University of Tartu

Faculty of Science and Technology

Institute of Technology

Farnaz Baksh

AN OPEN-SOURCE ROBOTIC STUDY COMPANION FOR UNIVERSITY STUDENTS

Master’s thesis (30 EAP)
Robotics and Computer Engineering

Supervisors:
Karl Kruusamäe, PhD
Sandra Schumann, MSc
Agnes Luhtaru, MSc

Tartu 2023

Abstract

An Open-Source Robotic Study Companion For University Students

This thesis presents an evidence-based approach and develops an affordable and effective Robotic Study Companion (RSC) prototype for university students. It addresses the lack of social robots tailored to higher education and the scarcity of open-source educational platforms. Through a comprehensive literature review on Human-Robot Interaction (HRI), social and companion robots, and Natural Language Processing (NLP) technologies, this work identifies trends and best practices for educational social robots.

The research systematically reviews select social robots, examining their applications, technical features, design, and human-centric interaction. It explores the human perspective of HRI, focusing on users in educational settings. Based on these insights, functional and non-functional requirements are established for the RSC, inspiring its design and development. The RSC prototype, built using off-the-shef components and leveraging OpenAI's large language models, demonstrates its potential to simplify complex concepts for students. The long-term goal is to enhance the RSC's design, durability, and commercial viability.

CERCS: T125 Automation, robotics, control engineering; S281 Computer-assisted education; P176 Artificial intelligence

Keywords: Robotic Study Companion (RSC), Social Robot, Natural Language Processing (NLP), University Students

Resümee

Avatud lähtekoodiga robotõpikaaslane üliõpilastele

Selles töös esitatakse tõenduspõhine lähenemisviis ja töötatakse välja taskukohane ja tõhus robotõpikaaslase prototüüp üliõpilastele. Töös käsitletakse kõrgharidusele kohandatud sotsiaalsete robotite puudumist ja avatud lähtekoodiga õppeplatvormide vähesust. Inimese ja roboti koostoime, sotsiaalsete robotite ja robotkaaslaste ning loomuliku keele töötlemise tehnoloogiaid käsitleva põhjaliku kirjandusülevaate põhjal määratakse selles töös kindlaks sotsiaalsete haridusrobotite suundumused ja parimad tavad.

Töös vaadeldakse süstemaatiliselt hulka sotsiaalseid roboteid, uurides nende rakendusi, tehnilisi omadusi, disaini ja inimsõbralikku suhtlust. Uuritakse inimese ja roboti koostoime inimperspektiivi, keskendudes kasutajatele hariduskeskkonnas. Nende arusaamade põhjal määratakse kindlaks robotõpikaaslase funktsionaalsed ja mittefunktsionaalsed nõuded, millel põhinevad selle disain ja arendus. Robotõpikaaslase prototüüp, mis on ehitatud kaubanduses olemasolevate komponentide abil ja OpenAI suurte keelemudelite abil, demonstreerib selle potentsiaali lihtsustada õpilaste jaoks keerulisi kontseptsioone. Pikaajaline eesmärk on parandada robotõpikaaslase disaini, vastupidavust ja kaubanduslikku tasuvust.

CERCS: T125 Automatiseerimine, robootika, juhtimistehnika; S281 Arvuti õpiprogrammide kasutamise metoodika ja pedagoogika; P176 Tehisintellekt

Märksõna: Robotõpikaaslane (RSC), sotsiaalne robot, loomuliku keele töötlus, üliõpilased

Contents

List of Figures

List of Tables

List of Abbreviations

RSC	Robot Study Companion
HRI	Human-Robot Interaction
AI	Artificial Intelligence
NLP	Natural Language Processing
AIY	Artificial Intelligence Yourself
RPi	Raspberry Pi
ER	Educational Robots
RO	Research Objective
LED	Light-Emitting Diode
LCD	Liquid Crystal Display
CAD	Computer-Aided Design
NLU	Natural Language Understanding
GPT	Generative Pre-trained Transformer
STT	Speech to Text
TTS	Text to Speech
PLA	Polylactic acid
GPIO	General Purpose Input/Output
PWM	Pulse Width Modulation
API	Application Programming Interface

1 Introduction

Anthropomorphism, the act of attributing human characteristics or emotions to non-living objects, has intrigued humans for many centuries. This concept is becoming increasingly relevant in the field of Human-Robot Interaction (HRI), particularly in the design of social robots [1]. In order to be considered social, these robots are often expected to exhibit human-like abilities such as voice communication, gestures, facial expressions, and eye contact, among others [2].

The use of social robots in education has gained significant attention in HRI research. Numerous robots have been developed for different educational purposes. There is evidence of promising results in language learning [3], [4], personalised instruction [5], [6], gamification [7], socialisation and collaboration [8] [9], and learning support for students with special needs or learning difficulties [10]–[12].

Although most research on social robots in education has focused on children and young adults (ages younger than 19), there are relatively few studies on how these robots can benefit learners and teachers in higher education [13]. Nevertheless, social robots have the potential to support self-motivation and focused attention in university settings where individual support from lecturers may be limited and where personal tutoring is often unavailable. Physical robots have been shown to elicit greater engagement, enjoyment, and information retention compared to on-screen or video representations of the same robot [5], [6]. Furthermore, university students perceive added value in the presence of a physical robot and report increased motivation, attention, and concentration [6], [13].

Further research is necessary to effectively integrate social robots into university settings in a way that supports learners and teachers. Social robots offer the advantage of adapting the learning experience to each individual's needs [6], but previous research has focused on utilising pre-existing robots rather than exploring the technical features needed for university students. Moreover, existing social robots on the market can be quite expensive [1], [14], which poses a challenge to their widespread adoption in higher education. Nonetheless, social robots could potentially serve as active participants in the learning process, highlighting the need to explore their role in university settings.

In light of this, this thesis aims to explore social and companion robot design trends and provide an evidence-based approach to designing a robotic study companion (RSC) platform for university students. The goal is to design and build a prototype of the RSC and contribute to the open-source community. This work can potentially improve university students' learning experience and contribute to the field of HRI research, ultimately furthering our understanding of the role of robotics technology in education.

1.1 Problem Statement

The current lack of social robots specifically designed as study companions for university students, combined with a scarcity of open-source desktop platforms tailored to the educational needs of higher education learners, presents a significant gap in the field of human-robot interaction. This study aims to fill this gap by providing evidence-based guidelines for the development of a social robot study companion, with a focus on addressing the unique educational needs and preferences of university students.

1.2 Research Objectives

RO1. To conduct a comprehensive review of the existing literature on Human-Robot Interaction (HRI), social and companion robots, and Natural Language Processing (NLP) technologies in order to identify key trends and best practices in the development of social robots for educational purposes.

RO2. To develop a set of functional, non-functional, and technical requirements for a university's robotic study companion synthesising the literature review and key findings to inform the design and development of an RSC.

RO3. To design and build a functional prototype of a social robot study companion that incorporates NLP technologies, with a focus on providing interactive and meaningful conversations to engage university students in their studies.

RO4. To lay the groundwork for future research in the field of social robots for educational purposes in a university setting. With a focus on exploring the potential of the social robot companion in adapting to individual student learning styles and providing personalised feedback to university students.

2 Background

This section presents a systematic review of a select few social robots, focusing on their applications, technical features, design, and human-centric interaction. It examines commercially available robots and explores their interaction modalities, mass, size, embodiment, shell, and release year, drawing from the extensive list covered in [1]. In addition, the section delves into the human perspective of human-robot interaction (HRI), with a particular emphasis on users in the educational settings. Furthermore, NLP technologies are also explored, providing a brief overview of their applications, given their use as conversational tools in social robots [1], [15].

According to [2], robotic companions are defined as robots that can maintain interactivity and assistivity for a longer period of time. However, to guide the search for existing research on social robots designed for educational purposes, a definition of robotic study companions (RSC) has been formulated, inspired by the definition of social robots [16].

A robotic study companion (RSC) is a type of social robot designed to interact with [university] students in a human-centric way and to operate in academic environments alongside them. RSCs can take on various forms, from humanoid to animal-like or even abstract, but they all share the goal of engaging students in an interpersonal manner. This involves communicating and coordinating their behaviour with students through various modalities, such as verbal, nonverbal, or affective means, with the aim of providing personalised support, feedback, and motivation for students to enhance their academic performance and learning outcomes.

2.1 Social Robots in Education

In recent years, social robots have emerged as a promising tool for enhancing the learning processes in various educational contexts [3]–[14]. These robots (Figure 2.1) come in different shapes and sizes, including humanoid (Nao, Asimo), animal-like (Aibo, Zooo), abstract and biomimetic (iRobi Q, Rubi-6, Sphero, MOCCA, Tabot).

Figure 2.1: Examples of social robots used in education research [1], [15]

Most of these educational robots are designed specifically for classroom use with a focus on enhancing the quality of education [1], [15], although the majority of research has been conducted with a focus on children [13]. Moreover, as explored in the following section on university students' preferences [17], most of these robots do not meet their needs.

Nao, for example, is a small humanoid robot used to support learning in computer and science classes ranging from primary to university courses [18]. It has been praised for its ability to adapt to different learners and learning materials [18] and for its valuable role in enhancing learning processes [15], [18]. However, commercially available humanoid robots like Nao are expensive (Openbox price from Robotlab [19] is USD 6,900 as of May 2023) and impractical for most students to own.

To fully understand the potential impact of social robots in education, it is important to consider the perspectives of those who interact with them in the educational realm. In the following sub-chapters, the viewpoints of both teachers and university students will be explored to gain insights into the benefits and challenges of using social educational robots.

2.1.1 Teachers' Perspectives

Based on [1], [20], educational robots can provide information on specific topics, query learned lessons, give advice on the learning process, correct errors, or provide feedback on students' progress. They can serve as assistants to teachers, helping to arrange lessons, or as personal tutors that help students in editing tasks and promote their individual learning process.

According to a study of German teachers' attitudes towards educational robots [18], teachers preferred using robots for individual or small-group learning, with 85% favouring using a robot as a tutor. They expected education robots to create a motivating learning environment, serve as a source of information, assist with assessing and monitoring students' progress, and be particularly beneficial for underachieving students who require individual support. However, they had concerns about the extra workload involved in programming, maintaining, and monitoring the robots, as well as high acquisition costs and the limited capacity of schools to purchase expensive tools. Teachers also expressed a desire for easy access to educational robots in the classroom, with integrated voice control making them easy to handle.

Other studies from different countries [21]–[25] have explored teachers' attitudes towards educational robots in school and classroom settings. While many agree that robots can be engaging tools for all students, potential issues highlighted include the fairness of access to technology, disruption of other classroom activities [22], privacy and data anonymisation [23], and cross-cultural acceptance since different countries' education systems may vary [24]. However, some studies emphasised their research limitations, such as biases due to specific regions, data collected mainly from pre-university educators, and focusing solely on using humanoid robots in the given scenarios [25].

2.1.2 University Students’ Perspectives

Reich-Stiebert et al. [17] examine the preferences of university students in relation to the physical design, interaction modalities, and personality traits of educational robots. The study recruited 116 undergraduate students, aged between 17 and 40 years, from different study fields, who were asked to complete an online questionnaire featuring visual examples of robot characteristics.

The findings revealed that most students favoured a machine-like robot (54.3%) over a human-like or animal-like robot [17]. Additionally, the students preferred a medium-sized robot (~100 centimetres) with gender-neutral colours such as white. The importance of speech as a mode of interaction was evident as it scored the highest (98.3%) among all the interaction modes presented to the students. Other preferred modes of interaction for an ideal educational robot were touch screen, gestures and touch, all scoring above 50%. It should be noted that previous research has shown that children's design preferences for robots may be influenced by their prior experience with robots [26]. Children with prior robot experience preferred smaller machine-like robots, while those without prior experience preferred adult-sized human-like robots. This might also apply to university-level students.

The robot's ability to distinguish between persons by recognising human faces or using people's names was considered an important feature[17]. Furthermore, 71.4% of the students believed that the ideal educational robot should recognise human emotions and display positive emotions. Negative personality characteristics such as pessimism, annoyance, boredom, or demotivation were also considered unfavourable by the students. [17] highlight the need to consider different learning contexts, such as schools versus universities, and personal factors, such as age, gender, and personality, when designing educational robots. Specifically, the study emphasised the need to address pending issues concerning the robot's ability to motivate and adapt to individual learners, its physical traits, privacy and safety aspects, and ease of handling.

Personalised assistance from robot tutors has been found to improve task performance and human-robot interactions in university students [6]. The study found that participants who received personalised assistance completed grid-based logic puzzles faster than those who received non-personalized assistance [6]. Additionally, [13] explored using a social robot (Pepper by Softbank Robotics) for additional tutoring in a university course. The study found that students were interested in using the robot tutor for exam preparation and desired more time with the learning environment. The study also identified several possibilities for individualising the learning experience, such as providing detailed explanations for incorrect answers and adapting the robot's movements based on learners' preferences [13]. Both studies [6], [13] highlight the potential benefits of designing educational robots with personalised features to enhance student learning outcomes and suggest that integrating social robots as personalised tutors can offer promising benefits in university education.

2.2 Design and Interaction Modalities of Social Robots

Designing social robots is a complex and multifaceted process that involves careful consideration of physical, aesthetic, and functional aspects. This section explores notable examples of abstract and biomimetic robots [1] were explored, each with unique designs and interaction modalities. These robots have been developed for a range of purposes, from companionship [27], [28] to educational and therapeutic applications [29], [30]. Many have been created with a focus on design thinking to maximise engagement with users and encourage long-term interaction [27], [30], [31]. By examining these robots and their design principles, valuable insights can be gained into how robots can be optimised for successful human-robot interaction.

2.2.1 Jibo - Design and Features

Jibo, created in 2014, is a personal assistant-style desktop robot designed to emotionally connect with users through its sophisticated voice interaction, expressive user interface, and physical movements of human-like poses and gestures [27]. Its design sought to create a new non-threatening experience while avoiding the stigma associated with robots.

The robot has a cylindrical body of approximately 30 centimetres (cm) in height and 2 kilograms (kg) in weight, with a 6.3×11 cm touch screen at the centre of its face. The screen usually displays Jibo's eye, which moves along with Jibo's body and can also display text, images, and videos. Two cameras are above the screen, a microphone array is across Jibo's head for identifying the direction of sounds, and two speakers are located on each side of Jibo's head for speech or music [32]. Its unique feature is its ability to express itself with body poses and gestures using a torso and head. The poses are created by motor-controlled rotational sections that are hidden in the inner core, making its movements silent and precise without the complexity of visible joints. Jibo's software is based on the Robot Operating System (ROS) and was in several pilot studies, such as implemented for children as young as four years old, intending to administer language, literacy, and speech pathology assessments [32].

Figure 2.2: Jibo’s external and internals [27]

Jibo is no longer owned by its creator Cynthia Breazeal and her MIT Media Lab team but was acquired by NTT Disruption in 2019 who plans on continuing research in the field with a focus on healthcare and education[1], they are planning on building a virtual form of Jibo, one that lives on a smartphone and can be accessed anywhere at any time [33]. Overall, Jibo's design and implementation have demonstrated its potential as a companion and educational tool for children, with its unique features and sophisticated movements creating a more natural and emotional interaction than other robots.

2.2.2 Haru - Minimalist Emotive Design

Haru is a tabletop robot developed by Honda Robotics in 2018 for research purposes [1]. Its minimalist design aims to maximise emotional engagement with users and encourage long-term interaction. To balance human expectations, physical affordance, and robot functionality, they researched popular anthropomorphised animated characters to analyse features and expressive possibilities (Figure 2.3) for incorporation into the robot's design [34]. The robot's expressive capabilities are enhanced by animated eyes and a customised structure that houses an addressable LED matrix display.

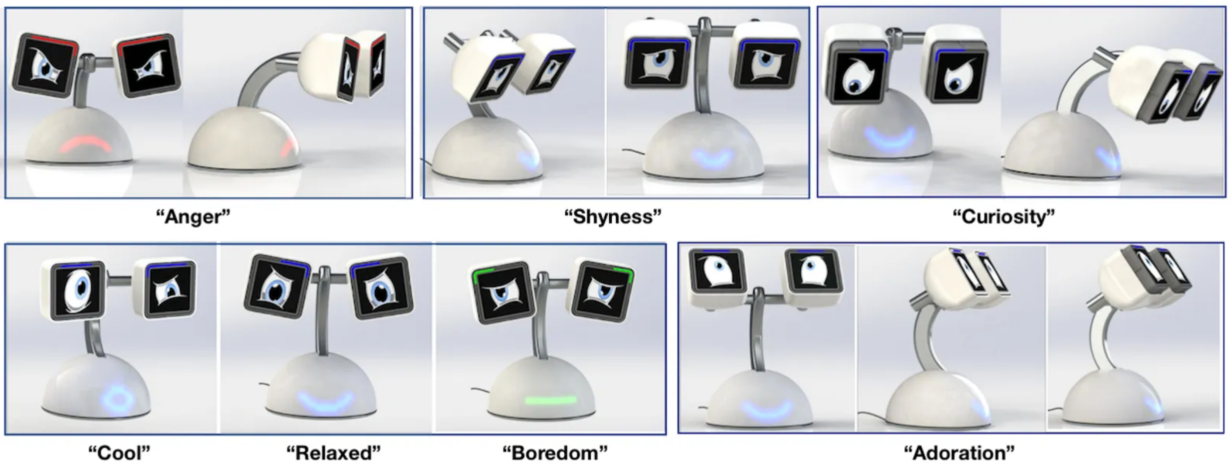


Figure 2.3: Haru's "emotive affordance." [34]

The size of Haru was carefully chosen to accommodate all the necessary components. It has a 22-centimetre diameter base that is about 15 centimetres tall, the minimum size required to fit all of the actuators, an addressable LED matrix display, a micro speaker, and an electronics control board. For uniform dispersion of light, the LED matrix display is housed in a customised structure that snugly fits the curvature of the external body casing [31]. Each of Haru's eyes consists of an outer shell measuring 95.5 mm x 95.5 mm x 61.19 mm and an inner part containing a 3-inch TFT display[31]. To reinforce the visual affordance of the eyes, the border of each eye is embedded with an addressable LED strip [31], [34].

Haru's emotive movements were determined through an in-depth participatory experiment with 50 volunteers, as empathy is a key consideration for a personal companion robot[31]. The robot's set of expressiveness and emotive movements, such as anger, shyness, curiosity, boredom, etc, are achieved through its five degrees of freedom (DOFs) [31], [34]. Haru has the movement capability of the LCD screens and LED rim eyes, as well as an LED matrix mouth in the base. It can also work with a range of movements, voice, and sound effects.

Honda Robotics envisions Haru as a multimodal communicative agent, interacting with users through nonverbal sounds (paralanguage), eye, face, and body movements (kinesics), and voice (language) [35]. Ongoing research with the Socially Intelligent Robotics Consortium [36] hopes Haru will drive research into robots as a new form of companion species and as a platform for creative content.

2.2.3 SciFi Inspired and Social Tabletop Robots

Science fiction has always fueled our imagination, and animated movies have played a large role in developing acceptance of widespread human-robot interaction (HRI) [31]. Two examples of beloved robots in recent years are Baymax [37], [38] and Wall-E [39], known for their friendly interactions with humans while fulfilling their purposes. Despite both being mobile robots, their exterior designs are vastly different, with Baymax having a soft exterior and Wall-E having a metal shell. This shows how the design of animated robots has evolved over time but still serves as inspiration for many present-day designs.

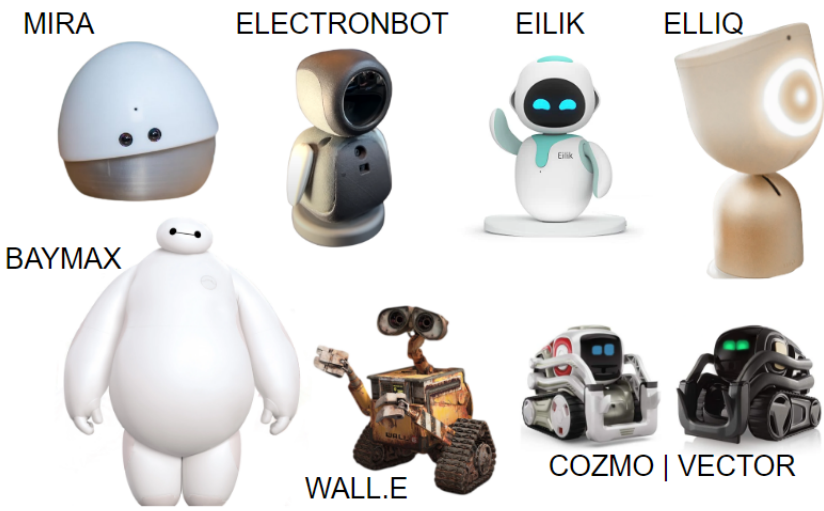


Figure 2.4: Inspiring robots (Mira [40], ElectronBot [41], Eilik [30], ElliQ [29], Baymax [38], [42], Wall.E [39], [43], Cozmo | Vector [28])

Cozmo and Vector [28] are small-sized robotic toys designed by Anki in 2016 and 2018, respectively. Both robots have a visual appearance of a robotic vehicle with animated eyes that they use to portray emotions. Their eyes, sounds, and behaviours give them a cute animated appearance. Cozmo was used to spark collaboration and engagement in various research [10] and even to teach mathematics [44].

ElectronBot [41] was created by Peng Zhihui and is an open-source project for building a small desktop robot that can be DIY 3D printed. It uses house motors for flipper movement and off-the-shelf modules for the circular LCD and camera.

Mira was created in 2017 by Alonso Martinez but was inspired by the open-source DIY 'Pia the robot' [40]. It has a simple spherical body and abstract egg-shaped tabletop robot design. It communicates through emotional rather than verbal language, using fluid responsive motions, and its emotion is conveyed through subtle movements and responses [45].

Eilik [30] was launched on Kickstarter in 2022 and is a work desk companion that responds to different touch interactions. Its size is 108 x 105 x 133 mm, and it weighs 230g. It has speakers, servos, an OLED display, touch sensors, and a microphone. You can play games via its display and use it to set a timer.

ElliQ [29] is a companion bot in healthcare designed mostly for elderly folks. It has a clean, modern design with a tablet extension interface and is an interactive tabletop health companion that helps older adults cope with required mental and social activity. Following five years of extensive research and insights from beta testing with older adults, Intuition Robotics launched ElliQ into the market in March 2022 [46].

Although these robots have multiple sensors and creative designs that explore abstract or less anthropomorphic social robot designs, they are primarily intended for entertainment purposes [28], [30], [40]. However, given their embodiment and potential for social interaction, there is room for improvement in their ability to meet high standards of sociality, function, and adaptability to the rhythms of everyday life as noted in [47].

2.3 NLP Technologies

Natural Language Processing (NLP) technologies have revolutionised the field of artificial intelligence, allowing computers to understand, interpret, and generate human language. Speech is among the most common interaction modalities used by social robots [1], as it enables more intuitive and human-like interactions between users and machines [48]. By recognising speech and processing language, social robots can understand user commands, respond to questions, and even engage in conversation.

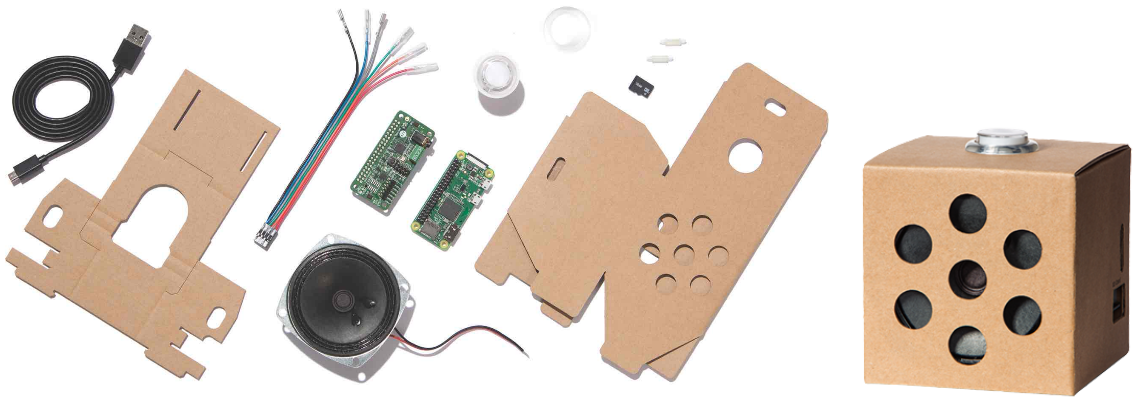
2.3.1 Conversational Agents: Chatbots and ITS

Conversational agents are a class of dialogue systems that have been researched in the field of communications for decades [49]. Chatbots are one category of conversational agents, which are software systems that mimic interactions with real people. They are intelligent and can accept natural language input in the form of text, voice, or both, and can provide conversational output in the form of verbal or text responses or even execute tasks [49]. One of the earliest chatbot systems was ELIZA by Weizenbaum in 1966, which was intended to demonstrate natural language conversation with a computer [50].

In recent years, conversational agents have found a new application in education [51] through the use of Conversational Intelligent Tutoring Systems (C-ITS), which provide personalised instruction to students through dialogue and can operate with or without the presence of a human teacher. An example, [51] evaluates the use of the Rasa [52], an open-source conversational toolkit for detecting student intents and building a new natural language interface in algebra tutoring.

2.3.2 Voice Chatbot: Google AIY Kit

An easy and hands-on way to start using a voice chatbot is through the Google AIY Voice Kit, a do-it-yourself artificial intelligence (AI) project kit [53]. This kit allows makers and developers to build their own intelligent speaker device with speech recognition and Google Assistant [54]. The latest version of the voice kit (v2) includes a Voice Bonnet that houses a microphone on the board itself, a speaker, an LED arcade-style button, and other components needed to assemble a voice-powered device using the Google Assistant SDK and Raspberry Pi. Google released the first AIY Voice Kit in collaboration with the Raspberry Pi Foundation in 2017 [53].



(a) List of materials inside the kit (b) Assembled Voice Kit hardware

Figure 2.5: Google AIY Voice Kit version 2 (v2) [53].

The Google AIY Voice Kit uses Dialogflow as its natural language processing (NLP) engine to understand and respond to user voice commands. Dialogflow [55] provides a platform for creating conversational interfaces and allows developers to build voice-enabled applications with NLU and integration with other Google services. This conversational toolkit is also utilised in other research, such as [56] which developed a Scratch tutorial chatbot (ScratchThAI) to teach young learners about MIT's Scratch programming language through a messaging platform. ScratchThAI uses DialogFlow to train data and build an NLU model.

2.3.3 OpenAI GPT Language Model

OpenAI [57] has advanced the NLP field with their research and models. Their transformer model, GPT-3, released in 2020, has 175 billion parameters, and is among one of the most advanced language models available [58]. One of the key features of this model is recognising learning capabilities, which allow it to learn from a small number of data examples. GPT-3 has shown impressive results in synthetic and qualitative tasks such as arithmetic, news article generation, and correcting English grammar [58]. The model generates text based on what it learned during its training period, during which it scanned vast amounts of text [59]. The GPT-3 model with 175 billion parameters is referred to as davinci.

To address the need for aligning language models with user intent across various tasks, researchers have explored the approach of fine-tuning with human feedback [60]. This involves collecting a dataset of labeller demonstrations that showcase the desired model behaviour and using it to fine-tune GPT-3 through supervised learning. Notably, this process has led to the development of InstructGPT [61], which outperforms GPT-3 at following English instructions, as evidenced by the results of human evaluations [60]. This improvement has led to OpenAI's GPT-3 model becoming a more effective and aligned general-purpose AI system that understands human intentions.

At the time of writing, the GPT-3.5 models are currently the most stable releases from OpenAI. The most capable and cost-effective model in the GPT-3.5 family is the gpt-3.5-turbo, which is the same model that powers ChatGPT [62]. GPT-3.5 is based on the improved InstructGPT model text-davinci-003 and has been optimised for chat but also works well for traditional completion tasks [63]. The training data for gpt-3.5-turbo is up to September 2021. On the other hand, the text-davinci-003 model can perform any language task with better quality, longer output, and more consistent instruction-following than the smaller GPT-3 models curie, babbage, or ada models. However, its training data is only up to June 2021 [63].

This new AI advancement holds promise for education, with potential use cases such as Khan Academy's AI-assisted tutor, Khanmigo [64]. The chatbot functions similarly to a real-life or online tutor, analysing students' work and providing assistance when they encounter difficulties. For instance, when solving a math problem, Khanmigo can identify not only if a student's answer is correct or incorrect but also where they may have made an error in their reasoning [64].

3 Requirements

This section defines the necessary requirements for developing the RSC. These requirements are derived from a comprehensive literature review synthesis, key findings from existing social robots in education, insights gathered from teachers' and university students' perspectives, and the utilisation of NLP technologies. By identifying and documenting these requirements, this thesis establishes the foundation for effectively designing and developing an RSC that can serve as a practical study companion for university students.

To achieve symbiosis, the RSC should be designed to seamlessly assist the user without serving as a distraction while providing approachable and engaging conversations [17], [18], [22]. Any accompanying gestures or movements should enhance the conversation aspect without causing any interference[13], as this is crucial to ensure that the RSC fulfils its purpose of facilitating meaningful conversations that aid in the learning process[20]. Therefore, the RSC must be able to recognise and interpret human speech using NLP technologies. Additionally, it should be equipped with sensors and actuators to interact with the environment and respond appropriately.

The presented functional and non-functional requirements are based on best practices discussed in the background section and informed by my personal experience as a university student.

3.1 Functional Requirements

I. The design of the RSC should facilitate its role as a study companion for university students without causing distractions. It should have a compact and aesthetically pleasing embodiment, with the ability to interact with students through multiple modalities. Specifically:
   A. The RSC should be a multimodal communicative companion, with various interaction modalities to complement each other, with speech as its primary mode of interaction, followed by touch, movement, and light indication.
   B. It should be capable of accurately interpreting the student's question using natural language processing.
   C. The embodiment of the RSC should be appropriately sized. The average size of Jibo *(figure 2.2)*, Haru *(figure 2.3)*, and others *(figure 2.4)*, estimated somewhere between 6cm to 30cm, is ideal. It should be compact and portable like Cozmo, ElectronBot and Eilik.
   D. Most students should be perceived the RSC as aesthetically appealing, which could be subjectively evaluated in future works. Additionally, an outer shell can be considered for protection and to bring out design curves and neutral colours.

II. The RSC must be able to engage in interactive conversations with university students to assist with their studies. It must be able to:
    A. Respond to students' questions related to their coursework and provide clear and accurate explanations of specific topics.
    B. Test or query students' learned lessons, track their learning progress and correct errors..
    C. Offer suggestions for study materials and resources based on students' preferences and learning styles.
    D. Provide feedback and encouragement to students based on their progress and performance.
    E. Maintain a record of the topics discussed with each student for personalisation and future reference.

III. The RSC should be capable of adapting to individual student learning styles, providing personalised feedback, and recommending customised study plans. It can achieve this by
    A. Analyzing student performance data to identify strengths and weaknesses.
    B. Adjusting its responses and feedback based on the individual student's learning style and pace.
    C. Providing customised study plans and recommendations based on the student's learning preferences and goals.
    D. Using gamification techniques to make studying more engaging and fun for students.

3.2 Non-Functional Requirements

I. Adaptability: The RSC should be adaptable for use in various learning environments, such as home, library, or group study settings. To enhance the RSC's suitability for a range of study environments, features such as adjustable volume levels or noise-cancellation technology may be incorporated.

II. Portability: The design of the RSC should be handy for students to easily transport it between different study environments. For example, moving it to and from a student's dorm room or to a private library cubical.

III. Security: The RSC platform must be designed to protect the privacy and confidentiality of the student's personal and academic data, and ensure that only authorised individuals have access to it.

IV. Reliability: The RSC should be reliable and available for use whenever needed and should not suffer from frequent downtime or system failures.

V. Performance: The RSC should respond to student queries within a reasonable amount of time, ideally under a few seconds, to prevent frustration and delay.

VI. Maintainability: The RSC should be easy to maintain and update, and should not require extensive manual intervention or specialised expertise.

VII. Scalability: The RSC design should incorporate multiple interaction modalities to accommodate various learning styles. It should allow for easy customisation of peripherals so that students can tailor their experience to their specific needs and preferences.

VIII. Durability: The RSC should be designed to withstand regular use by university students without suffering from wear and tear. The system should be able to operate reliably over an extended period of time without requiring significant maintenance or repairs.

IX. Usability: The RSC should be easy to use and navigate and should not require extensive training or technical knowledge to operate. The system should be designed to minimise confusion and provide clear feedback to the user. The conversational aspect of the RSC should be engaging and helpful without being overly complex or confusing.

X. Curriculum support: The RSC should be designed to support various subjects or courses, providing adequate learning materials and resources that align with the curricula of the academic institutions or programs it is intended for.

4 Design and Development

This thesis adopts a design approach informed by the Design Science Research Method [65]. This approach involves creating an innovative, purposeful artefact for a specific problem domain while simultaneously generating knowledge about the design process and the artefact itself. A systematic design process was followed that involved identifying the problem domain, specifying the requirements, and designing and prototyping the RSC. Design decisions were informed by the extensive background and requirements.

The design thinking model involves a constant iteration process, where each cycle builds upon the understanding gained from the previous cycle [31]. Ideally, a design research study would have taken place to survey students' preferences in this region and educational settings. However, based on prior surveys and discussions on the intended use [13], [17], [18], [23], a prototype was developed to contribute to the open-source community and facilitate future iterations.

4.1 RSC Overview

The Robotic Study Companion (RSC) is an open-source tabletop robot (Figure 4.1) designed for university students to enhance their learning experience. It provides various interaction modalities, including voice, feedback, and touch, to facilitate engaging study sessions. The RSC balances cost and technical performance, making it affordable for educational institutions and students. The first version, presented in this thesis, utilises off-the-shelf components and offers reconfigurability, allowing users to add and replace sensors, computation units, and other components as needed. With its sleek design and friendly appearance, the RSC creates an inviting and approachable presence. It serves as a study companion, offering academic support and aiding students in their learning journey.

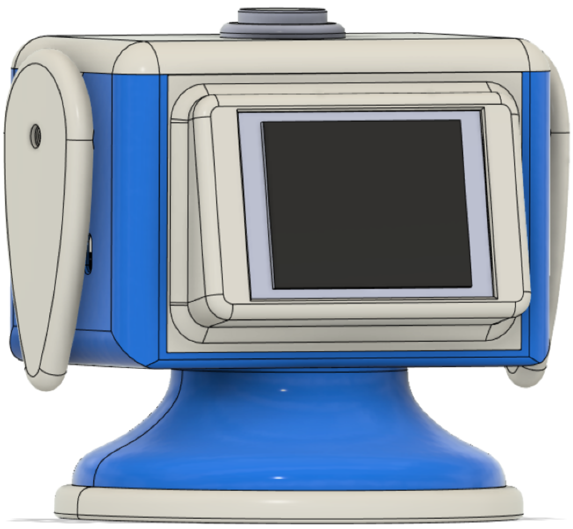


Figure 4.1: Side-by-side view of the CAD and the printed prototype of the RSC.

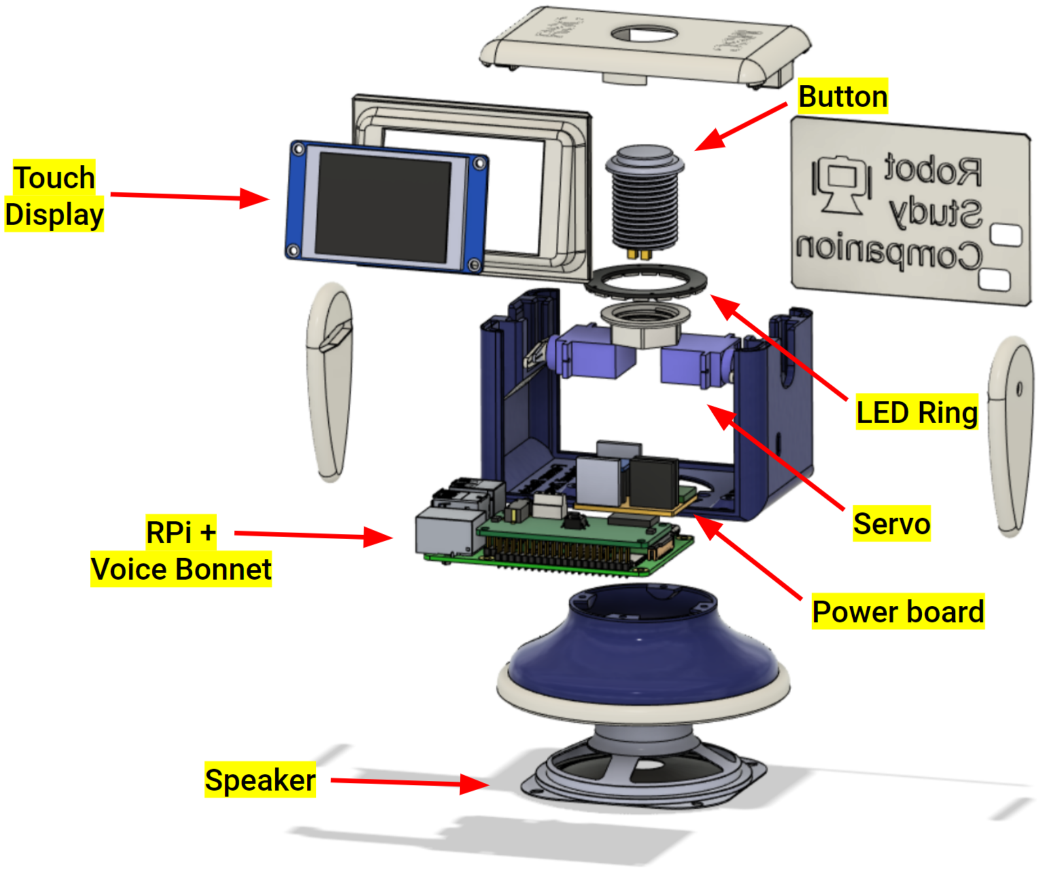


Figure 4.2: Exploded view of the RSC showcasing its internal components and structure.

4.1.1 Modalities of User Interaction

This sub-section provides a breakdown of the interaction modalities currently employed by the RSC and those that can be implemented in the future. To fulfil the functional and non-functional requirements of a multimodal communicative companion, the RSC incorporates a variety of modalities that encompass different ways for users to engage and interact during study sessions. Depending on user preference, not all modalities need to be activated simultaneously, and the possibilities extend beyond these suggested use cases.

I. Voice-based: The RSC incorporates a speaker and microphones to recognise and respond to voice commands. Users are prompted to engage in conversation with the RSC, and their speech is captured by the microphone using Speech-to-Text (STT) software. The RSC then generates responses using Text-to-Speech (TTS) software and delivers them through the speaker. Currently, the RSC's responses are generated by OpenAI's GPT model.

II. Gesture: The RSC is designed to visually communicate and interact with users through gestures. In the current design, flipper-like structures (Figure 4.2) attached to servo motors are utilised for this purpose. These flippers are programmed to indicate the robot's readiness to engage with users. For instance, the RSC may wave to greet the user and indicate that it is ready to assist.

III. Tactile interaction: One mode of touch input on the RSC is the arcade button, which users can press for various cues, such as indicating the robot's readiness to listen. Ensuring clear visibility of the robot's activated or deactivated sensors is crucial to alleviate any feelings of constant surveillance and provide users with a sense of security and control [6], [18].

In addition to the arcade button, an easily accessible power button for power control should be included in the RSC design to contribute to the prevention of potential damage to the robot. However, it is worth noting that this power button was not incorporated in the current iteration.

IV. Display screen: The RSC is equipped with a touch display that can allow users to adjust volume and access other settings through a menu. Furthermore, the display can be utilized to present information or visually explain concepts to users [17].

V. Feedback mechanisms: The RSC currently employs three primary methods to provide feedback to users:

A. Light: The RSC incorporates RGB LED lights as indicators for users. These lights can display the loading stage of the robot, indicate engagement in a conversation, or signal malfunctions, such as by illuminating a red colour.

B. Motion: The RSC utilises servo motors to showcase interactivity and expressiveness. Physical movements, such as rotating flippers, demonstrate the robot's responsiveness and liveliness.

C. Sound: In addition to its speech capabilities, the RSC can employ various sound indications to convey its status or state. This feature, commonly found in existing robots [10], [45], can enhance user understanding and interaction with the RSC.

VI. Other modalities: To fulfil the remaining requirements, a suggested approach involves the use of a web application that is synced with the RSC. This enables users to access information such as their study progress and a record of topics discussed, providing a means to measure their performance. It is important to note that the web interface can be considered as a potential future addition.

4.2 Mechanical Design

In the process of designing the RSC, the mechanical aspect played a crucial role in realising the identified functionalities, particularly verbal and non-verbal interaction. A primary challenge revolved around integrating the selected components into a compact form while ensuring the necessary protection and convenient access to power charging. This section delves into the conceptualisation, CAD and prototyping phases, highlighting the key considerations and decisions made during the mechanical design process.

4.2.1 Conceptualization and Sketching

The design of the RSC was inspired by existing solutions, including Jibo *(figure 2.2)*, Haru *(figure 2.3)*, Mira, ElectronBot, Eilik, ElliQ *(figure 2.4)* and the AIY voice kit cardboard box *(figure 2.5-b)*. Robots that are already available on the market or used in research were chosen to leverage their strengths and limitations.

As part of the design process, an informal evaluation was conducted to select the most common shapes and features of the listed robots (under section 2.2). This involved incorporating elements such as curved edges, circular shapes, and neutral-coloured external shell designs with an internal structure (Figure 4.3). Additionally, considerations were made regarding the requirement for the robot to function as a tabletop device (Figure 4.4) rather than a larger, more mobile one, in order to fulfill the desired portability and scalability requirements.

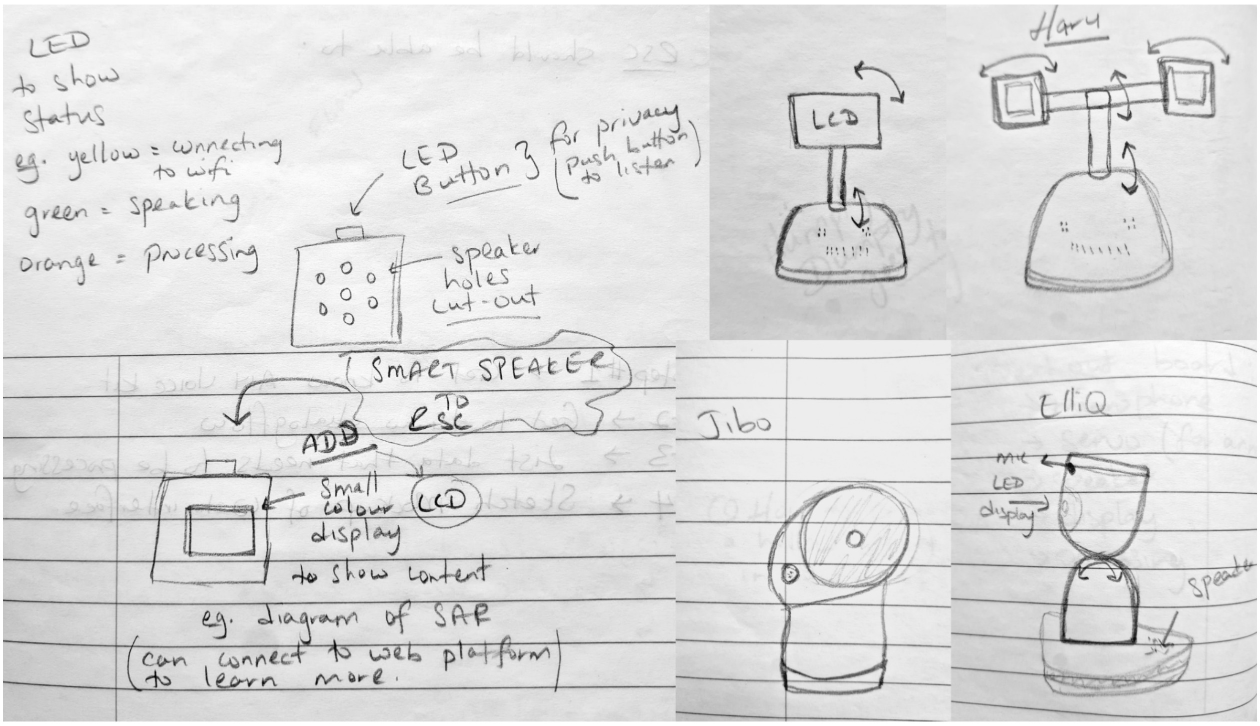


Figure 4.3: Sketches of Existing Solutions. Brainstorming ideas from AIY Voice Kit, Haru, Jibo, and ElliQ for inspiration in the design process.

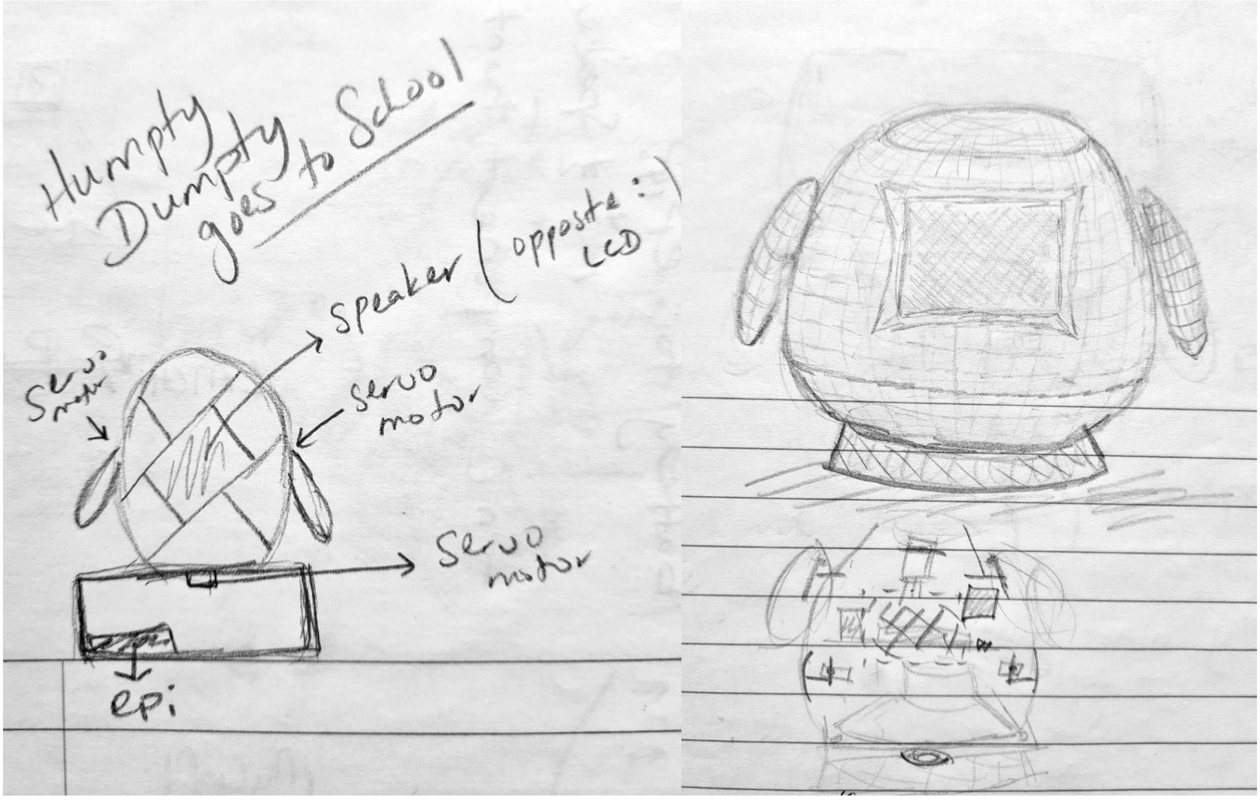


Figure 4.4: Initial design concept for a spherical robot body inspired by existing solutions. Due to challenges in housing all the components, the attempt to model this design was unsuccessful and earned the nickname 'Humpty Dumpty' due to its resemblance to an egg sitting on a wall.

The final design sketch (Figure 4.5) of the RSC incorporates several design principles that were identified [17]. Students preferred a practical, clean machine-like robot with limited human characteristics, especially in its physical appearance. They also emphasised the importance of easy handling to prevent interruptions during learning processes. Additionally, the chosen design had to be a social robot capable of interacting and communicating with humans in a natural and intuitive way.

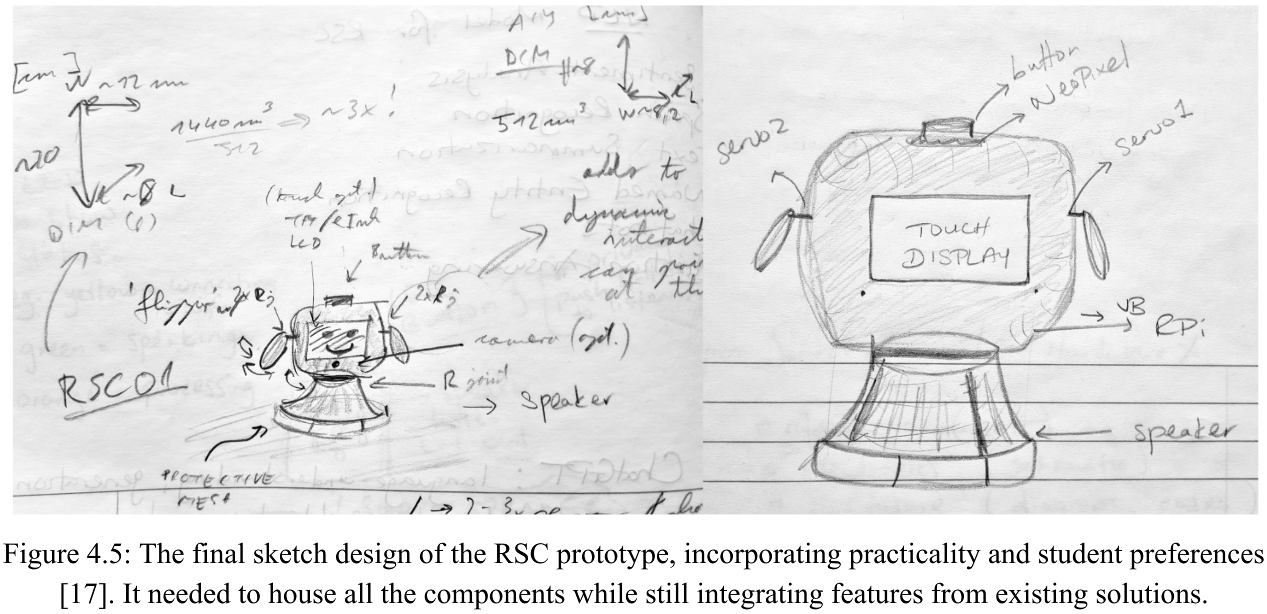


Figure 4.5: The final sketch design of the RSC prototype, incorporating practicality and student preferences [17]. It needed to house all the components while still integrating features from existing solutions.

Although students expressed a desire for educational robots to recognise and display basic human emotions [17], this feature was not included in the current RSC prototype but can be considered for future iterations. Verbal interaction can convey basic emotions, but the absence of a camera and facial recognition technology prevents the recognition of emotions through facial expressions. The decision not to include a camera was made to address privacy and security concerns.

4.2.2 CAD Modeling

All the chosen peripherals listed in the BOM (table 4.1) had to be sourced or created in CAD. This means that each model (Figure 4.6, Figure 4.7) had to be carefully dimensioned and almost accurate before start designing the physical RSC in CAD software.

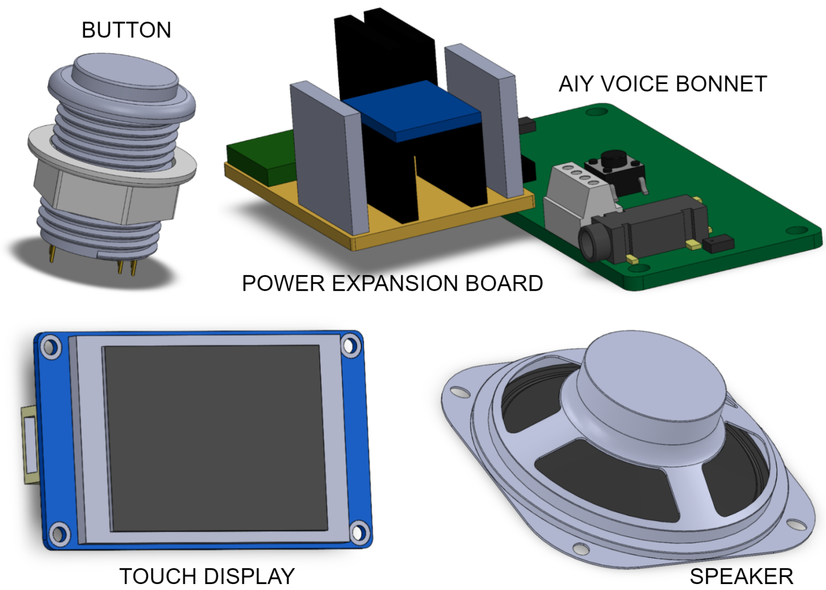


Figure 4.6: CAD models of various peripherals created using SolidWorks 2022.

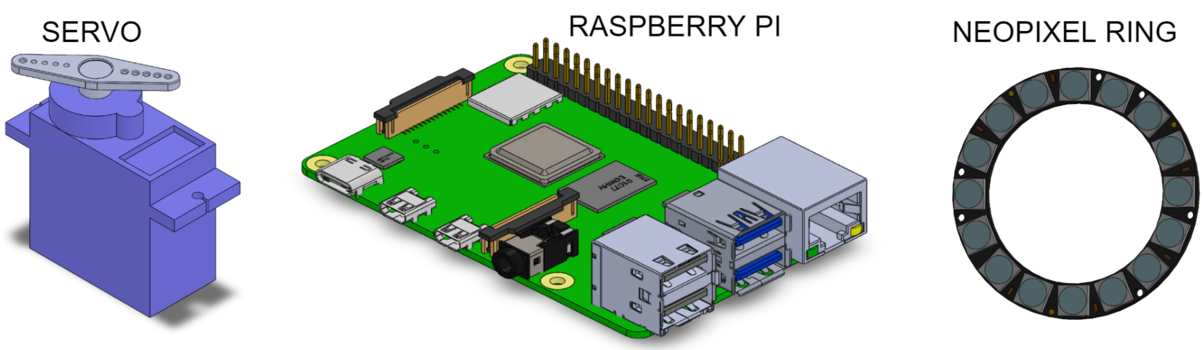


Figure 4.7: CAD models of other components used in the RSC obtained from Grabcad, namely: servo [66], Raspberry Pi 4 [67], and NeoPixel Ring [68].

To tackle the challenge of determining the ideal size and placement of peripherals, a rapid prototype was developed as an initial step (Figure 4.8). This prototype proved valuable in identifying design and practical issues that would have gone unnoticed in purely theoretical research. Furthermore, the prototype discovered a tangled and disorganised wiring layout, which highlighted the need for a design that minimizes the potential for wire interferences.

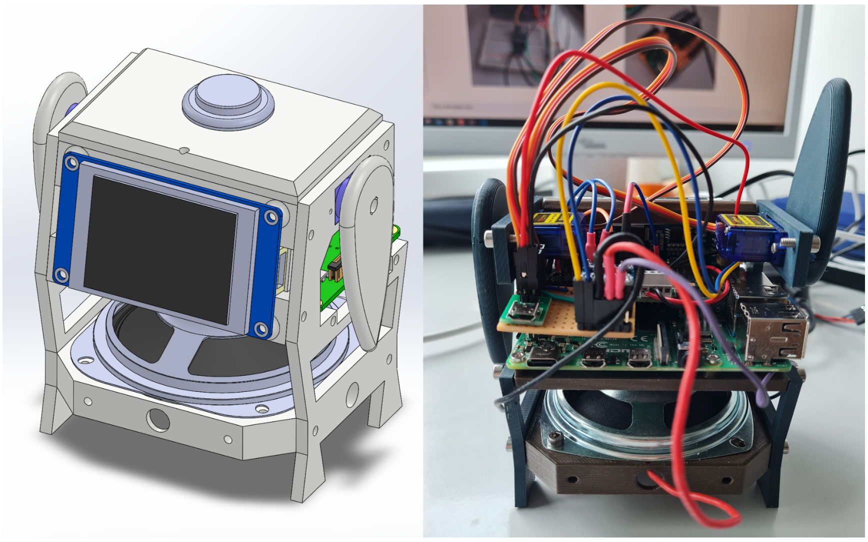


Figure 4.8: Side-by-side comparison of the CAD model and the first printed prototype, highlighting the wiring mess and the need for better design to help mitigate cross-talk.

One design concept was to have a more approachable aesthetic with a rounded or ball-like exterior. Figure 4.9 shows an attempt at this design, named 'Humpty Dumpty', with a cut-out spherical exterior intended to be printed with flexible filaments. However, the design process was challenging and time-consuming. Moreover, it was concluded that the extra bulkiness would make the RSC less practical for daily use, particularly when handling the device. Consequently, this version was quickly deprecated in favour of a more cost-effective approach.

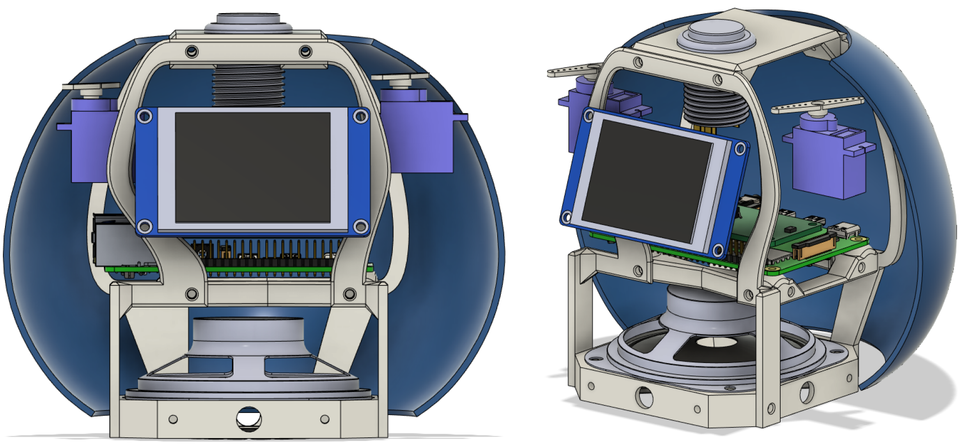


Figure 4.9: Initial design concept for the spherical body shell.

The physical design of the RSC prototype (Figure 4.10) prioritises portability, with a height of 11cm and a weight of approximately 374g. The device features a circular base shroud with a diameter of 11cm, which conceals and protects the speaker. The speaker, being the heaviest component, is located at the base, while the rectangular Raspberry Pi 4 and voice bonnet, responsible for computing, power distribution, and peripherals, are positioned above it. Integrating the touch display into the curved exterior has proven challenging due to its rectangular shape. Thus the overall shape was determined by the need to accommodate all of these peripherals.

Servos are used to actuate small flippers or ears, and they fit into bespoke slots within the main chassis. The lid of the robot features an arcade button that provides additional user interaction options. The locking nut of the button also retains a neoPixel ring against the lid internally, which illuminates the robot's interior and enables custom light animations.

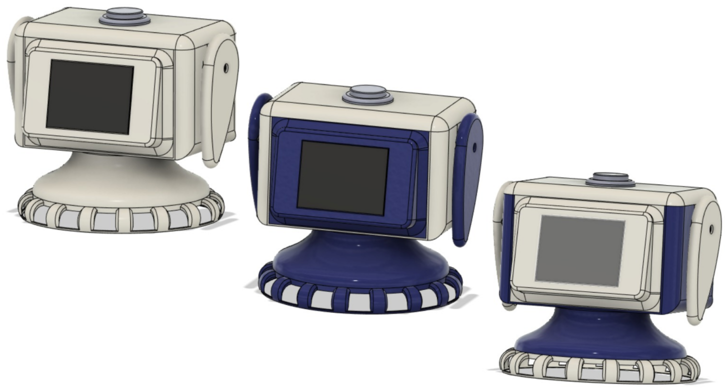


Figure 4.10: CAD models showcasing different colour combinations for the RSC prototype.

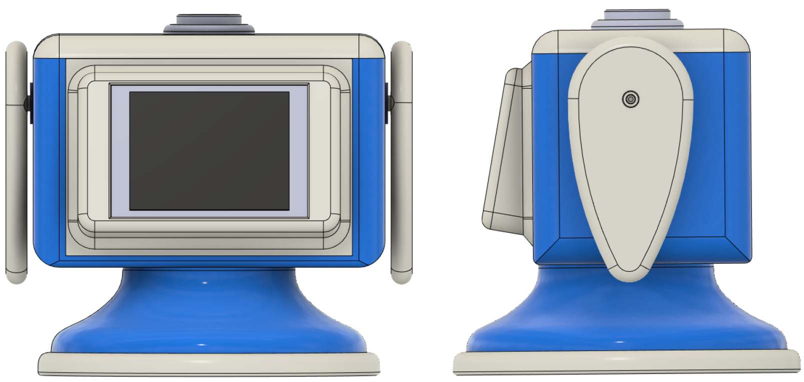


Figure 4.11: Front and side view of the final CAD model of the RSC.

All of the components are part of a snap-fit construction that protects the robot's internals. The hourglass shape incorporates both rounded and rectangular features. The blue and white colour palette (Figure 4.1) was chosen to create a more appealing and friendly look [69].

All CAD files can be found under the folder name 'MechDesign' via the GitHub link [70].

4.2.3 Prototype Development

The RSC prototype is fabricated using 3D printing technology. In the initial printed prototype, the structure relied on an excessive number of M3 screws for assembly. Subsequent iterations (Figure 4.12) aimed to streamline the construction process by reducing the number of screws used. The final 3D-printed version of the RSC is composed of white and azure blue PLA filament. The components are primarily held together through snap-fit connections, ensuring a secure assembly. The only set of screws present in the prototype is used to connect the base, where the speaker is located, to the upper/head section, which houses the remaining peripherals.

It is worth noting that in a later iteration, the base was replaced with a simpler design (Figure 4.13), printed using TPU (Thermoplastic Polyurethane) material, as this offers better grip and stability on tabletop surfaces.

Figure 4.12: Printed iterations of the RSC prototype demonstrating the design evolution.

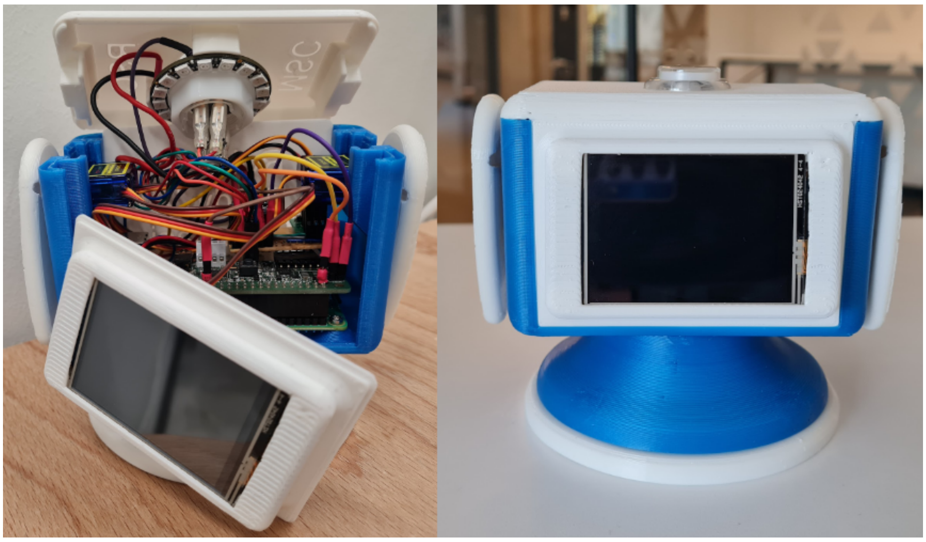


Figure 4.13: Internal and external looks of the final 3D-printed RSC prototype.

4.3 Electronics Components

In order to accomplish the objective of building a functional prototype of the RSC, the plan was to explore the use of open-source technologies with an emphasis on affordability. The plan was to incorporate all the necessary multimodal interaction features using off-the-shelf hardware components.

The peripherals for the RSC prototype were selected based on their ability to provide seamless interaction modalities with the user, as well as their reliability and compatibility with the system. Each component was carefully chosen to ensure that it provided the necessary functionality required for the robot to operate effectively.

- **Raspberry Pi 4 Model B 4GB RAM**

The Raspberry Pi 4 (RPi4) Model B (Figure 4.7) [71] with 4GB of RAM has been chosen as the primary processing and development component for the RSC. This 85.6mm × 56.5mm single board computer is equipped with a 64-bit quad-core ARM Cortex-A72 1.8GHz CPU, which is well-suited for software computing and provides optimal performance during data analysis. It also features a Micro-SD card slot (opted for a 32 GB card) for loading the operating system and storing data. The RPi4 is powered by a USB-C power supply and has multiple ports, as well as GPIO headers for connecting external peripherals. In addition, it provides various connectivity options, such as Bluetooth 5.0, Gigabit Ethernet, and dual-band 802.11ac wireless networking. As the main processing unit for the RSC, it is capable of running the necessary software components required to achieve the outlined requirements.

- **AIY Voice Bonnet**

The AIY Voice Bonnet version 2 (Figure 4.14) [53] serves as the primary hardware for voice recognition on the RSC. This pHAT form-factor bonnet is equipped with features such as stereo microphones, speaker terminals, a headphone jack, an 8-pin button connector, UART breakout pins, and four expansion GPIO pins called PIN_A, PIN_B, PIN_C, and PIN_D [72]. With its communication over I2C and the need for 5V and 3V3 power, it uses seven GPIO pins [72] when attached to the RPi4, thereby allowing for easy integration with other components.

To ensure efficient and accurate voice recognition and enhance the user experience, hardware capable of hearing the user's voice from a distance was necessary. The Voice Bonnet v2 features two SPH1642HT5H-1 microphones [73], enabling far-field voice recognition. By connecting the voice bonnet to the Raspberry Pi via its GPIO pins, the RSC's voice recognition and interaction capabilities were seamlessly integrated with room for other peripherals.
After rigorous testing, the AIY Voice Bonnet v2 was determined to be the optimal choice for the first prototype of the RSC due to its compact size, excellent microphone performance, and compatibility with the Raspberry Pi.

- **Speaker**

The AIY Voice Kit [53] is equipped with an NH 4Ω 3W 76mm loudspeaker (Figure 4.14), which delivers excellent sound quality. The AIY Voice Bonnet v2 [53] features stereo speaker terminals for connection and specialised sound drivers, such as 'modinfo rl6231 rt5645 snd_aiy_voicebonnet' [73], to enable optimal performance. Following extensive testing, this speaker was selected for its superior ability to provide high-quality voice responses and its easy integration with the voice bonnet.

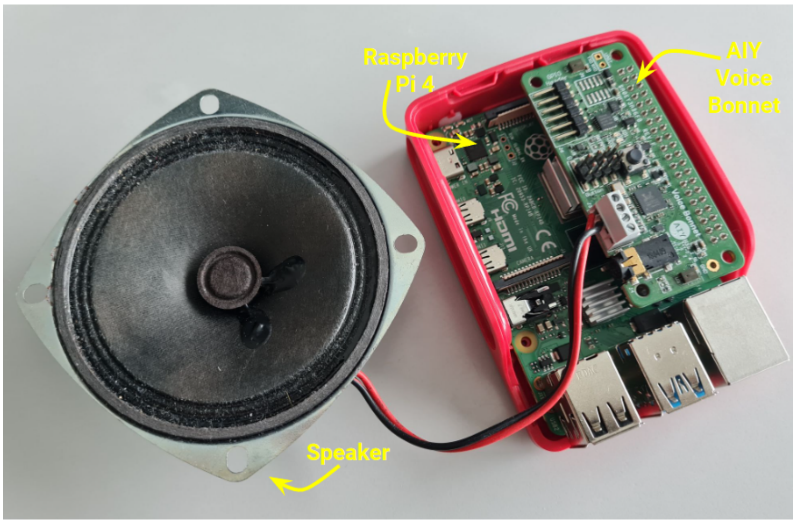
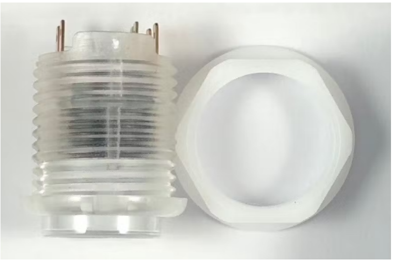


Figure 4.14: The RPi4 with the voice bonnet and speaker connected for speech recognition testing.

- **Arcade Button**

To enable a user-friendly interaction, the RSC incorporates a large arcade-style button (Figure 4.15-a) [53] connected to the 8-pin button connector via the Voice Bonnet, featuring an integrated LED that can be programmed to perform various functions, such as initiating a conversation. The button was selected for its ability to activate the "listening" mode, thus addressing privacy and security concerns and enabling a first-touch user interaction modality.



(a) AIY Aracde Button [53]   (b) Nexopixel Ring 16 [74]

Figure 4.15: AIY Kit Button for touch, and Adafruit NeoPixel for LED indication.

- **Neopixel Ring 16 x 5050 RGB LED**

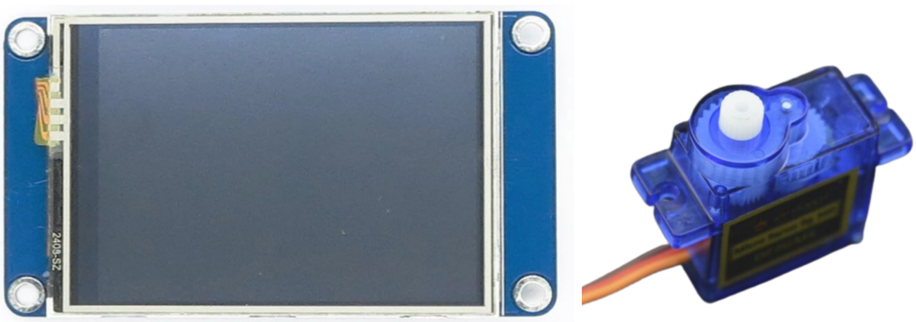
The neopixel Ring 16 (Figure 4.15-b) [74] is a circular board with 16 individually addressable RGB LEDs arranged in a circle, weighing 3.03g and with an outer diameter of 44.5mm. It requires a 5VDC power source and can be controlled through a single data line, making it compatible with most microcontrollers, including the RPi4. The LEDs have a constant current drive of ~18mA, ensuring consistent colours even if the voltage fluctuates [74]. The neopixel Ring is programmable and can create various visual effects like patterns and animations.

To integrate it with the RSC, the neopixel signal wiring is connected to one of the PWM GPIO pins on the voice bonnet. The neoPixel ring inside the RSC illuminates its interior, not only providing an aesthetic look but also serving as a visual indicator of its state. Additionally, it can function as a mood lamp [75], [76] when powered.

- **Nextion Enhanced NX3224T024 - Generic 2.4" HMI 320*240 Touch Display**

The Nextion Enhanced NX3224T024 is a 2.4-inch touch display (Figure 4.16a) with a high resolution of 320x240 pixels [77]. It comes with a powerful 8-bit microcontroller running at 200 MHz and 16MB of internal flash memory for storing user interfaces and touchscreen projects and requires 5VDC to operate. The display has a built-in LED backlight and a resistive touch panel for user interaction. It is also equipped with a UART interface, making it easy to connect to a microcontroller or single-board computer like the RPi. In the current RSC setup, the display is connected to the voice bonnet via UART breakout ports [72].

This display was chosen for its compact size, high resolution, and HMI capabilities, making it ideal for enhancing touch modality and meeting the RSC's adaptability requirements, such as displaying a menu. The Nextion displays are programmed using a proprietary language called Nextion Instruction Set (NIS) [78], which allows developers to create custom user interfaces and control logic.



(a) Nextion Touch Display [77] (b) SG90 Servo [79]

Figure 4.16 : HMI Touch Display for interaction and SG90 Servos for movement.

- **SG90 Servo Motors**

The TowerPro SG90C 360-degree Micro Servo (Figure 4.16-b) [79] was chosen for basic movement in the RSC prototype of this thesis. These servos weigh approximately 10g and measure 22.6mm by 30mm. Two servos were purchased and are connected to the voice bonnet via the expansion GPIO pins called PIN_A-D [72]. The servos are capable of producing a torque of 1.2 ~ 1.6 kg x cm, given the power consumption on a 4.8 ~ 6V power supply [79]. While it is possible to connect two servos to the RPi4, it is not recommended to power them directly from the 5V pins on the RPi, as the 5V pins may not be able to supply enough current to power both servos simultaneously. These servos were selected for testing purposes to determine the capabilities of the RSC setup.

4.3.1 BOM Summary

A comprehensive bill of materials (Table 4.1) was obtained for the RSC prototype, reflecting prices as of February 2023. The Raspberry Pi 4, its power supply, micro SD card, and the AIY Voice Bonnet were already in possession. The 3D printed filaments were provided through the university's makerlab but an estimated cost is given based on usage.

Table 4.1: Bill of Materials (BOM) for the RSC, including component names, descriptions, quantities, costs in euros, and website to source the items.

Component	Description	Base QTY	Cost Per Unit €	Amount EUR	Source of materials
RPi Power Supply	USB-C power supply with EU adapter	1	10	10.00	https://thepihut.com/products/raspberry-pi-psu-uk
Raspberry Pi 4B/4GB		1	45	45.00	https://www.digikey.ee/en/products/detail/raspberry-pi/RASPBERRY-PI-4B-2GB/10258782
AIY Voice Bonnet v2	Microphone: SPH1642HT5H-1 Speaker: NH 4ohm 3W 76mm LED Button	1	cost per mic: $1.4 cost per speaker: 1.95	60.00	https://www.adafruit.com/product/4080
Micro SD Card	RW, 32GB	1	9	9.00	https://thepihut.com/products/noobs-preinstalled-sd-card
Servos	SG90	2	3.91	7.82	https://mou.sr/3Wm1oPl
NeoPixel Ring	16 x 5050 RGB LED	1	9.95	9.95	https://www.adafruit.com/product/1463#technical-details
LCD Touch Display	Nextion Enhanced NX3224T024 - Generic 2.4" HMI 320*240 Touch Display	1	52.55	52.55	https://mou.sr/45fVW4q
Logic Level Shifter	4-Channel, Bidirectional	1	3	3.00	https://www.pololu.com/product/2595
3D Printed Filament	White & Blue PLA White Polyflex	~170+ g	-	20.00	https://blog.prusa3d.com/calculator_3416/
TOTAL	[Price as of February 2023]	**9**	**€**	**€217.32**	

4.3.2 Electronics Connection

The voice bonnet serves as the central hub, connecting all the peripherals on top of the Raspberry Pi, as shown in Figure 4.17. It acts as the data connection point for all the components of the RSC, however the RPi does the main processing.

To power the RSC's peripherals (i.e. neopixel, touch display, servos), a power expansion board was designed and implemented. Since the RPi power supply alone could not provide sufficient power to all these components, the Total Dissipated Power (TDP) of each component was calculated and added up. The combined TDP was estimated at around 3A, with a 5V supply, drawing 15W of power. By preparing the expansion board to handle the power distribution, the RSC was able to operate at full capacity.

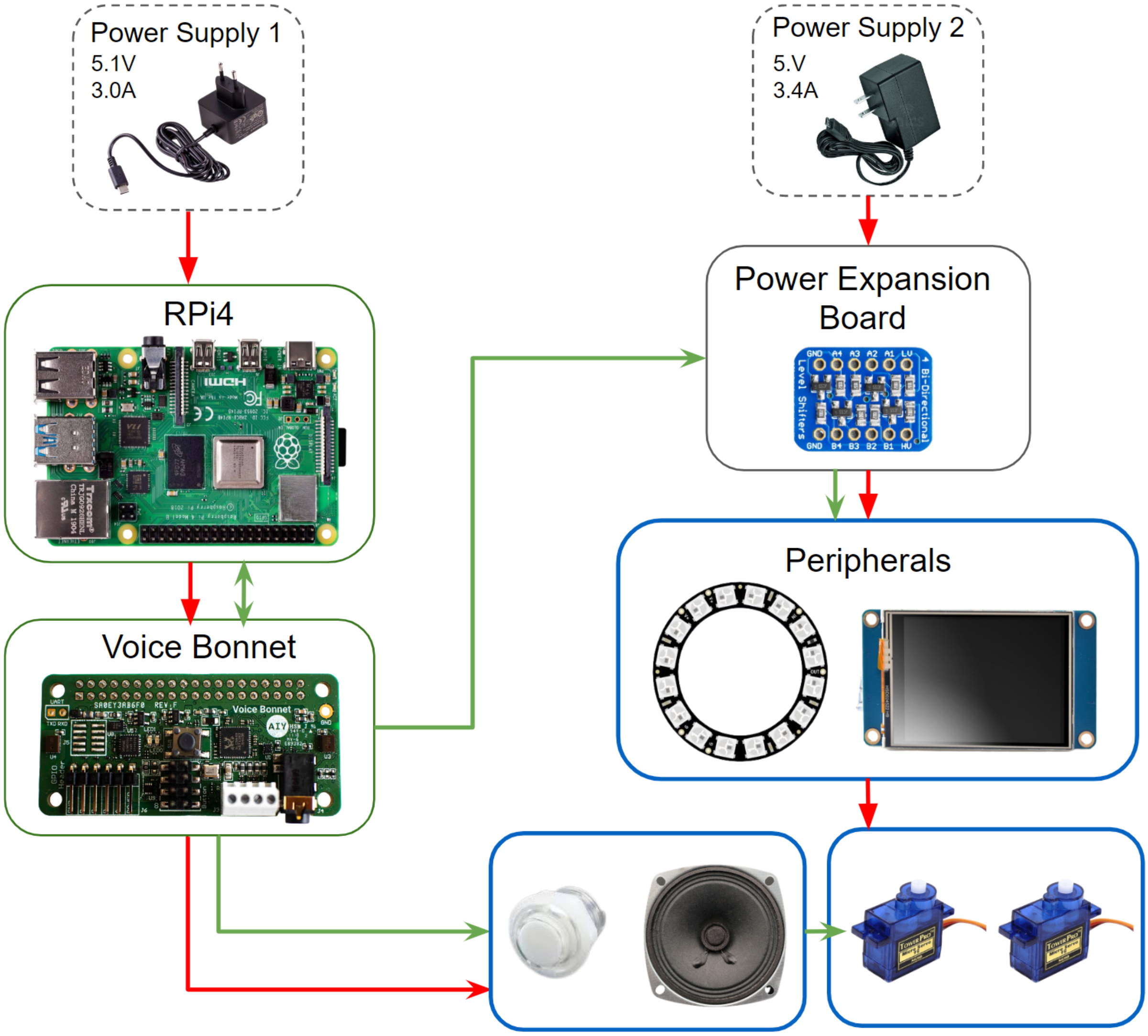


Figure 4.17: Connection Architecture of the RSC. The red arrows represent power delivery, while the green arrows denote data communication.

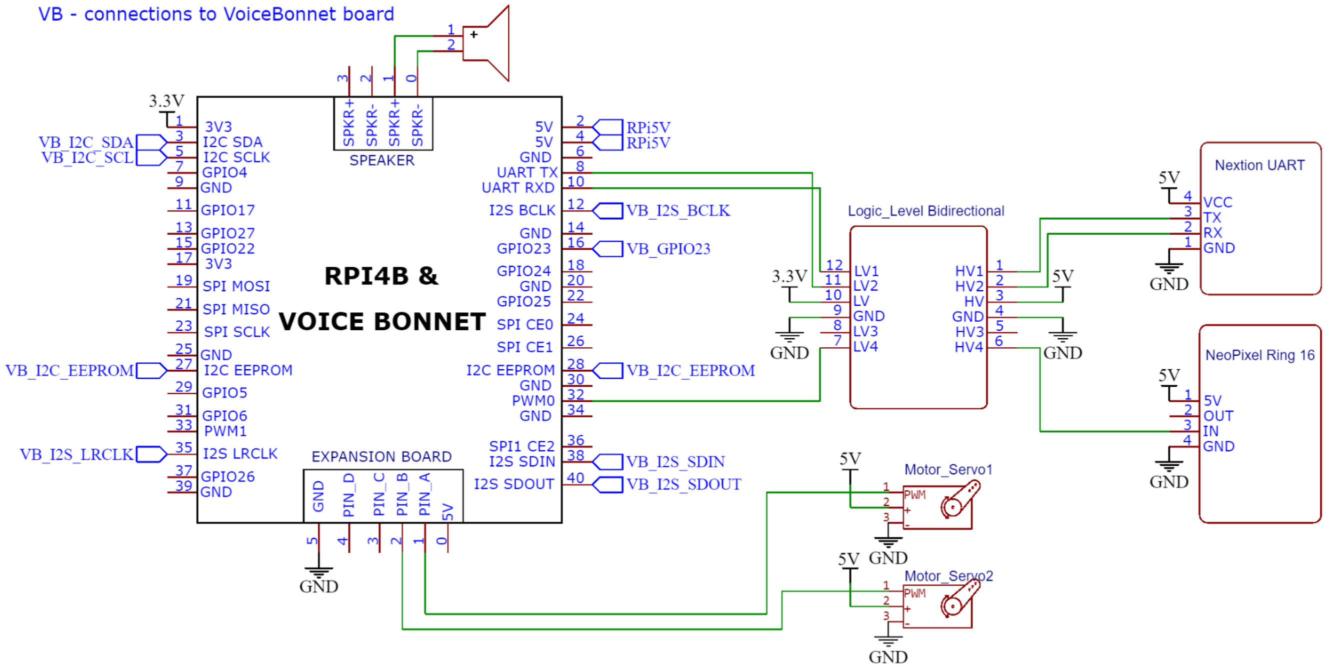


Figure 4.18: Electronic schematic of the RSC showing the wiring connections of the peripherals.

The power expansion board is equipped with a Logic Level Shifter to ensure the proper functioning of peripherals. It is physically connected to the voice bonnet through female headers that connect to the Voice Bonnet expansion PWM GPIO for the servos. The board also includes specific soldering connections, such as a GPIO pin for the neopixel, UART connections for the touch display, and a ground pin (Figure 4.18). Precise soldering techniques were employed to establish reliable connections between these pin headers and the voice bonnet on top the RPi (Figure 4.19).

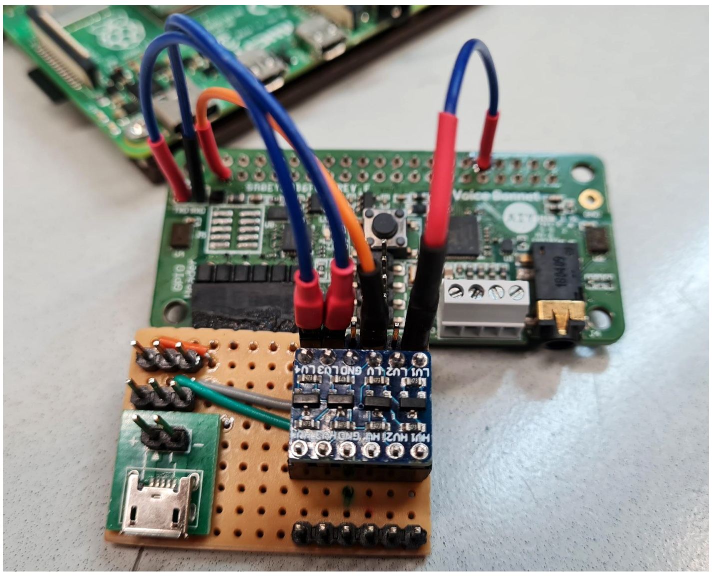


Figure 4.19: Second version of the power expansion board connected to the AIY Voice Bonnet.

4.4 Software Architecture

Creating the RSC prototype involved the integration of various components and software libraries to enable its functionality and provide seamless interaction with users. This section provides an overview of the software architecture and key libraries utilised in the development process.

The RSC's processing power and storage capacity are provided by the Raspberry Pi, utilising a special version of the RPi operating system, Raspbian GNU/Linux, known as the AIY system image [80]. This system image incorporates additional AIY software and offers easy-to-use Python APIs for the Voice Kit, particularly the Voice Bonnet. It's important to note that as of February 9, 2023, the associated GitHub repository for the AIY projects has been publicly archived [81], as Google has ceased production of the AIY kits.

Python's versatility and extensive ecosystem of open-source software libraries were instrumental in the prototyping process. Python version 3.7.3 [82] was the main programming language used, and Python3 packages were leveraged in the creation of the first prototype of the RSC.

4.4.1 Speech Processing

PyPI (Python Package Index) is a repository of software packages for the Python programming language. The SpeechRecognition library version 3.9.0 [83] enabled speech recognition, allowing the RSC to interpret user inputs by transcribing spoken language into written text. PyAudio version 0.2.11 [84] captured speech from users by providing the functionality for accessing the RSC audio devices, including both microphones and the speaker.

The pyttsx3 library version 2.90 [85] was used for text-to-speech conversion. This library allowed the RSC to convert text-based responses into spoken words, enhancing the interactive experience with users. With pyttsx3, you can control various aspects of speech synthesis, such as voice selection, speech rate, and volume.

These speech-processing libraries were simple to use, easy to run, and compatible with the Linux-based AIY operating system.

4.4.2 NLP API

The NLP aspect of the RSC's software relied on the text_davinci003 variant of the GPT (Generative Pre-trained Transformer) model, which served as a conversational API [63]. Specifically chosen for its advanced language processing capabilities, the text_davinci003 model can generate coherent and contextually relevant responses, making it an ideal choice for the RSC to enable engaging in interactive conversations with users.

To seamlessly integrate the text_davinci003 model, the RSC utilised the OpenAI Python Library version 0.27.4 [86]. This library provided a convenient and user-friendly interface to interact with

OpenAI models and APIs, simplifying the utilisation of GPT-3 for natural language understanding and generation. By leveraging the library’s capabilities, the RSC effectively incorporated the text_davinci003 model as a conversation API, enabling it to comprehend user inputs and generate contextually appropriate responses. This integration empowered the RSC to deliver engaging and meaningful interactions tailored to the specific context.

4.4.3 Peripheral Libraries

The current state of the RSC does not utilise the touch display. However, it uses the button, neopixel ring, servos, microphones and speaker for speech recognition (Figure 4.20).
Several libraries were utilised to control and interact with the RSC's peripherals.

- **CircuitPython**

The adafruit-circuitpython-neopixel library version 6.3.8 [87] integrated and controlled the RSC prototype’s neopixel RGB LED ring. This library provided a straightforward and user-friendly interface for managing the colour and brightness of individual pixels within Neopixel strips or matrices. By leveraging the capabilities of this CircuitPython library, the RSC can create dynamic lighting effects, enhancing the visual experience and interactivity of the device.

- **AIY API**

The AIY API library [73], [81] played a crucial role in the software integration of the RSC prototype, particularly in accessing and controlling the GPIO pins and interacting with the servos and arcade buttons. For the servos, the AIY API library provided a convenient interface to utilise the GPIO pins for servo control, enabling precise movement and positioning. Additionally, the library facilitated the integration of the button functionality, allowing users to interact with the RSC through physical button presses.

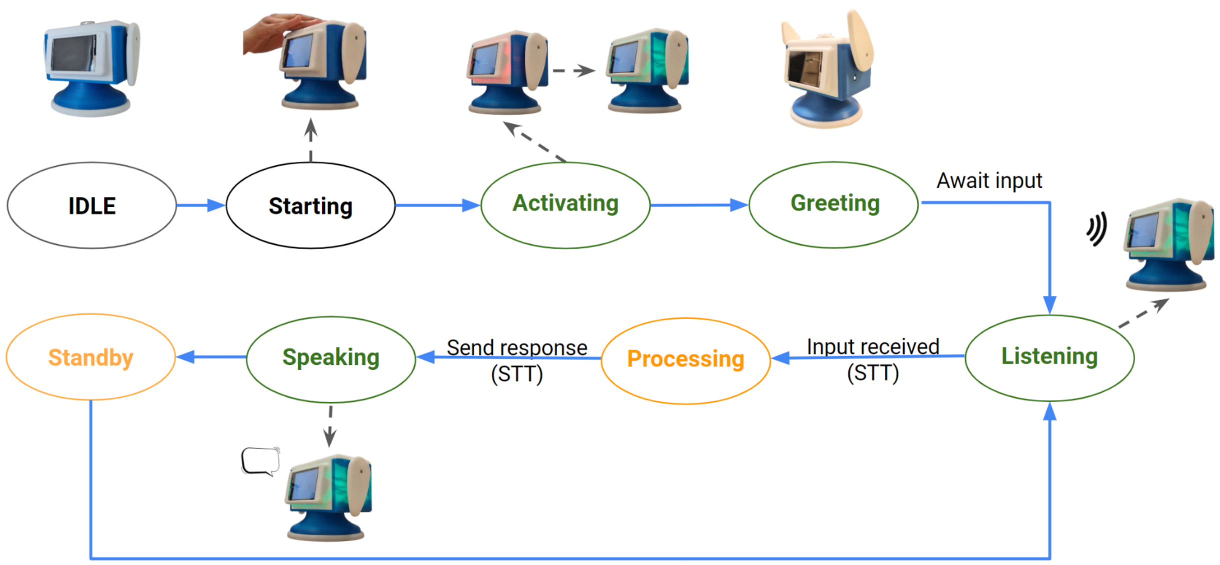


Figure 4.20: Functional state diagram of the RSC in its current state.

To connect to the RSC, software tools like Visual Studio (VS) Code IDE to SSH into the RPi can be used. There are two methods to obtain the IP address for connecting to the voice bonnet [53]: either use the AIY Projects app on an Android device or connect a monitor, mouse, and keyboard directly to the RPi's ports.

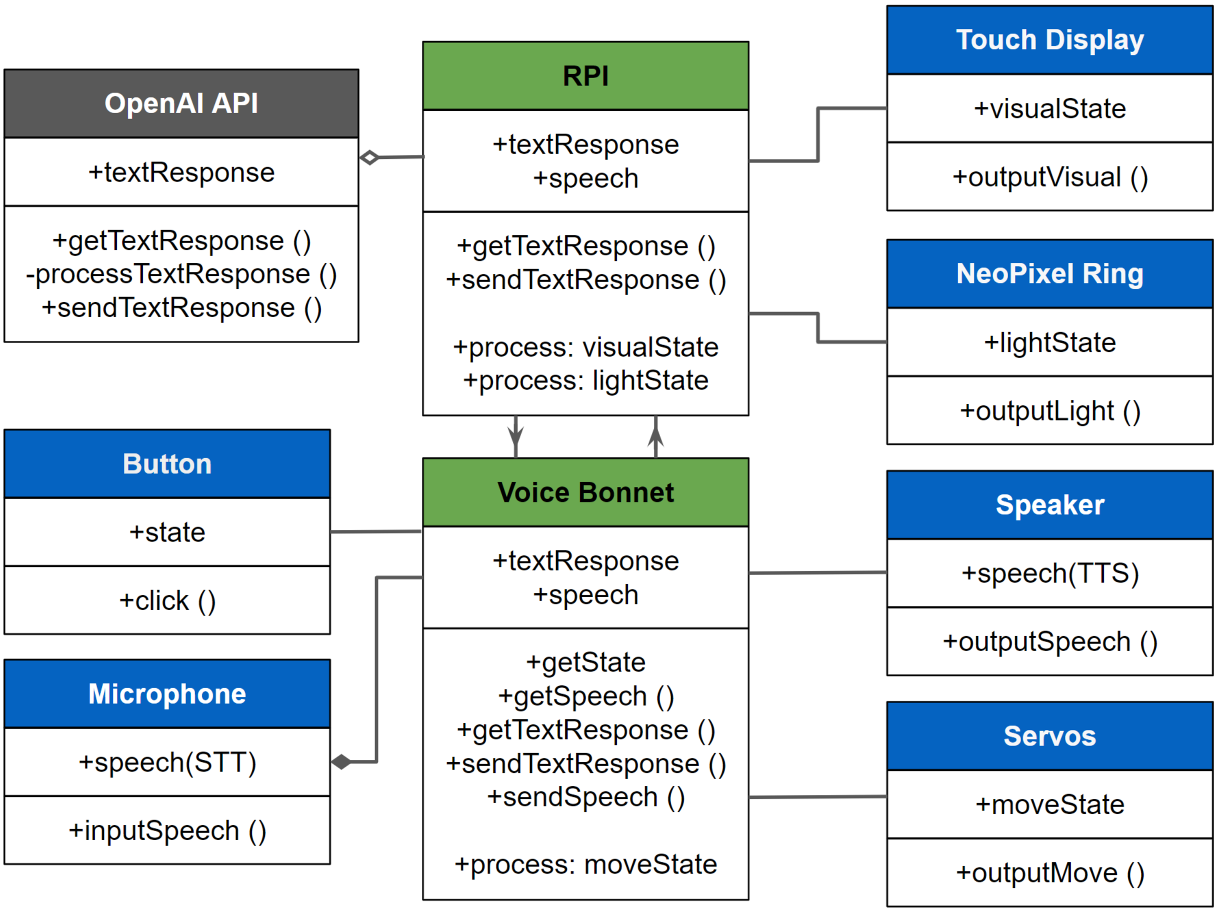


Figure 4.21: Class diagram illustrating all the peripherals of the RSC, including various objects, their attributes, and data operations.

There is significant room for improvement in the software development for the RSC. The class diagram (Figure 4.21) illustrates how all the peripherals communicate with their qualities (attributes) and data operations (methods). The visibility of an attribute or method sets the accessibility for that attribute or method: private (-) cannot be accessed by another class, whereas another class can access public (+). The lines represent the relationship between objects and classes. The empty diamond straight line denotes that OpenAI API operates independently of the RPi, due to the offloaded computation to the OpenAI services. Meanwhile, the filled line ending shows that audio input via the microphone depends on the Voice Bonnet.

All current and future development will be hosted on GitHub [70] to ensure efficient code management and also provides a video demo of the RSC in action.

5 Discussion

In this thesis, all research objectives were addressed, and some success was achieved. A comprehensive review of the existing literature was conducted. However, it is essential to acknowledge the vastness of the research on social robots for educational purposes, leaving much more to explore. Nonetheless, efforts were made to be inclusive and cover a wide range of relevant literature.

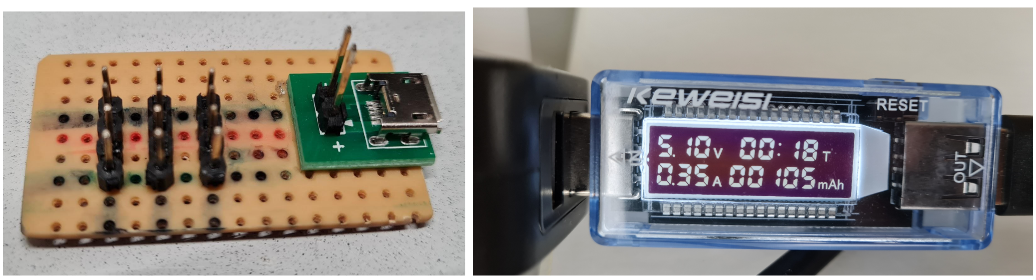
The functional and non-functional requirements for a university's robotic study companion (RSC) were outlined based on the literature review and key findings, guiding the design and development of the RSC. A working prototype incorporating NLP technologies was successfully designed and built. The hourglass shape and colour palette of the RSC provide a friendly aesthetic and the chosen off-the-shelf peripheral components facilitated rapid prototyping.
However, not all of the outlined requirements were fully met, with room for improvement in future iterations of the RSC. Particularly, the challenge of creating an RSC that can adapt to individual student learning styles and provide personalised feedback requires further work and investigation.

An official evaluation is still necessary to assess the effectiveness of the RSC. Nevertheless, the RSC demonstrates a compact and aesthetically pleasing embodiment, capable of interacting with students through multiple modalities.

5.1 Challenges

5.1.1 First Power Expansion Board

An initial power expansion board was created, but it caused numerous issues, such as intermittent servo and neopixel ring malfunctions. Despite careful soldering of the pin headers and excessive testing with a multimeter, the peripherals were still non-functioning, attributed to later discovered cross-talk and a voltage logic level mismatch. The RPi is a 3.3V logic appliance, whereas most peripherals operate at 5V logic, which was overlooked in the planning stage.



(a) First Breakout Board (b) USB Port Voltmeter

Figure 5.1: Old power expansion board and the USB voltmeter. When the RSC is assembled, the voltmeter measured power consumption, showing 5.10V and 0.35A during standby.

To address this problem, a Logic Level Shifter 4-Channel, Bidirectional [88], was added, and a second breakout or expansion board (Figure 4.19) was prepared. During soldering, considerations were taken in the layout planning to protect the wiring from cross-talk and interference by minimising path length and connection routing. However, adding the new level shifter did not solve the logic level issue. After rigorous testing and probing using an oscilloscope, it was concluded that the component was malfunctioning after connecting the device and measuring incorrect voltage levels at various test points. After replacing it with a spare, the logic voltage level mismatch was resolved successfully.

5.1.2 Unnatural Sounding Voice

A software challenge encountered during the development of the RSC was the issue of unnatural voice communication. Although the pyttsx3 library [85] controlled various aspects of speech synthesis, such as voice selection, speech rate, and volume, the generated voice still sounded robotic and was at times, difficult to understand.

Alternative solutions were explored [83], including experimenting with different speech recognition libraries. However, each alternative came with its own challenges, making it difficult to find a suitable solution that effectively addressed the issue of unnatural voice communication. Further research and experimentation are necessary to overcome this challenge and find a more natural and engaging voice communication solution for the RSC.

5.2 Limitations

The RSC implementation has some limitations, including technical constraints, software and lack of user evaluation. The design process encountered limitations due to technical constraints, such as integrating various components into a compact form factor, which required careful consideration of space and power requirements. Optimising the accuracy and reliability of the speech recognition system, handling diverse accents and dialects, and ensuring seamless integration with the overall functionality of the RSC were also significant hurdles. The now archived AIY Python library and AIY Voice bonnet also limited the RSC's potential by lacking certain software and hardware documentation.

Despite meeting most of the outlined design objectives, a user evaluation of the RSC regarding student preferences and documentation of their interaction with the RSC would have provided invaluable insights and outlined improvement and design refinement areas. The evaluation preparation process would require time-consuming and comprehensive documentation, with several accompanying committee approvals, even at a small scale. As a consequence, delegating the evaluation to future work was preferred. This limitation restricts the ability to gather valuable insights on user preferences, experiences, and areas of improvement, hindering the refinement process.

5.2 Future work

Long-term goals:

1. Improve and expand the design of the RSC to make it more durable and commercially viable.
    - ***1.1. Hardware***
        - 1.1.1. Use smaller peripherals with the aim of redesigning for a smaller form factor and portability.
        - 1.1.2. Implement a single power supply solution to simplify operation and maintenance. This involves consolidating the present multiple circuit boards (RPi4, Voice Bonnet & power expansion board) into a single, more efficient board.
    - ***1.2. Software***
        - 1.2.1. Utilise a more natural-sounding Text-to-Speech (TTS) voice to enhance responses to natural language inputs from users.
        - 1.2.2. Address potential privacy and security concerns associated with user data and interactions.
        - 1.2.3. Enhance the Automatic Speech Recognition (ASR) component to improve understanding of the user's dialogue state.
        - 1.2.4. Train customised language models specifically tailored for particular courses to optimise the RSC's performance in personalised learning.
        - 1.2.5. Develop a sequence model that can effectively map input words to slot fillers, domain, and intent.
        - 1.2.6. Customize prompt engineering to enhance the user experience and engagement.
        - 1.2.7. Incorporate customised light, sound, and movement indications to facilitate more natural and intuitive interaction with the user.

2. Add affective emotion reading capabilities to the RSC, enabling it to detect and respond to the user's emotional state. This can be achieved by integrating emotion detection technologies, such as voice analysis, into the RSC's software.
   The RSC should then be able to display appropriate motions and sounds based on the detected emotion with the aim to stimulate positive emotions in the user.

3. Develop a web-based application (Figure 5.2) to support students in various courses and track their progress. The application will enable the RSC to interact with students and gather data on their understanding of course materials. This data will be stored in a database that can be accessed by both the lecturer and students. The application will also provide tools for tracking students' progress, such as quiz scores and performance metrics, and allow for customised prompts and feedback based on individual learning styles.

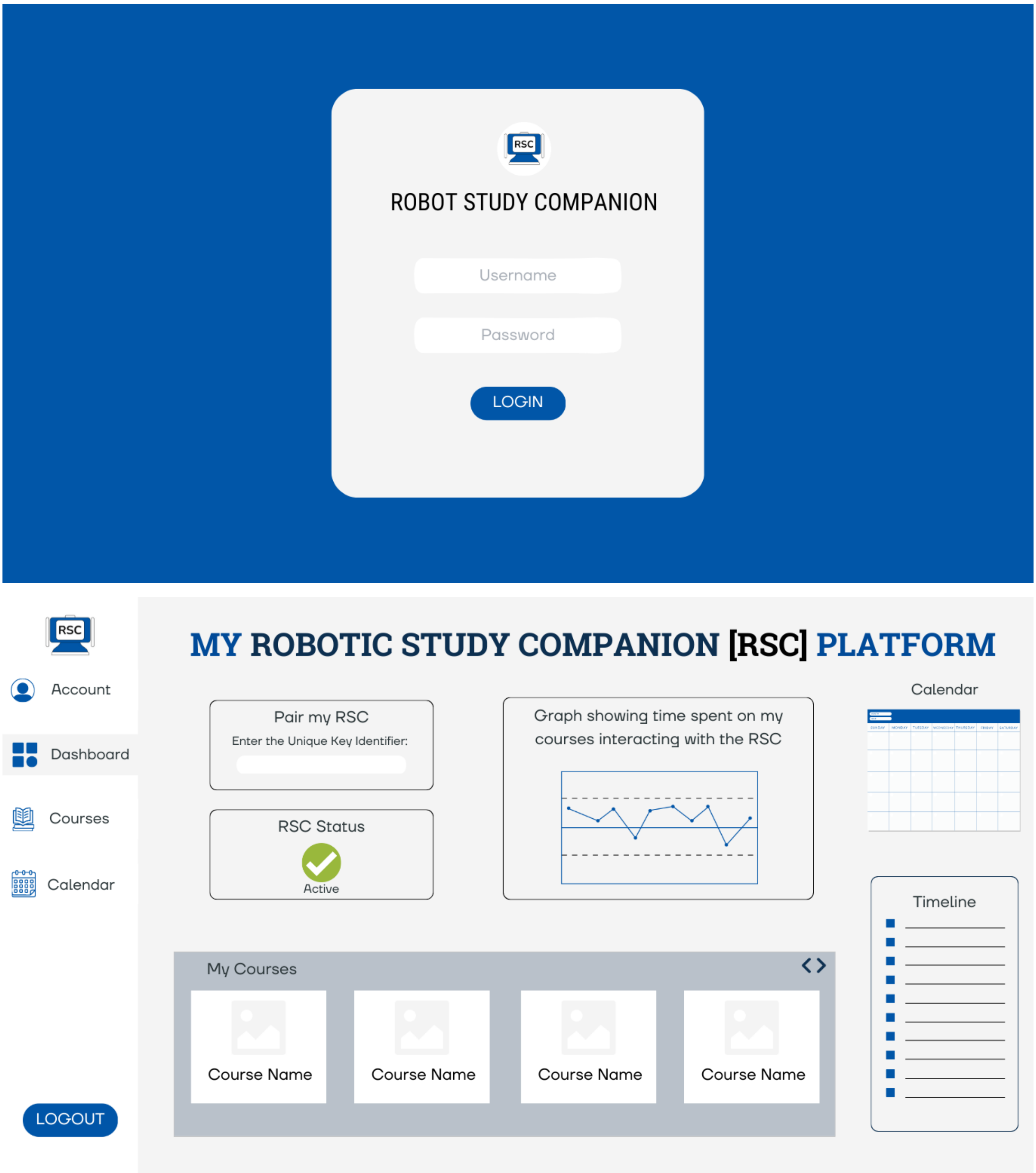


Figure 5.2: A work-in-progress mockup of the Web Application Interface for RSC platform. Login page and dashboard mockup for students.

The visual mockup showcases the intuitive design of the web app, featuring a side menu that provides convenient access to various functionalities. Users can easily navigate to their account settings, view their personalised dashboard, access enrolled courses that the RSC can assist with, and check upcoming exams and deadlines through the integrated calendar. The dashboard presents key elements, including the option to pair the robot, the RSC status updates, and a graphical representation of interaction time with the RSC.

5.2.1 Implementation & Evaluation

The potential implementation and evaluation of the RSC can be approached in several ways to assess its effectiveness and usability. Validating the RSC with university students across diverse learning environments and courses is crucial to gain insights into its real-world applicability. The primary focus of this evaluation should be on assessing the RSC's impact on students' study practices and their perception of its usefulness.

One approach is to conduct user testing with students, as demonstrated in previous studies [7][13]. Engaging students preparing for exams or studying for specific tests and providing them with the RSC as a supportive tool during their study sessions can yield valuable feedback and insights. By observing their interactions and gathering their perspectives, the extent to which the RSC enhances students' learning experiences can be evaluated.

Evaluating the RSC's impact on students' performance is an important aspect of its assessment. Measuring improvements in academic performance, such as grades and comprehension of course materials, can serve as indicators of its efficacy. Additionally, it is essential to gather feedback from university students, measure their acceptance of the RSC, and ensure its relevance and effectiveness in supporting their educational journey.

6 Conclusion

In summary, this thesis has focused on addressing the gap in social robots designed for university students by developing an open-source robotic study companion (RSC) platform. Through extensive background research, valuable insights were gathered from existing educational solutions, which aided in carefully formulating the requirements and guiding the design and development of the RSC. A prototype was constructed using off-the-shelf components and leveraging OpenAI's conversational API, which has demonstrated its potential to simplify complex concepts for students.

While the RSC prototype represents a significant achievement, there is still room for improvement in terms of the RSC's hardware and software, particularly to enhance its personalisation capabilities. Therefore, the RSC platform is currently under active development, and external contributions are welcomed to this open-source project. More information can be found in the project's repository on GitHub [70].

Acknowledgements

I would like to extend my sincere appreciation to my supervisors, Karl Kruusamäe, Sandra Schumann, and Agnes Luhtaru, for providing their valuable time, feedback, and guidance.

Special thanks to Teet Tilk for procuring hardware parts and dedicating time to help diagnose electronics issues. I would like to acknowledge Renno Raudmäe for obtaining the azure blue PLA and white polyflex filament; and Ulrich Norbisrath for sharing his knowledge on TTS & STT and providing a loaned RPi4 and white PLA filament.

Huge thanks to Matevž Zorec for input, assistance, and guidance in design, electronics, and software feedback. And a shout out to my family (parents & sister) for their support and willingness to listen during my moments of stress, despite being 5,710 miles away.

Grammarly [89] and ChatGPT [62] helped proofread and improve the writing style of this thesis. With the help of these tools, I corrected my usage from American to British English, identified instances of active or passive voice, and received suggestions to enhance clarity and readability. In specific sections like 4.4 Software Architecture, ChatGPT provided helpful suggestions for software diagrams and recommended terminology to effectively convey information. It also assisted me in understanding, differentiating, and organising Requirements (section 3) in a logical manner.

Finally, this thesis started with the letter "A" in the Introduction and now concludes with a "Z."

Farnaz Baksh
20.05.2023

References

[1] H. Mahdi, S. A. Akgun, S. Saleh, and K. Dautenhahn, “A survey on the design and evolution of social robots — Past, present and future,” *Robot. Auton. Syst.*, vol. 156, p. 104193, Oct. 2022, doi: 10.1016/j.robot.2022.104193.

[2] J. Kędzierski, P. Kaczmarek, M. Dziergwa, and K. Tchoń, “Design for a Robotic Companion,” *Int. J. Humanoid Robot.*, vol. 12, pp. 1550007–1, Jan. 2015, doi: 10.1142/S0219843615500073.

[3] J. Kanero, V. Geçkin, C. Oranç, E. Mamus, A. C. Küntay, and T. Göksun, “Social Robots for Early Language Learning: Current Evidence and Future Directions,” *Child Dev. Perspect.*, vol. 12, no. 3, pp. 146–151, 2018, doi: 10.1111/cdep.12277.

[4] “Learning French with Sophie – Personal Robots Group.” https://robots.media.mit.edu/portfolio/learning-french-sophie/ (accessed Apr. 26, 2023).

[5] D. Leyzberg, S. Spaulding, M. Toneva, and B. Scassellati, “The Physical Presence of a Robot Tutor Increases Cognitive Learning Gains”.

[6] D. Leyzberg, S. Spaulding, and B. Scassellati, “Personalizing robot tutors to individuals’ learning differences,” in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, in HRI ’14. New York, NY, USA: Association for Computing Machinery, Mar. 2014, pp. 423–430. doi: 10.1145/2559636.2559671.

[7] A. Riedmann, P. Schaper, and B. Lugrin, “Integration of a social robot and gamification in adult learning and effects on motivation, engagement and performance,” *AI Soc.*, pp. 1–20, Jun. 2022, doi: 10.1007/s00146-022-01514-y.

[8] H. Brock, J. Ponce Chulani, L. Merino, D. Szapiro, and R. Gomez, “Developing a Lightweight Rock-Paper-Scissors Framework for Human-Robot Collaborative Gaming,” *IEEE Access*, vol. 8, pp. 202958–202968, 2020, doi: 10.1109/ACCESS.2020.3033550.

[9] “ABii’s World by Van Robotics,” *ABii’s World by Van Robotics*. https://www.smartrobottutor.com (accessed May 08, 2023).

[10] S. S. Balasuriya, L. Sitbon, M. Brereton, and S. Koplick, “How can social robots spark collaboration and engagement among people with intellectual disability?,” *Proc. 31st Aust. Conf. Hum.-Comput.-Interact.*, pp. 209–220, Dec. 2019, doi: 10.1145/3369457.3370915.

[11] G. A. Papakostas *et al.*, “Social Robots in Special Education: A Systematic Review,” *Electronics*, vol. 10, no. 12, Art. no. 12, Jan. 2021, doi: 10.3390/electronics10121398.

[12] n.d., “Why do children with autism learn better from robots?,” *LuxAI S.A.*, Aug. 11, 2021. https://luxai.com/blog/why-children-with-autism-learn-better-from-robots/ (accessed May 08, 2023).

[13] M. Donnermann, P. Schaper, and B. Lugrin, “Integrating a Social Robot in Higher Education – A Field Study,” in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, Aug. 2020, pp. 573–579. doi: 10.1109/RO-MAN47096.2020.9223602.

[14] M. Ferguson, N. Webb, and T. Strzalkowski, “Nelson: a low-cost social robot for research and education,” in *Proceedings of the 42nd ACM technical symposium on Computer science education*, in SIGCSE ’11. New York, NY, USA: Association for Computing Machinery, Mar. 2011, pp. 225–230. doi: 10.1145/1953163.1953230.

[15] T. Belpaeme, J. Kennedy, A. Ramachandran, B. Scassellati, and F. Tanaka, “Social robots for education: A review,” *Sci. Robot.*, vol. 3, no. 21, p. eaat5954, Aug. 2018, doi: 10.1126/scirobotics.aat5954.

[16] C. Breazeal, K. Dautenhahn, and T. Kanda, “Social Robotics,” in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds., in Springer Handbooks. Cham: Springer

International Publishing, 2016, pp. 1935–1972. doi: 10.1007/978-3-319-32552-1_72.
[17] N. Reich-Stiebert, F. Eyssel, and C. Hohnemann, “Exploring University Students’ Preferences for Educational Robot Design by Means of a User-Centered Design Approach,” *Int. J. Soc. Robot.*, vol. 12, no. 1, pp. 227–237, Jan. 2020, doi: 10.1007/s12369-019-00554-7.
[18] N. Reich-Stiebert and F. Eyssel, “Robots in the Classroom: What Teachers Think About Teaching and Learning with Education Robots,” in *Social Robotics*, A. Agah, J.-J. Cabibihan, A. M. Howard, M. A. Salichs, and H. He, Eds., in Lecture Notes in Computer Science. Cham: Springer International Publishing, 2016, pp. 671–680. doi: 10.1007/978-3-319-47437-3_66.
[19] C. Vaudel, “NAO Robot V5 Open Box Edition for Schools.” https://www.robotlab.com/store/naov5-open-box-edition (accessed May 14, 2023).
[20] N. Reich-Stiebert and F. Eyssel, “Learning with Educational Companion Robots? Toward Attitudes on Education Robots, Predictors of Attitudes, and Application Potentials for Education Robots,” *Int. J. Soc. Robot.*, vol. 7, no. 5, pp. 875–888, Nov. 2015, doi: 10.1007/s12369-015-0308-9.
[21] J. Kennedy, S. Lemaignan, and T. Belpaeme, *The Cautious Attitude of Teachers Towards Social Robots in Schools*. 2016.
[22] S. Serholt *et al.*, *Teachers’ Views on the Use of Empathic Robotic Tutors in the Classroom*, vol. 2014. 2014. doi: 10.1109/ROMAN.2014.6926376.
[23] G. Ewijk, M. Smakman, and E. Konijn, *Teachers’ perspectives on social robots in education: an exploratory case study*. 2020, p. 280. doi: 10.1145/3392063.3394397.
[24] J. Han, E. Hyun, M. Kim, H.-K. Cho, T. Kanda, and T. Nomura, “The Cross-cultural Acceptance of Tutoring Robots with Augmented Reality Services.,” *JDCTA*, vol. 3, pp. 95–102, Jan. 2009.
[25] J. Ceha, E. Law, D. Kulić, P.-Y. Oudeyer, and D. Roy, “Identifying Functions and Behaviours of Social Robots for In-Class Learning Activities: Teachers’ Perspective,” *Int. J. Soc. Robot.*, vol. 14, no. 3, pp. 747–761, Apr. 2022, doi: 10.1007/s12369-021-00820-7.
[26] M. Obaid, W. Barendregt, P. Alves-Oliveira, A. Paiva, and M. Fjeld, *Designing Robotic Teaching Assistants: Interaction Design Students’ and Children’s Views*. 2015, p. 511. doi: 10.1007/978-3-319-25554-5_50.
[27] “JIBO - by DESIGN TEAM/HUGE DESIGN / Core77 Design Awards,” *Core77*. https://designawards.core77.com/home/award_permalink?id=34998&category_id=1 (accessed Apr. 27, 2023).
[28] “Cozmo vs Vector - 5 Differences between them, and which one is better. The answer might surprise you! - What is the difference between the two?” https://ankicozmorobot.com/cozmo-vs-vector/ (accessed May 01, 2023).
[29] “ElliQ, the sidekick for healthier, happier aging,” *ElliQ*. https://elliq.com/ (accessed May 01, 2023).
[30] “Eilik,” *Energize Lab*. https://store.energizelab.com/products/eilik (accessed May 01, 2023).
[31] R. Gomez, D. Szapiro, K. Galindo, and K. Nakamura, “Haru: Hardware Design of an Experimental Tabletop Robot Assistant,” in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, Chicago IL USA: ACM, Feb. 2018, pp. 233–240. doi: 10.1145/3171221.3171288.
[32] G. Yeung *et al.*, “A robotic interface for the administration of language, literacy, and speech pathology assessments for children”.
[33] A. Carman, “Jibo, the social robot that was supposed to die, is getting a second life,” *The Verge*, Jul. 23, 2020. https://www.theverge.com/2020/7/23/21325644/jibo-social-robot-ntt-disruptionfunding (accessed Apr. 27, 2023).
[34] “Haru: An Experimental Social Robot From Honda Research - IEEE Spectrum.”

https://spectrum.ieee.org/haru-an-experimental-social-robot-from-honda-research (accessed Feb. 26, 2023).
[35] "Meet Haru, the Unassuming Big-Eyed Robot Helping Researchers Study Social Robotics - IEEE Spectrum." https://spectrum.ieee.org/honda-research-institute-haru-social-robot (accessed Feb. 26, 2023).
[36] "Embodied Mediator for Cross Cultural Understanding among Children across the World," *Socially Intelligent Robotics Consortium*. https://mypersonalrobots.org/pilot (accessed Apr. 30, 2023).
[37] G. C. Gavero and A. P. S. Guerrero, "My Teacher Baymax: Lessons from the Film Big Hero 6," *Acad. Psychiatry*, vol. 41, no. 5, pp. 601–605, Oct. 2017, doi: 10.1007/s40596-017-0803-4.
[38] *Big Hero 6*, (2014).
[39] *Wall - E*, (2008).
[40] Marco, "Mira the robot waits for you to smile," *Personal Robots*, Sep. 29, 2019. https://www.personalrobots.biz/mira-the-robot-waits-for-you-to-smile/ (accessed May 01, 2023).
[41] 稚晖, "ElectronBot: Mini Desktop Robot." Apr. 30, 2023. Accessed: May 01, 2023. [Online]. Available: https://github.com/peng-zhihui/ElectronBot
[42] "Baymax," *Big Hero 6 Wiki*, May 02, 2022. https://bighero6.fandom.com/wiki/Baymax (accessed May 16, 2023).
[43] "WALL•E (character)," *Pixar Wiki*. https://pixar.fandom.com/wiki/WALL%E2%80%A2E_(character) (accessed May 16, 2023).
[44] M. I. Ahmad, M. Khordi-moodi, and K. S. Lohan, "Social Robot for STEM Education," in *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, in HRI '20. New York, NY, USA: Association for Computing Machinery, Apr. 2020, pp. 90–92. doi: 10.1145/3371382.3378291.
[45] "Mira the Baby Robot – Background Research – 16-223 Work." https://courses.ideate.cmu.edu/16-223/f2018/work/2018/09/12/mira-the-baby-robot-background-research/ (accessed May 01, 2023).
[46] S. world, "ElliQ is an emotionally intelligent, intuitive robot sidekick for the ageing." https://www.stirworld.com/see-features-elliq-is-an-emotionally-intelligent-intuitive-robot-sidekick-for-the-ageing (accessed May 01, 2023).
[47] C. Caudwell, C. Lacey, and E. B. Sandoval, "The (Ir)relevance of Robot Cuteness: An Exploratory Study of Emotionally Durable Robot Design," in *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*, in OzCHI '19. New York, NY, USA: Association for Computing Machinery, Jan. 2020, pp. 64–72. doi: 10.1145/3369457.3369463.
[48] I. Fadelli and T. Xplore, "A robotic planner that responds to natural language commands." https://techxplore.com/news/2020-03-robotic-planner-natural-language.html (accessed Dec. 22, 2022).
[49] N. Radziwill and M. Benton, "Evaluating Quality of Chatbots and Intelligent Conversational Agents".
[50] J. Weizenbaum, "ELIZA—a computer program for the study of natural language communication between man and machine," *Commun. ACM*, vol. 9, no. 1, pp. 36–45, Jan. 1966, doi: 10.1145/365153.365168.
[51] R. S. Albornoz-De Luise, P. Arnau-González, and M. Arevalillo-Herráez, "Conversational Agent Design for Algebra Tutoring," in *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Oct. 2022, pp. 604–609. doi: 10.1109/SMC53654.2022.9945524.
[52] "Conversational AI Platform | Superior Customer Experiences Start Here," *Rasa*, Dec. 01, 2020. https://rasa.com/ (accessed May 15, 2023).

[53] "Voice Kit," *Google AIY Voice Kit*. https://aiyprojects.withgoogle.com/ (accessed May 02, 2023).
[54] P.-C. Lin, B. Yankson, V. Chauhan, and M. Tsukada, "Building a speech recognition system with privacy identification information based on Google Voice for social robots," *J. Supercomput.*, vol. 78, no. 13, pp. 15060–15088, Sep. 2022, doi: 10.1007/s11227-022-04487-3.
[55] "Dialogflow," *Google Cloud*. https://cloud.google.com/dialogflow (accessed May 15, 2023).
[56] K. Katchapakirin and C. Anutariya, "An Architectural Design of ScratchThAI: A conversational agent for Computational Thinking Development using Scratch," in *Proceedings of the 10th International Conference on Advances in Information Technology*, in IAIT 2018. New York, NY, USA: Association for Computing Machinery, Dec. 2018, pp. 1–7. doi: 10.1145/3291280.3291787.
[57] "OpenAI." https://openai.com/ (accessed May 15, 2023).
[58] T. B. Brown *et al.*, "Language Models are Few-Shot Learners." arXiv, Jul. 22, 2020. Accessed: May 15, 2023. [Online]. Available: http://arxiv.org/abs/2005.14165
[59] "Introducing ChatGPT and Whisper APIs." https://openai.com/blog/introducing-chatgpt-and-whisper-apis (accessed May 15, 2023).
[60] "Aligning language models to follow instructions." https://openai.com/research/instruction-following (accessed May 20, 2023).
[61] L. Ouyang *et al.*, "Training language models to follow instructions with human feedback." arXiv, Mar. 04, 2022. doi: 10.48550/arXiv.2203.02155.
[62] "Introducing ChatGPT." https://openai.com/blog/chatgpt (accessed May 20, 2023).
[63] "OpenAI API." https://platform.openai.com (accessed May 15, 2023).
[64] "Axios Login: Khan's AI tutor," *Axios*. https://www.axios.com/newsletters/axios-login/ (accessed May 15, 2023).
[65] J. vom Brocke, A. Hevner, and A. Maedche, "Introduction to Design Science Research," 2020, pp. 1–13. doi: 10.1007/978-3-030-46781-4_1.
[66] "Servo Motor | 3D CAD Model Library | GrabCAD." https://grabcad.com/library/servo-motor-101 (accessed May 06, 2023).
[67] "Raspberry Pi 4 Model B | 3D CAD Model Library | GrabCAD." https://grabcad.com/library/raspberry-pi-4-model-b-1 (accessed May 06, 2023).
[68] "Adafruit NeoPixel Ring 16 | 3D CAD Model Library | GrabCAD." https://grabcad.com/library/adafruit-neopixel-ring-16-1 (accessed May 06, 2023).
[69] E. Liberman-Pincu, Y. Parmet, and T. Oron-Gilad, "Judging a Socially Assistive Robot by Its Cover: The Effect of Body Structure, Outline, and Color on Users' Perception," *ACM Trans. Hum.-Robot Interact.*, vol. 12, no. 2, p. 23:1-23:26, Apr. 2023, doi: 10.1145/3571717.
[70] "RoboticStudyCompanion," *GitHub*. https://github.com/Farnaz03/RoboticStudyCompanion (accessed May 18, 2023).
[71] "Buy a Raspberry Pi 4 Model B," *Raspberry Pi*. https://www.raspberrypi.com/products/raspberry-pi-4-model-b/ (accessed May 06, 2023).
[72] "AIY Voice Bonnet at Raspberry Pi GPIO Pinout." https://pinout.xyz/pinout/aiy_voice_bonnet (accessed May 06, 2023).
[73] "Voice Kit overview — AIY Projects 2018-11-16 documentation." https://aiyprojects.readthedocs.io/en/latest/voice.html#voice-bonnet-voice-kit-v2 (accessed May 06, 2023).
[74] A. Industries, "NeoPixel Ring - 16 x 5050 RGB LED with Integrated Drivers." https://www.adafruit.com/product/1463 (accessed May 06, 2023).
[75] C. Castro and E. Ramos, "Mood Lamps," in *Arduino and Kinect Projects: Design, Build, Blow Their Minds*, E. R. Melgar, C. C. Díez, and P. Jaworski, Eds., Berkeley, CA: Apress, 2012, pp. 133–165. doi: 10.1007/978-1-4302-4168-3_7.

[76] divilife, “What is Mood Lighting? Mood Lighting vs. Lighting for Your Mood,” *BIOS Lighting*, Dec. 06, 2021. https://bioslighting.com/what-is-mood-lighting/wellness-lighting/ (accessed May 06, 2023).
[77] Patrick, “NX3224T024,” *Nextion*. https://nextion.tech/datasheets/nx3224t024/ (accessed May 06, 2023).
[78] Patrick, “The Nextion Editor Guide,” *Nextion*. https://nextion.tech/editor_guide/ (accessed May 06, 2023).
[79] “DFRobot DF9GMS 360 Degree Micro Servo (1.6Kg).” https://core-electronics.com.au/towerpro-sg90c-360-degree-micro-servo-1-6kg.html (accessed May 06, 2023).
[80] “Releases · google/aiyprojects-raspbian,” *GitHub*. https://github.com/google/aiyprojects-raspbian/releases (accessed May 20, 2023).
[81] “AIY Projects.” Google, May 18, 2023. Accessed: May 20, 2023. [Online]. Available: https://github.com/google/aiyprojects-raspbian
[82] “Python Release Python 3.7.3,” *Python.org*. https://www.python.org/downloads/release/python-373/ (accessed May 20, 2023).
[83] A. Zhang (Uberi), “SpeechRecognition: Library for performing speech recognition, with support for several engines and APIs, online and offline.” Accessed: May 20, 2023. [MacOS :: MacOS X, Microsoft :: Windows, Other OS, POSIX :: Linux]. Available: https://github.com/Uberi/speech_recognition#readme
[84] “PyAudio: Cross-platform audio I/O with PortAudio.” Accessed: May 20, 2023. [Online]. Available: https://people.csail.mit.edu/hubert/pyaudio/
[85] N. M. Bhat, “pyttsx3: Text to Speech (TTS) library for Python 2 and 3. Works without internet connection or delay. Supports multiple TTS engines, including Sapi5, nsss, and espeak.” Accessed: May 20, 2023. [MacOS :: MacOS X, Microsoft :: Windows, POSIX]. Available: https://github.com/nateshmbhat/pyttsx3
[86] “OpenAI Python Library.” OpenAI, May 20, 2023. Accessed: May 20, 2023. [Online]. Available: https://github.com/openai/openai-python
[87] “CircuitPython Essentials,” *Adafruit Learning System*. https://learn.adafruit.com/circuitpython-essentials/circuitpython-neopixel (accessed May 20, 2023).
[88] “Pololu - Logic Level Shifter, 4-Channel, Bidirectional.” https://www.pololu.com/product/2595 (accessed May 06, 2023).
[89] “How Does Grammarly Work? | Grammarly Spotlight,” *How To Use Grammarly | Grammarly*, Dec. 04, 2019. https://www.grammarly.com/blog/how-does-grammarly-work/ (accessed May 20, 2023).

Non-exclusive licence to reproduce thesis and make thesis public

I, Farnaz Baksh

1. grant the University of Tartu a free permit (non-exclusive licence) to

reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright, my thesis

AN OPEN-SOURCE ROBOTIC STUDY COMPANION FOR UNIVERSITY STUDENTS

supervised by Karl Kruusamäe, Sandra Schumann, Agnes Luhtaru

2. I grant the University of Tartu a permit to make the thesis specified in point 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 4.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in points 1 and 2.

4. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Farnaz Baksh
20.05.2023