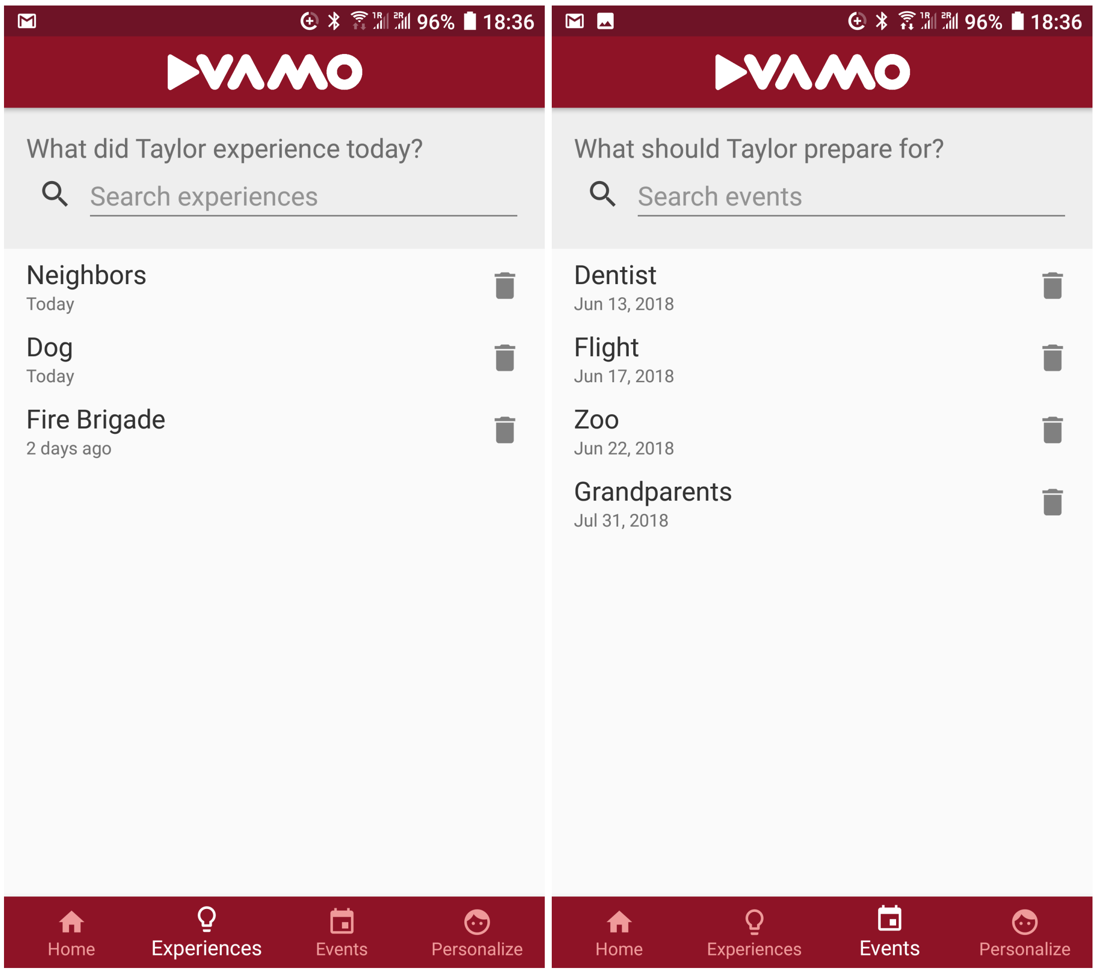


Figure 5.13: The Experiences and Events Fragment let you edit the list of experiences and events.

5.5.3 Personalize Fragment

In this fragment (Figure 5.14) the name and birthday of the child can be set. There are some sliders, which let the parents describe the personality of the child. These are based on the Big Five personality traits [10]. This is followed by a list of activities VAMO is set up to help with and a button to add more.

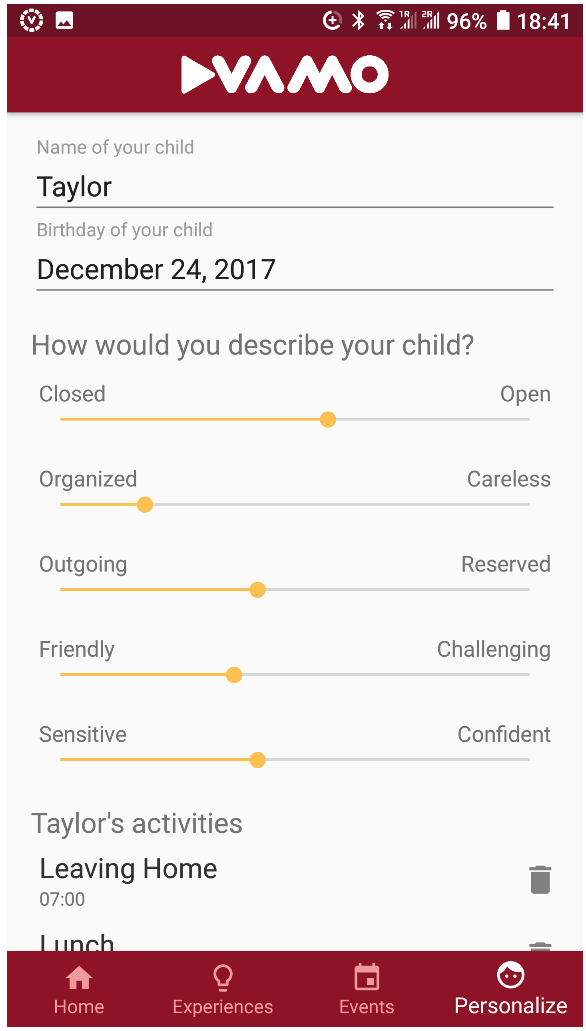


Figure 5.14: The Personalize fragment of the Android app lets the parents enter data about the child and configure activities.

5.5.4 Notifications

Using the Android `AlarmManager` the app schedules alarms for Buffer Time start and story start, which trigger notifications. Additionally, a notification (fig. 5.15) is triggered 10 minutes after the story ended in which the parents can answer if the story went well or not, without starting the app.

5.5.5 Backend

The app data is locally saved on the phone using a `SharedPreferences` object. Using HTTP requests it is communicated to the Milestones, in which the story will be played when each individual Milestone should become available. Android's `NsdManager` is used to implement the service discovery using Zeroconf.

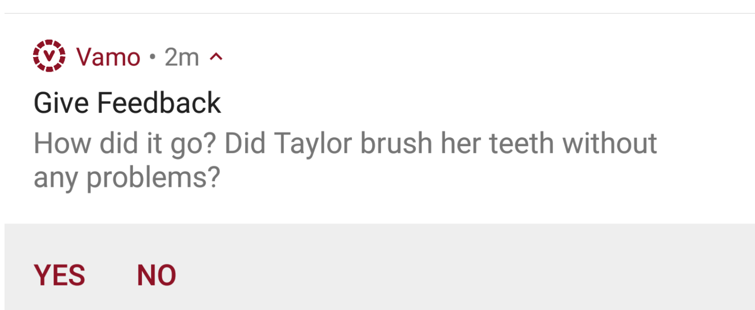


Figure 5.15: The notification shown by the Android app after the story is finished.

6 Key Learnings

Throughout the year long project, a number of important knowledge and experience were learned.

6.1 3D Printing

For all the Milestones that we printed, we tried three different printing methods: Fused Deposition Modeling (FDM), Stereolithography (SLA) and PolyJet. We found the pros and cons of these printing methods for our parts.

6.1.1 FDM

Fused Deposition Modeling (FDM) is an additive manufacturing (AM) technology commonly used for modeling, prototyping, and production applications as trademarked by Stratasys Inc [18]. This printing method is fast and has decent resolution (up to 0.05" for FDM F370 Figure 6.1). The problem with this method is that, certain features in vertical direction is easy to break. In addition, the support in printing process is hard to remove.

Figure 6.1: Stratasys F370 FDM Printer

6.1.2 SLA

Stereolithography (SLA) is a form of 3D printing technology used for creating models and prototypes in a layer by layer fashion using photopolymerization, a process by which light causes chains of molecules in liquid resin to link, forming polymers. The machine we used in printing is Form 2 (Figure 6.2) from FormLab. It has a nice surface finishing and supports are easy to remove. For the white resin we used, the resolution goes up to 0.05mm. However, this printing method is much slower because of high resolution, as well as because it takes time for liquid resin to stabilize after each layer of printing.

Figure 6.2: Form 2 SLA Printer

6.1.3 PolyJet

We contracted out to Autotiv Manufacturing Corporation to print our parts using PolyJet (c.f. Figure 6.3). With microscopic layer resolution and accuracy down to 0.1 mm, it can produce thin walls and complex geometries using the widest range of materials available with any technology [19]. We used white vero as our material in this printout. The advantage is that it can produce very complex structures and has very few design limitations. However, it is more expensive compared with the other two methods we tried.

6.2 Energy Consumption

The key learning about the energy consumption of our Milestone is the difference between different types of batteries and their most suitable use cases (Figure 6.4). Alkaline batteries are cheaper compared to Lithium batteries. However, their internal resistance increases significantly as the current raises. Therefore, a lot of energy is converted to heat energy and battery life is very short for our application as the current draw is around or over 500mA. Lithium batteries is better in applications with large current draw. The internal resistance increases only slightly as the current goes over 500mA. Therefore, it is the preferred battery for our use case.

Figure 6.3: PolyJet Printer

Figure 6.4: Comparison of AAA Alkaline (left) and Lithium (right) batteries

6.3 Software

We had some technical issues shortly before and during EXPE that could have been prevented through further testing and more preparation: The story failed often failed to play at the second Milestone, while before it had always worked. Through more stress testing under difficult conditions and also repeating the tests more often, we may have been able to spot that mistake earlier and taken corrective measures. We also added a hidden debug screen to the Android app just one day before EXPE. This screen allowed us to trigger resets, some notifications and show available Milestones, which would have been useful for us earlier. Lastly, we programmed a functionality for EXPE in the hour before, because we thought it would be very easy, but a bug appeared we could not fix, so we had to spend some time on a workaround. This made the preparation a lot more stressful and could have been avoided.

6.4 Story Preparation

We noticed that for story generation, we could not reuse existing stories as they were because we were using a fragment based approach to combine stories. This is a weak point of the system since stories have to be specifically designed for this system. We also found it comparatively hard to write stories, that last only for a short amount of time (2 to 5 minutes) and still convey a coherent narrative that has enough content to fully engage the kid. Our system thus faces a trade off: On the one hand, we want to keep the stories short not to prolong the daily routine unnecessarily, which is also not wanted by the parents, on the other hand, the stories should contain enough content and context information to engage the kid.

6.5 EXPE Feedback

Figure 6.5: For EXPE we set up a hallway connecting a play area with a bathroom, where people could listen to an example story played back by three Milestones.

During EXPE we also got some valuable feedback about our solution, which we showcased in the form of a hallway connecting a play area and bathroom, equipped with three Milestones and a very short story. The hallway with some people listening to the story can is pictured in Figure 6.5. Most people liked our solution and we even heard from other teams, that their visitors told them how nice our booth was. Many also wanted to know, whether and when our solution will be released as a product. We also received some new ideas, such as putting the Milestone directly on an outlet or doubling it as a nightlight. Many visitors also would have liked us to continue the story during and after the brushing started, which we thought about before, but did not thoroughly evaluate because of time constraints. The possibly most positive feedback, however, we received in the form of 5 year old child, who enthusiastically pressed all buttons and listened to the story until he finally reached the bathroom as seen in Figure 6.6.

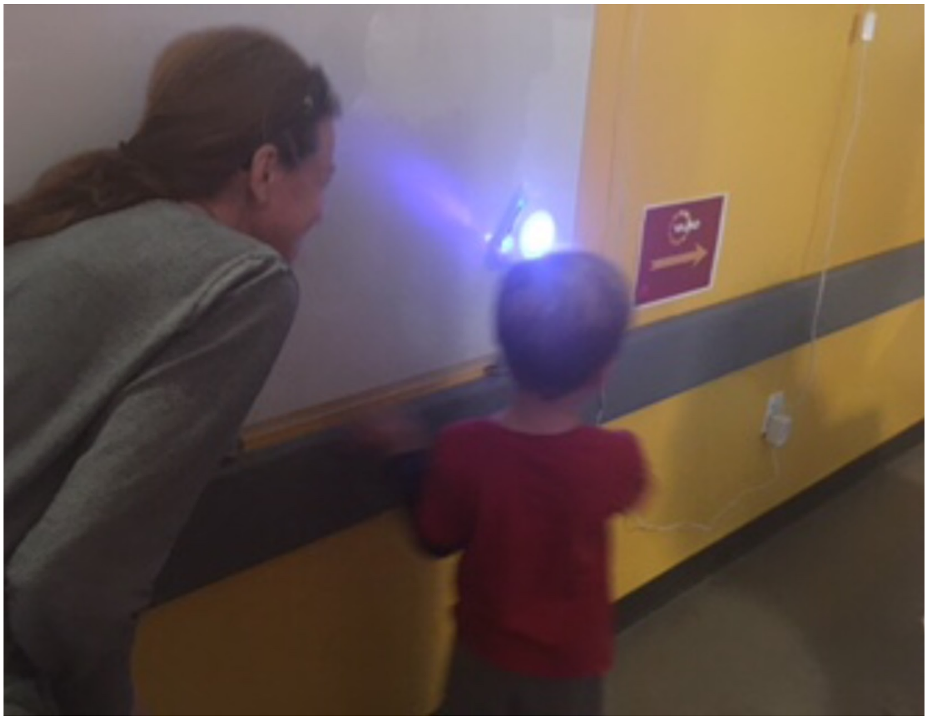


Figure 6.6: A 5 year old testing VAMO during EXPE

7 Next Steps

In this chapter we will describe, what would still be necessary to release VAMO to a broad audience and would can still be done to improve it in the future.

7.1 Setup Process

Before usage VAMO needs to be setup: All the Milestones need to be connected to the same WiFi router and linked to the smartphone app. Also, their order or for setup with multiple activities, their locations need to be recorded in the app. WiFi setup of components with no screen and limited input capabilities is essentially a solved problem also required for many other commercial products such as Google Chromecast Audio[1] or smart lighting technology. While ordering the button is possible by simply clicking pressing the buttons in order during the setup, recording their location without requiring the user to input the floor plan might be challenging experience to design.

7.2 Mass Production

The housing components are expected to be injection molded in mass production. Compared with 3D printing in our prototyping process, injection molding produces the parts at a much cheaper price and faster speed, as well as consistency. The electronic components can also be customized in mass production and only keep the functions that are related to our device. For example, the current Raspberry Pi has built-in Bluetooth, which we are not using, and far too much processing power. The speaker pHAT has an LED bar showing the volume and in our application, this function is not used. Therefore, in mass production, we will have a customized board that consists of microprocessor, WiFi module, speaker driver and amplifier. This might also lead to greater power savings, so that the battery can last longer. It also means that the form factor can be updated to a more compact form since the customized board can have a specifically designed shape and will likely be smaller than the Raspberry Pi Zero and speaker pHAT arrangement.

7.3 Obtaining Stories

Since stories cannot be fully created by artificial intelligence yet, someone needs to write story fragments for our product. There are many stories for preparing children for events or expanding on their experiences already, so it might be possible to make a contract with either their authors or the publishing companies behind them. Additionally, characters of popular children's books could be incorporated, which the child might already be familiar with. These sources have the benefit of only providing stories from trusted entities, so the parents would not need to listen to every story and check for inappropriate content.

[1] https://store.google.com/product/chromecast_audio

Users might also want to write their own stories and even publish them for other users, so a platform for user generated content could be setup. Here, it might be possible to add reviews and ratings, to circumvent the problem of stories being added that are not appropriate for children.

7.4 Artificial Intelligence

For now, our story creation algorithm is very simple and relies on pre-written fragments that match an exact interface. With recent advances in text processing and creation, our algorithm could become more and more advanced, as our product is developed further. The following challenges could potentially be solved by AI in the future:

- Finding out if fragments fit together, so that the story still makes sense.
- Connecting two fragments that do not fit together by creating a transition.
- Checking stories for inappropriate content.
- Creating an entire story solely based on the user input.

Also we have not obtained any training data yet, so lots of user testing needs to be conducted to make AI algorithms work for our problem.

7.5 Data Privacy

Especially when using child data for machine learning, measures need to be put in place to protect their data. It should also be possible for parents to opt-out of data sharing altogether and never upload any data to a remote server.

8 Project Management

The work presented in this documentation is surrounded by a self-organized project management including time management, a team budget and the use of communication tools. This chapter will give an overview about our internal management and conclude with personalized reflections of the project.

8.1 Project Timeline

Throughout the course, different assignments are due, which were team-external milestones. Table 8.1 visualizes the deliverables, the due dates and which university requested the corresponding submission.
In order to synchronize between the two local teams, we scheduled a weekly video call every Wednesday using appear.in (c.f. subsection 8.2.6).

Within the team, we mainly concentrated on three milestones in the winter quarter:

- prototype without any limitations to get an understanding of radical ideas and what users think about these (dark horse prototypes).
- focus on a user target group and understanding of their needs.
- create prototypes for the newly identified user group and gaining some feedback.

The second and third milestone were quite important for us. One of the goals of winter quarter was to get the Stanford and HPI teams in sync with respect to a similar user group. We selected this goal in order to ease the convergence process on a specific need. While agreeing on a user target group was very easy, it took both teams some time to get familiar with the selected user group and conduct interviews with representatives from this group.

Instead of creating a Funky prototype, the HPI team focused on additional interviews to gain a better understanding of families and their communication preferences. While many previous prototypes already included the key-requirements for Funky prototypes (c.f. subsection 3.2.11), this step gave us the opportunity to learn more about communication within families.

The HPI students presented the overall progress of the project at Karlsruhe in March. We were required to prepare a 2 minute pitch in front of the large audience and a booth, which we could use to gather feedback from a large number of interested and creative visitors. In five rounds, our booth was visited by 10 to 20 people each. We presented the robot arm prototype and collected feedback for it, but also showed other prototypes according to the interests of the visitors.

Date	**Description**	**Submission at**
11/02/2017	Corporate Context / Needfinding handout	Stanford
11/06/2017	Needfinding / Benchmarking handout	HPI
11/07/2017	Launch Venture Report	Stanford
11/14/2017	Benchmarking / CEP handout and presentation	Both
11/28/2017	Benchmarking / CEP writeup	Stanford
11/29/2017	CFP handout and presentation	HPI
11/30/2017	CFP handout and presentation	Stanford
12/07/2017	Brochure	Stanford
12/07/2017	Fall presentation	Stanford
12/12/2017	Fall documentation	Both
12/13/2017	Fall presentation	HPI
01/23/2018	Dark Horse 1 handout and presentation	Stanford
01/30/2018	Dark Horse 2 handout and presentation	Stanford
01/31/2018	Dark Horse handout and presentation	HPI
02/13/2018	Funky System handout and presentation	Stanford
02/14/2018	Funky System requirements and additional interviews	HPI
03/06/2018	Functional System handout and presentation	Stanford
03/07/2018	Functional System handout and presentation	HPI
03/12/2018	Winter presentation	HPI
03/15/2018	Winter presentation	Stanford
03/20/2018	Winter documentation	Both
04/12/2018	Spring Hunting Plan handout and presentation	Both
04/17/2018	X is Finished handout and presentation	HPI
04/19/2018	X is Finished handout and presentation	Stanford
04/24/2018	Manufacturing Plan writeup	Stanford
04/24/2018	Design Requirement section writeup	Stanford
04/25/2018	Manufacturing Plan presentation	Both
05/03/2018	Final Project Definition	Stanford
05/23/2018	Penultimate testing and presentation	HPI
05/23/2018	Penultimate Briefing Packet	Stanford
06/02/2018	EXPE Brochure	Both
06/04/2018	EXPE Poster	Both
06/05/2018	Convergence presentation	Stanford
06/07/2018	Spring presentation	Both
06/07/2018	EXPE	Both
06/12/2018	Spring documentation	Both
06/28/2018	Spring presentation	HPI

Table 8.1: Overview of all assignments and deliverables

The goal of the Stanford visit at the end of March was to further converge. The agenda of the Stanford visit is attached as Figure A.2 in the chapter A. In addition, our goal was to create a first version of the Hunting Plan (see Figure A.3 in the chapter A) to finalize the concept. The most essential part of the week working together was to decide very precisely which final product we were going to build, collect the requirements for building it and determine its concrete components. Because of our different backgrounds, the HPI students focused more on the software components of our final product and the Stanford students on the hardware components throughout the whole process up to completion. In all phases up to completion, we reflected our progress and made necessary changes according to our gained insights during testing. Nevertheless, each team built a complete Milestone with the working software to get feedback from testing and to get an overall impression of the system.
During the final travel of the HPI team to Stanford, we completed our product and built several Milestones. In addition, we finalized our branding (such as our logo and overall design) and created the presentation, brochure and posters.

8.2 Distributed Team Management & Communication

Working in a global team opens many possibilities because various cultural background and different skills may be included. The two teams that worked together for this project were located in two different time zones separated by nine hours and roughly 9,000 kilometers. Besides the global kick-off early November and the CEP we worked on together face-to-face, other work was performed team-wise and shared. This slightly changed for the last quarter after agreeing on a final concept as we increased the frequency of our exchange. For our internal communication and team organization, we used the following tools:

8.2.1 Sharing Documents: Google Drive

Any project related work was collected in a team-drive hosted by Google Drive. Each folder contained different content and files, such as notes, handouts, photos or videos from the work conducted so far. Many documents were created using Google Docs to store information from different Brainstorming sessions and to keep track of our progress. In addition, we primarily used the Docs and Slideshow applications to create the deliverables, such as handouts and presentations.

8.2.2 Sharing Code: Git and GitHub

In order to work on the source code simultaneously, we decided to use a Git repository hosted on GitHub[1]. Git[2] its version control system mainly used for tracking changes in source code that supports easy collaboration. We used one private repository to check in the code for the Milestones and for the app. In addition, we relied on the git repository to deploy code changes to the Milestones.

[1] GitHub: Popular Git server - https://www.github.com
[2] https://git-scm.com

8.2.3 Scientific Writing: ShareLaTeX

Instead of using Google Docs, we created the documentations with ShareLaTeX[3], an free online LaTeXeditor. LaTeXis a typesetting software and document preparation system using a markup language to describe a document [16]. This allowed us to easily meet the formal requirements for the ME310 course by using the specified LaTeXtemplate and still work together on the same document.

8.2.4 Chat: Facebook Messenger

During the global kick-off and the final travel for the EXPE in Stanford, we used the Facebook Messenger[4] group to stay in touch. It was the easiest way to connect and a good solution for the HPI team's stay in Stanford. Due to message flow in one single chat room without the possibility to concentrate on a specific topic, we switched to Slack (see subsection 8.2.5) for other project-related communication.

8.2.5 Chat: Slack

The main communication was done through the team messenger service Slack[5]. This allowed a more agile communication than writing emails and thus it was one of our main communication tools. Besides local communication to make appointments or to contact the fellow students, we mainly shared some important information within the global team outside of our weekly video call. We used Slack both, asynchronously and synchronously, depending on the time of the day and the type of question.

8.2.6 Video conferencing: appear.in

Besides using Slack as a written chat service for global and local team communication, we used the video conferencing tool appear.in[6] to have weekly video calls. The conference mainly served as a way to mutually synchronize with respect to the latest findings, prototypes and the next steps to take. In contrast to Skype[7], appear.in does not require any registration or the installation of a software as it completely runs in a web browser. In addition to video calls within the global team, we also used the appear.in service to stay in touch with our liaison and to get feedback on our ideas.

8.3 Budget

Each team had a budget to be used in context of the nine-month long project, especially for prototyping materials. The Stanford team had a budget of up to $8,000 while the HPI team could spend up to €5,000. In total, both teams together spent $8,214.37 during the whole project.
The following Table 8.2 gives an overview of how the Stanford team used the budget during the course. The Stanford team spent $6,816.35 and most of the budget was used in the last quarter to build several Milestones. Furthermore, the Stanford team payed for most expenses related to the EXPE (e.g. for the posters and the booth supplies). In addition, the expenses include several

[3] ShareLaTeX: Free online editor for LaTeX - https://sharelatex.com
[4] Facebook Messenger - https://messenger.com
[5] Slack: A team messenger - https://slack.com
[6] appear.in - one click video conversations - https://appear.in
[7] Skype: An instant messaging service with video support - https://skype.com

devices for prototyping or Benchmarking, such as a Raspberry Pi, Arduino or an Amazon Echo Dot.

Date	**Project Period**	**Description**	**Purpose**	**Cost**
10/31/2017	Benchmarking	Conference: Human AI Collaboration: A Dynamic Frontier	Learning about the field of AI and the state of current research	$15.00
11/11/2017	Benchmarking	Amazon Echo Dot	Benchmarking	$54.49
11/27/2017	Decoration	6 posters	Team space decoration	$72.76
11/28/2017	Prototyping	Paper, Binders, and folders	Prototyping organization	$49.87
01/10/2018	SUDS Catering	Ike's Place	Team Duty for SUDS	$290.30
01/22/2018	Prototyping	Plasma Ball, Baby monitor and Tent	Dark Horse Prototype	$159.63
02/10/2018	Prototyping	Raspberry Pi and Electronic Components	Funky System Prototype	$209.91
02/21/2018	Prototyping	Smart Robot Kit, Lights and accessories for the arduino car	Functional System Prototype	$272.92
06/05/2018	Manufacturing	3D Printing	Final Prototype	$1897.17
06/02/2018	Presentation	Booth Supplies	EXPE Set Up	$859.84
06/05/2018	Presentation	Posters	EXPE Set Up	$980.07
06/05/2018	Manufacturing	Electronics	Prototyping and Final Prototype	$1,389.46

Table 8.2: Budget used by the Stanford team

The HPI team spent €1,165.02 for the entire project period (see Table 8.3 for more details), which is about $1,398.02 (using a fixed conversion rate of $1.20 = €1). Most of the budget has been spent to build a Milestone and to get a Smart Lights to integrate to the home system. This Milestone has been used to create the software and to test the user experience with parents and children. For EXPE, the HPI team bought six T-shirts with our logo and design. A small portion of the budget has been used for various incentives as small presents for long and informative interviews or for intensive time spent with the prototype. They have been given after the corresponding interview or testing session to express our gratitude to participants and keep them in the loop for future tests. In addition, we bought a book to learn more about communication models and concepts. Most prototypes were either software-based or the required material was already present in the HPI work space. During winter quarter, the most expenses were connected with to the winter presentation in the south of Germany. There, we presented our current progress and got a lot of feedback and some new ideas which we considered while working on our final concept.

Date	**Project Period**	**Description**	**Purpose**	**Cost**
10/23/2017	Needfinding	Incentive	Presents for interview partners	€10.90
12/04/2017	Benchmarking	Book	Communication Models for the Study of Mass Communications	€22.14
12/06/2017	Prototyping	Incentive	Presents for interview partners	€15.89
02/15/2018	Needfinding	Incentive	Presents for interview partners	€5.64
02/21/2018	Needfinding	Taxi ride	Getting to the army barracks for our interview	€11.80
03/09/2018	Presentation	Train ticket	One-way ticket for Simon to Karlsruhe for the Winter presentation	€34.40
03/10/2018	Presentation	Train ticket	One-way ticket for Sebastian to Karlsruhe for the Winter presentation	€44.90
03/11/2018	Presentation	Train ticket	One-way ticket for Winfried to Karlsruhe for the Winter presentation	€29.90
03/11/2018	Presentation	Hotel	Accommodation for the Winter presentation in Karlsruhe	€188.00
03/13/2018	Presentation	Train ticket	Return ticket to Berlin from the Winter Presentation	€81.95
03/29/2018	Convergence	Team Dinner	Team Dinner during Stanford's visit at HPI	€114.10
03/30/2018	Prototyping	Electrical Components	Accessories for the Raspberry Pi	€65.92
04/01/2018	Prototyping	Electrical Components	Different Raspberry Pis and accessories	€91.52
04/19/2018	Testing	Incentive	Presents for interview partners	€7.53
05/11/2018	Prototyping	Electrical Components	Raspberry Pi and accessories for building a milestone	€52.50
05/14/2018	Prototyping	Philips Hue	Bridge and LED strip to integrate as smart home system	€79.95
05/18/2018	Prototyping	Pixel ring	Original Adafruit Pixelring for the milestone	€24.95
05/31/2018	Prototyping	Electrical Components	Various Power and USB adapters for the EXPE	€39.98
06/02/2018	Presentation	T-Shirts	Six customized T-Shirts for the EXPE	€243.05

Table 8.3: Budget used by the HPI team

8.4 Team Profiles and Reflections

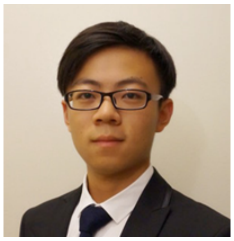
So far, this documentation reflects our global team: We worked together on all parts of this documentation including the various aspects of our work from the previous months. The following section is different, as every team member describes his or her own thoughts on the project and individual reflections.

8.4.1 Corporate Liaison

- **Jianming Dong**
  Chief User Experience (UX) Architect and Senior Design Director
  FutureWei Technologies, Inc.
  dongjianming@huawei.com

- **Jennifer (Wangjing) Wong**
  Director R&D
  FutureWei Technologies, Inc.
  jennifer.wangjing@huawei.com

8.4.2 Stanford Team



Lucas Zhiheng Zhou
Status: ME. Graduate Student
Contact: zzh@stanford.edu
Skills: SOLIDWORKS, MATLAB, C++, Arduino, Microsoft Office, DIAdem, manufacturing
Language: English and Mandarin

I am from China, I attended University of Michigan as an undergraduate for my degrees in Mechanical Engineering and Materials Science Engineering. My interests include material science, product design, world geography, transportation system study. I interned at Dow Chemical as improve engineer and ZF TRW as product engineer during the past year.

Reflection: We were lucky to be assigned to this cutting edge project topic, that is Artificial Intelligence. Last quarter, I was frustrated with the big project prompt. Unlike my previous design and manufacturing class that gives a very specific prompt and sometimes even give a potential solution, this class wants us to define the problem, create the user group and explore all kinds of solutions before moving to the step of design. This is really a learning and unlearning process. With the design methodology taught in ME310, my Design Thinking is changed to: (know the topic) find potential user needs, benchmarking, narrow down users, build simple prototype to test functionality, finalize the design idea and start to build fully functional prototype. The experience during the Autumn quarter gave me a fresh view of Design Thinking, especially for the design challenge without constraint solutions, and I believe this new Design Thinking will generate a lot more interesting ideas and concepts. This quarter, we explored our design space in psychological care and children care. All the activities including needfinding, interview and prototyping gave us an insight into make a product for young children. We found a solid need in daily communication between parents and their children. I can really feel that need because I have the experience of babysitting my 16 year younger sister when she was at 3-5 years old. We have successfully built a working prototype and demonstrate the functionality of our system during the EXPE. It was a very precious learning experience for me in the area of design thinking and building a portable or consumer electronic device.

Felipe Cabral
Status: M.E. Graduate Student
Contact: fcabral@stanford.edu
Skills: SOLIDWORKS, Microsoft Office, MATLAB, C++, Arduino, experience with machine shop tools
Languages: English and Portuguese

I'm originally from São Paulo in Brazil and came to the USA for my Undergraduate degree, which I completed at Stanford in Mechanical Engineering. I enjoy swimming, basketball, performing, teaching and watching improv and movies.

Reflection: This project has been a huge learning experience. Tackling a broad prompt and working together in a team across two continents proved to be quite a challenging albeit rewarding experience. Winter quarter was particularly challenging to me since we spent a long time in uncertainty of who our user was. I felt like our prototypes were going in random directions and weren't necessarily the most productive activities. Once we took a step back and did Needfinding with a more specific user group in mind, our direction started to become a lot clearer and with that we managed to land on a compelling need that propelled our subsequent prototyping and design activities in a much more meaningful way. I am very pleased with the final direction we pursued and our realization of our solution. It was a pleasure to work with my teammates and to be challenged by our teaching team every step of the way. I am very grateful and proud of the work we have accomplished in the past few months.

Simran Shah
Status: M.E. Graduate Student
Contact: simshah@stanford.edu
Skills: Experience in AutoCAD, DS Catia, ANSYS, MATLAB, Microsoft Office, Adobe Photoshop, Adobe Illustrator
Languages: English, Hindi, Gujarati, Marathi

I was born and raised in Mumbai, India. I earned my B.E. (Hons.) in Mechanical Engineering at Birla Institute of Technology and Science, Pilani, Dubai. I enjoy art, solving puzzles, listening to music and travelling.

Reflection: The past year has been a great learning experience. Personally, I had a very different perspective of Mechanical design and had a different expectation of the skills required to excel in this area. This project was my first experience with the Stanford school of thought for Design Thinking and it was a great opportunity to dive right in and have a long term hands on experience of the process from ideation to realisation. This project was extremely important to me in helping me realize my own strengths and weaknesses. It helped me become more aware of what I bring to a team as well as the skills that I would like to develop and reach a higher level of proficiency in. I find that the Design Thinking philosophy is a part of who I am now and has hugely impacted the way I approach things in my personal life as well. The project itself went through a series of ups and downs, unexpected road bumps and large amounts of ambiguity. The experience taught me different ways to deal with various situations and more importantly that things find a way of coming together eventually. I am extremely happy with the end result and the positive feedback we received at EXPE definitely made the months of effort worth it. I feel fortunate and grateful for having had this opportunity and also for being part of a team that found a way to work with each person's strengths and allowed each of our individual identities to find a space in the project.

8.4.3 HPI Team

Simon Krogmann
Status: C.S. Graduate Student
Contact: Simon.Krogmann@student.hpi.de
Skills: Python, C++, LaTeX, OpenGL, algorithm engineering
Languages: English, German, basic Spanish

I grew up in Berlin, Germany, after which I moved to Potsdam, Germany. There, I earned my B.Sc. in IT-Systems Engineering from Hasso Plattner Institute. In my free time I travel and do a wide variety of sports including volleyball and bouldering.

Reflection: I feel that over the last year we built a great product together. Although in the beginning, I had some problems with our vague project prompt, I think that we managed to converge on a single need quite well. There are still things to make VAMO into a final product, but I am satisfied with handing over this project to Futurewei and hope to see it developed further. But even if it will not get picked up in the future I learned a lot about Design Thinking, about teamwork, and about some technical aspects of our product. In a lot of cases I really had to leave my comfort zone and did things, I never would have imagined I would ever do, like interviewing such a wide variety of people, sometimes in very unusual places.
It was also incredibly enlightening seeing the variety of products built by the other teams and also the projects created in other courses at Stanford. When talking to their creators about their design process, the problems they had, and how they overcame them I learned a lot to hopefully apply to my future projects.

Winfried Lötzsch
Status: C.S. Graduate Student
Contact: Winfried.Loetzsch@student.hpi.de
Skills: C#, SQL, Python, PHP, Git, LaTeX, Tensorflow
Languages: English, German

I was born in the ore mountains, graduated in Computer Science at the university of Chemnitz in fall 2017 and moved to Berlin afterwards to start the Master IT-Systems Engineering in Potsdam. For the last three years, I worked as a student employee at the IT department of the airbag manufacturer Takata in Saxony. In my spare time I like to make music (piano) and go swimming.

Reflection: During the last nine months, we explored a wide variety of different problems and explored the Design Space in many possible directions. Only few months ago, we decided on the final User Group for our project and still later, we finally decided on the product to build. While there was no guarantee, that the long exploration in different directions we did before making these decisions would lead to a valuable outcome, I am very happy with our final solution. The final product addresses a real need, that exists today and also exists for a large range of people. I also like the fact, that I learned a lot about how teams function locally from my experience with the HPI team, but I also learned a lot about how to work together over a large distance. The collaboration over a long time span with people from another country was new to be and is in my opinion a very unique experience. As I was new to Design Thinking when we started nine months ago, I am now excited to know that the Design Thinking framework in general, but also the work we invested to explore the Design Space and the prototypes we implemented and completely abandoned later, finally led to the steps we took towards designing our final solution. The final testing at EXPE also showed me once again, that there is an infinite amount of ideas left to explore, and although we were able to present a fully functioning product, it can only be seen as a first prototype towards the goal to build a consumer ready product. At this point in time I feel more confident of my own creativity and am eager to continue the journey of Design Thinking by participating in new Design projects.

Sebastian Serth
Status: C.S. Graduate Student
Contact: Sebastian.Serth@student.hpi.de
Skills: Agile Software Development with Ruby, Scala and Python.
Experience with Git, SQL, LaTeX, Microsoft Office
Languages: English, German, French

After I graduated in IT-Systems Engineering from Hasso Plattner Institute, Potsdam, Germany, I did two technical internships before returning to university: One at SAP Labs, Sofia, Bulgaria and another at Lexmark International, Sydney, Australia. In my spare time, I enjoy meeting friends, jumping the trampoline and travelling.

Reflection: In my opinion, we had an exciting journey during the nine-month project and I am happy with the final outcome. Especially during EXPE, we got a lot of positive feedback confirming that it was worth putting that much effort into the project. After the visit of our global partner team in March, I struggled with defining the concrete concept even though I saw the parent's need we agreed on. Therefore, I was happy about the additional information we gathered from the interviews, which allowed us to shape the final concept. In order to deal with the uncertainty in the meantime, I had to left my comfort zone many times. On these occasions, I was able to learned a lot about project management in a global team and got more confident in the methods of Design Thinking. Even though the implementation and realization phase felt more similar to software engineering project, I was thrilled to include so many (new) technologies in a project of that size. As I started to work on a topic I felt familiar with, the interplay of different parts paved the way for me to topics with which I had had little contact so far. During our final travel and in preparation to the EXPE, I also got practical insights to mechanical engineering and different types of 3D printing, which I only knew from lectures so far. Overall, I am satisfied with my learnings, the project outcome and I look forward to seeing how FutureWei will benefit from our result.

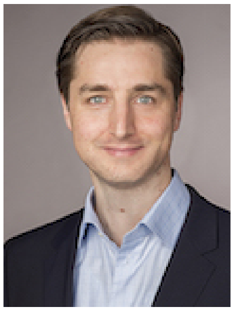
8.4.4 Stanford Coaches

Larry Leifer
Status: Professor, Mechanical Engineering
Contact: leifer@cdr.stanford.edu

Mark Cutkosky
Status: Professor, Mechanical Engineering
Contact: cutkosky@stanford.edu

George Toye
Status: Consulting Professor, Mechanical Engineering
Contact: deputy@me310.stanford.edu

8.4.5 HPI Coaches



Matthias Uflacker
Status: ME310 Programm Director
Contact: Matthias.Uflacker@hpi.de

Keven Richly
Status: ME310 Programm Manager
Contact: Keven.Richly@hpi.de

Christopher Schmidt
Status: ME310 Programm Manager
Contact: Christopher.Schmidt@hpi.de

8.4.6 Expert Coaches

Li Jiang
Status: PhD, Lecturer of Robotics, AI and Design of Future Education, Stanford
Contact: lijiang@stanford.edu

Dylan James Moore
Status: ME310 alumnus & PhD Candidate, Mechanical Engineering (Design), Stanford
Contact: djmoore3@stanford.edu

8.4.7 HPI Teaching Assistant

Georg Berecz
Status: ME310 alumnus, C.S. Graduate Student & Dedicated TA for our Team
Contact: Georg.Berecz@student.hpi.de

Bibliography

[1] Apple, Inc. The future is here: iPhone X - Apple. https://www.apple.com/newsroom/2017/09/the-future-is-here-iphone-x/, 2017.

[2] Andy Boxall. Huawei Kirin 970: Everything you need to know. *Digital Trends*, 09 2017.

[3] Tim Brown. *Change by design*. Collins Business., 2009.

[4] T Carleton and L Leifer. Stanford's ME310 course as an evolution of engineering design. In *Proceedings of the 19th CIRP Design Conference–Competitive Design*. Cranfield University Press, 2009.

[5] Aaron Chalfin, Oren Danieli, Andrew Hillis, Zubin Jelveh, Michael Luca, Jens Ludwig, and Sendhil Mullainathan. Productivity and selection of human capital with machine learning. *American Economic Review*, 106(5):124–27, 2016.

[6] S. Cheshire and M. Krochmal. DNS-Based Service Discovery. RFC 6763 (Proposed Standard), February 2013. https://www.rfc-editor.org/rfc/rfc6763.txt.

[7] S. Cheshire and M. Krochmal. Multicast DNS. RFC 6762 (Proposed Standard), February 2013. https://www.rfc-editor.org/rfc/rfc6762.txt.

[8] Mostafa M. El-Bermawy. Your Filter Bubble is Destroying Democracy. *Wired*, 2016.

[9] Glassdoor, Inc. Futurewei Technologies Overview. https://www.glassdoor.com/Overview/Working-at-Futurewei-Technologies-EI_IE409316.11,33.htm, 2017.

[10] Robert Hogan, Joyce Hogan, and Brent W Roberts. Personality measurement and employment decisions: Questions and answers. *American psychologist*, 51(5):469, 1996.

[11] Huawei Investment & Holding Co., Ltd. Huawei Investment & Holding Co., Ltd. 2016 Annual Report. http://www.huawei.com/en/about-huawei/annual-report/2016, 2016.

[12] L. Crystal Jiang. Absence makes the communication grow fonder: Geographic separation, interpersonal media, and intimacy in dating relationships, 2013. http://onlinelibrary.wiley.com/doi/10.1111/jcom.12029/full.

[13] David Kelley and Tom Kelley. *Creative confidence: Unleashing the creative potential within us all*. Crown Pub, 2013.

[14] Annica Kristoffersson, Silvia Coradeschi, and Amy Loutfi. A review of mobile robotic telepresence. *Advances in Human-Computer Interaction*, 2013:3, 2013.

[15] Vijay Kumar. *101 design methods: A structured approach for driving innovation in your organization*. John Wiley & Sons, 2012.

[16] Leslie Lamport and D Bibby. Latex, A Document Preparation System, 1986.

[17] Gideon Lewis-Kraus. The Great A.I. Awakening. *The New York Times*, 2017.

[18] Stratasys Ltd. Stratasys legal information. http://www.stratasys.com/legal/legal-information.

[19] Stratasys Ltd. Stratasys polyjet technology. http://www.stratasys.com/polyjet-technology.

[20] Andrea Löbbecke. Die vielen fallen einer fernbeziehung, 2012. https://www.welt.de/gesundheit/psychologie/article111363645/Die-vielen-Fallen-einer-Fernbeziehung.html.

[21] Catherine Matacic. Google's new translation software is powered by brainlike artificial intelligence. *Science*, 09 2016.

[22] Rim Razzouk and Valerie Shute. What is design thinking and why is it important? *Review of Educational Research*, 82(3):330–348, 2012.

[23] Matthew R Sanders. Triple p-positive parenting program: Towards an empirically validated multilevel parenting and family support strategy for the prevention of behavior and emotional problems in children. *Clinical child and family psychology review*, 2(2):71–90, 1999.

[24] Rafia Shaikh. After Pixels, It's Time for Google-Branded Earbuds! Potentially Transform the Travel Industry with Realtime Translation [How It Works]. *WCCF Tech*, 10 2017.

[25] Natasha Singer. In a Mood? Call Center Agents Can Tell. *The New York Times*, 2013.

A Appendix

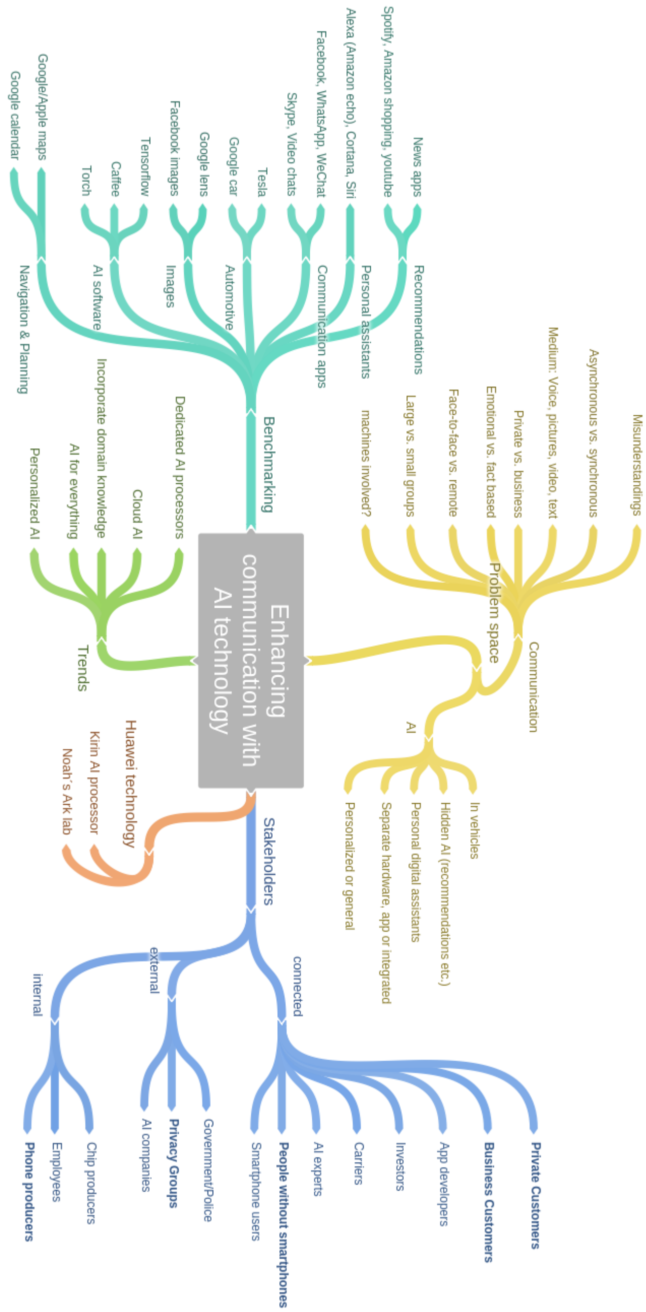


Figure A.1: The Design Space Map summarizes important design directions.

```
#define LT_R !digitalRead(10)
#define LT_M !digitalRead(4)
#define LT_L !digitalRead(2)

#define ENA 5
#define ENB 6
#define IN1 7
#define IN2 8
#define IN3 9
#define IN4 11

#define carSpeed 140

const int ledPin =  13;
const int ledPin2 =  1;
int ledState = LOW;
int ledState2 = LOW;
unsigned long previousMillis = 0;
const long interval = 200;


void forward(){
  analogWrite(ENA, carSpeed);
  analogWrite(ENB, carSpeed);
  digitalWrite(IN1, HIGH);
  digitalWrite(IN2, LOW);
  digitalWrite(IN3, LOW);
  digitalWrite(IN4, HIGH);
  Serial.println("go forward!");
}

void back(){
  analogWrite(ENA, carSpeed);
  analogWrite(ENB, carSpeed);
  digitalWrite(IN1, LOW);
  digitalWrite(IN2, HIGH);
  digitalWrite(IN3, HIGH);
  digitalWrite(IN4, LOW);
  Serial.println("go back!");
}

void left(){
  analogWrite(ENA, carSpeed);
  analogWrite(ENB, carSpeed);
  digitalWrite(IN1, LOW);
  digitalWrite(IN2, HIGH);
```

```
  digitalWrite(IN3, LOW);
  digitalWrite(IN4, HIGH);
  Serial.println("go left!");
}

void right(){
  analogWrite(ENA, carSpeed);
  analogWrite(ENB, carSpeed);
  digitalWrite(IN1, HIGH);
  digitalWrite(IN2, LOW);
  digitalWrite(IN3, HIGH);
  digitalWrite(IN4, LOW);
  Serial.println("go right!");
}

void stop(){
   digitalWrite(ENA, LOW);
   digitalWrite(ENB, LOW);
   Serial.println("Stop!");
}

void setup(){
  Serial.begin(9600);
  pinMode(LT_R,INPUT);
  pinMode(LT_M,INPUT);
  pinMode(LT_L,INPUT);
    pinMode(ledPin, OUTPUT);
}

void loop() {
  if(LT_M){
    forward();
  }
  else if(LT_R) {
    right();
    while(LT_R);
  }
  else if(LT_L) {
    left();
    while(LT_L);
  }
//  else if (!LT_L && !LT_R) {
//    stop();
//  }

  unsigned long currentMillis = millis();
```

```
  if (currentMillis - previousMillis >= interval) {

    previousMillis = currentMillis;
    if (ledState == LOW) {
      ledState = HIGH;
      //ledState2 = LOW;
    } else {
      ledState = LOW;
      //ledState2 = HIGH;
    }
    digitalWrite(ledPin, ledState);
    //digitalWrite(ledPin2, ledState2);
  }

}
```

Source Code A.1: Arduino Code for the RC Car Functional System Prototype

Exploration Directions	*Human remote communication*	*Popular AI Apps*	*Personal assistants on smartphones*	*Personal assistants on separate hardware*
Examples	• Telephone • Skype • WhatsApp • WeChat • Facebook	• Spotify • Snapchat • BBC News • YouTube	• Apple Siri • Microsoft Cortana • Google Assistant • Amazon Alexa • Samsung Bixby	• Amazon Echo • Google Home • Apple HomePod
Experience	• reduced quality of communication because of time delay, cellular service and language • a response is sometimes not guaranteed	• users like emotional and personalized apps • strong direct communication preference for emotional situations • fear of "Filter Bubbles"	• many misunderstandings • only for simple questions • tight bound to one vendor • people feel uncomfortable to talk to Siri • organizational help from AI (calendar...) welcomed	• smart home hub • always present / available, but many privacy concerns • personal assistant as addition to a speaker • similar problems as smartphone assistants
Extreme Users	• ATC with pilots • Call center • Secretary • CEOs, lawyer • Remote workers • Long-distance relationship • Phone call for disruptive life events	• Phone-addict	• business users allowing hands-free interaction, e.g. while driving a car	• Smart home residents • Rachel Neasham, travel agent. Trains a travel agent AI and works with it to help customers (Centaur relationship)
Results / Opportunity	• have some system to facilitate or bridge the communication	• build apps that make users feel comfortable • the physical distance might help to speak out problems	• put efforts to make voice assistant more human like	

Table A.1: Benchmarking of AI-powered apps and solutions

Date	Time	Activity	Room	Address	Participants
Monday, 26th March 2018	09:00	Breakfast		Tegel Airport, Zufahrt zum Flughafen Tegel, 13405 Berlin, Germany	Stanford Team
Monday, 26th March 2018	11:00	Campus Tour		Hasso Plattner Insitute, Prof.-Dr.-Helmert-Str. 2-3, 14482 Potsdam, Germany	Team FutureWei
Monday, 26th March 2018	12:30	Lunch		Ulf's Café im HPI, Prof.-Dr.-Helmert-Str. 2-3, 14482 Potsdam, Germany	Team FutureWei
Monday, 26th March 2018	13:30	Design team flex time	E-U.1-1	Hasso Plattner Insitute, August-Bebel-Str. 88, 14482 Potsdam, Germany	Team FutureWei
Monday, 26th March 2018	17:00	Design team flex time	E-U.1-1	Hasso Plattner Insitute, August-Bebel-Str. 88, 14482 Potsdam, Germany	Team FutureWei
Monday, 26th March 2018	19:00	Dinner			Team FutureWei
Monday, 26th March 2018	22:00	Return to Hostel		Beethovenweg 27, 14532 Kleinmachnow, Germany	Stanford Team
Tuesday, 27th March 2018	09:00	Breakfast		Beethovenweg 27, 14532 Kleinmachnow, Germany	Stanford Team
Tuesday, 27th March 2018	10:00	Meeting to discuss convergence	E-U.1-1	Hasso Plattner Insitute, August-Bebel-Str. 88, 14482 Potsdam, Germany	Team FutureWei
Tuesday, 27th March 2018	12:30	Lunch		Ulf's Café im HPI, Prof.-Dr.-Helmert-Str. 2-3, 14482 Potsdam, Germany	Team FutureWei
Tuesday, 27th March 2018	14:00	SGM	E-U.1-2	Hasso Plattner Insitute, August-Bebel-Str. 88, 14482 Potsdam, Germany	Team FutureWei, TA Georg Berecz, ME310 Program Managers Keven Richly and Christopher Schmidt
Tuesday, 27th March 2018	12:30	"I like / I wish" feedback session	E-U.1-1	Hasso Plattner Insitute, August-Bebel-Str. 88, 14482 Potsdam, Germany	Team FutureWei
Tuesday, 27th March 2018	14:30	Prototyping Session	E-U.1-1	Hasso Plattner Insitute, August-Bebel-Str. 88, 14482 Potsdam, Germany	Team FutureWei
Tuesday, 27th March 2018	19:00	Dinner @ Restaurant		Restaurant Chi Keng, Luisenpl. 3, 14471 Potsdam	Team FutureWei
Tuesday, 27th March 2018	22:00	Return to Hostel		Beethovenweg 27, 14532 Kleinmachnow, Germany	Stanford Team
Wednesday, 28th March 2018	09:00	Breakfast		Beethovenweg 27, 14532 Kleinmachnow, Germany	Stanford Team
Wednesday, 28th March 2018	10:00	User Testing	E-U.1-1	Hasso Plattner Insitute, August-Bebel-Str. 88, 14482 Potsdam, Germany	Team FutureWei
Wednesday, 28th March 2018	12:30	Lunch		Ulf's Café im HPI, Prof.-Dr.-Helmert-Str. 2-3, 14482 Potsdam, Germany	Team FutureWei
Wednesday, 28th March 2018	13:30	Final convergence	E-U.1-1	Hasso Plattner Insitute, August-Bebel-Str. 88, 14482 Potsdam, Germany	Team FutureWei
Wednesday, 28th March 2018	19:00	Dinner			Team FutureWei
Wednesday, 28th March 2018	22:00	Return to Hostel		Beethovenweg 27, 14532 Kleinmachnow, Germany	Stanford Team
Thursday, 29th March 2018	09:00	Breakfast		Beethovenweg 27, 14532 Kleinmachnow, Germany	Stanford Team
Thursday, 29th March 2018	10:00	Design team flex time	E-U.1-1	Hasso Plattner Insitute, August-Bebel-Str. 88, 14482 Potsdam, Germany	Team FutureWei
Thursday, 29th March 2018	12:30	Lunch		Ulf's Café im HPI, Prof.-Dr.-Helmert-Str. 2-3, 14482 Potsdam, Germany	Team FutureWei
Thursday, 29th March 2018	15:15	LGM preparation	E-U.1-2	Hasso Plattner Insitute, August-Bebel-Str. 88, 14482 Potsdam, Germany	Team FutureWei
Thursday, 29th March 2018	16:00	LGM	E-U.1-1	Hasso Plattner Insitute, August-Bebel-Str. 88, 14482 Potsdam, Germany	All Teams (FutureWei, Hella, HUK), all TAs (Georg Berecz, Jan Selke, Tamara Slosarek) and ME310 Program Managers Keven Richly and Christopher Schmidt
Thursday, 29th March 2018	18:00	SUDS dinner	E-U.1-2	Hasso Plattner Insitute, August-Bebel-Str. 88, 14482 Potsdam, Germany	All Teams (FutureWei, Hella, HUK), all TAs (Georg Berecz, Jan Selke, Tamara Slosarek) and ME310 Program Managers Keven Richly and Christopher Schmidt
Thursday, 29th March 2018	22:00	Return to Hostel		Beethovenweg 27, 14532 Kleinmachnow, Germany	Stanford Team
Friday, 30th March 2018	09:00	Breakfast		Beethovenweg 27, 14532 Kleinmachnow, Germany	Stanford Team
Friday, 30th March 2018	10:00	Design team flex time			Team FutureWei
Friday, 30th March 2018	12:30	Lunch			Team FutureWei
Friday, 30th March 2018	13:30	Future Direction & Summary			Team FutureWei
Friday, 30th March 2018	19:30	Depart from TXL		Tegel Airport, Zufahrt zum Flughafen Tegel, 13405 Berlin, Germany	Stanford Team

Figure A.2: Agenda for Stanford Visit at HPI

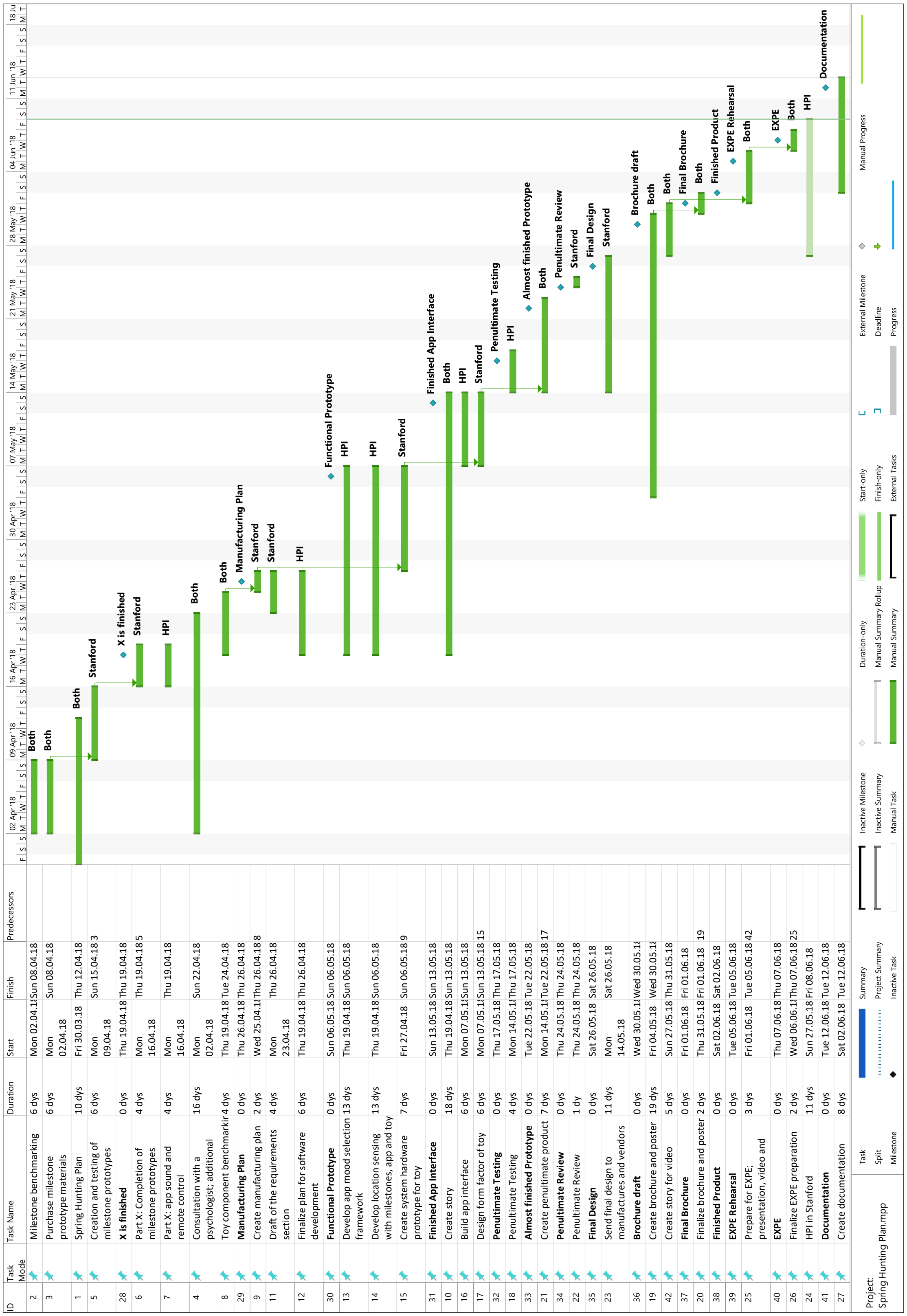


Figure A.3: Spring Hunting Plan

```
import Cocoa
import CoreBluetooth

let lineFollowerCBUUID = CBUUID(string: "0xFFF0")
//let lineFollowerCBUUID = CBUUID(string:
"0000ffe1-0000-1000-8000-00805f9b34fb")
let LineFollowerService = CBUUID(string: "0xFFF0")
let LineFollowerCharacteristic = CBUUID(string: "0xFFE1")


class ViewController: NSViewController {

    var centralManager: CBCentralManager!
    var lineFollowerPeripheral: CBPeripheral!

    override func viewDidLoad() {
        super.viewDidLoad()

        centralManager = CBCentralManager(delegate: self, queue: nil)
        // Do any additional setup after loading the view.
    }

    override var representedObject: Any? {
        didSet {
        // Update the view, if already loaded.
        }
    }


}

extension ViewController: CBCentralManagerDelegate{
    func centralManagerDidUpdateState(_ central: CBCentralManager) {
        switch central.state {
        case .unknown:
            print("central.state is .unknown")
        case .resetting:
            print("central.state is .resetting")
        case .unsupported:
            print("central.state is .unsupported")
```

Figure A.4: First page of the Swift Code for the RC Car app

```
        case .unauthorized:
            print("central.state is .unauthorized")
        case .poweredOff:
            print("central.state is .poweredOff")
        case .poweredOn:
            print("central.state is .poweredOn")
            centralManager.scanForPeripherals(withServices: [lineFollowerCBUUID])
            //centralManager.scanForPeripherals(withServices: nil)
        }
    }

    func centralManager(_ central: CBCentralManager, didDiscover peripheral:
CBPeripheral, advertisementData: [String : Any], rssi RSSI: NSNumber) {
        print(peripheral)

        //if (peripheral.name == "HC-08"){
            lineFollowerPeripheral = peripheral
            lineFollowerPeripheral.delegate = self
            centralManager.stopScan()
            centralManager.connect(lineFollowerPeripheral)
        //}
    }
    func centralManager(_ central: CBCentralManager, didConnect peripheral:
CBPeripheral) {
        print("Connected!")
        lineFollowerPeripheral.discoverServices([LineFollowerService])
        //lineFollowerPeripheral.discoverServices(nil)
    }
}

extension ViewController: CBPeripheralDelegate{
    func peripheral(_ peripheral: CBPeripheral, didDiscoverServices error: Error?) {
        guard let services = peripheral.services else { return }

        for service in services {
            print(service)
            peripheral.discoverCharacteristics([LineFollowerCharacteristic], for:
service)
            //peripheral.discoverCharacteristics(nil, for: service)
        }
    }

    func peripheral(_ peripheral: CBPeripheral, didDiscoverCharacteristicsFor
```

Figure A.5: Second page of the Swift Code for the RC Car app

```
service: CBService, error: Error?) {
        guard let characteristics = service.characteristics else { return }

        for characteristic in characteristics {
            print(characteristic)

            if characteristic.properties.contains(.write) {
                print("\(characteristic.uuid): properties contains .write")

                let data = Data("1".utf8)
                print(data)
                peripheral.writeValue(data, for: characteristic, type:
CBCharacteristicWriteType.withResponse)
                print("wrote 1")

            }

        }
    }
    func peripheral(_ peripheral: CBPeripheral, didWriteValueFor characteristic:
CBCharacteristic, error: Error?) {
        if (error != nil){
            print("Error")
            return
        }
        print("Updated!")
    }

}
```

Figure A.6: Third page of the Swift Code for the RC Car app