KLM Virtual Reality and Adaptive AI

Feasibility Study of The Use Of Adaptive AI For Enhanced Employee Training

Joint Interdisciplinary Project

Team 1.4.1

KLM Virtual Reality and Adaptive AI

Feasibility Study of The Use Of Adaptive AI For Enhanced Employee Training

by

Team 1.4.1

Student Name	Student Number
Deepak Sogasu	5792355
Ehsan Sadegh	4597052
Henk Jekel	5609593
Kshitij Sreenivasa Roopa	5795591
Ramon Mozaffarian	4964470
Ujjayan Dhar	5791073

Company Coach:	Jae Maloney
JIP Programme Manager:	Birgit de Bruin
Project Duration:	Sept. 2023 - Nov. 2023

Cover: KLM Royal Dutch Airlines

TU Delft

KLM

Management Summary

In an era where technology is a driving force behind innovation, a multidisciplinary team at Delft University of Technology recognized the potential to harness the power of adaptive AI to modernize training methods. This report presents the project progress of an interdisciplinary team at TU Delft conducted as part of the program - Joint Interdisciplinary Project (JIP). The team includes second-year master's students working in collaboration with KLM Royal Dutch Airlines, one of the oldest and most innovative airlines in the world. The initiative focused primarily on the feasibility of redesigning the training environment of the KLM sales department using adaptive AI. The journey began with an initiation phase that laid a solid foundation for collaboration and an introspection of KLM's existing training methods. This was followed by a comprehensive research phase to test the feasibility of an adaptive AI phone application to train the sales employees at KLM. Before testing the feasibility of the phone application for the employee training task, the prototype underwent rigorous testing to ensure its effectiveness. The prototype was carefully developed in collaboration with AI experts and feedback from people outside of the development team played a key role in refining the prototype. The final feasibility test had an emphasis on qualitative and quantitative indicators to measure its feasibility.

The product successfully demonstrated the transformative capabilities of adaptive AI in training, as reflected in its overwhelmingly positive reviews. Its ability to engage in in-depth conversations, simulate human-like voices, provide unpredictability, and deliver near-real-time responses underscored the feasibility of the project. This compelling outcome prompted KLM to extend their research efforts through a Master's thesis. In light of the results from the feasibility test, the forthcoming research will focus on immediate enhancements, including improvements in audio quality, refining response flow for a smoother user experience, enabling user interruptions, and reducing latency. It's worth noting that a significant limitation of the application lies in its reliance on third-party solutions, emphasizing the need for future research to explore the development of in-house alternatives to mitigate this dependency.

This project goes beyond the mere introduction of new training methods; its ambition lies in setting pioneering standards, which encompass achieving a new level of unpredictability during training and fostering high levels of trainee motivation. The team envisions a revolutionary transformation in how sales training is both perceived and delivered at KLM. The overarching objective is to extend this paradigm shift across various sectors within the company.

The strategic roadmap that was formulated stands as the cornerstone of this transformative expedition. The fusion of adaptive AI and VR capabilities paves the way for a training environment at KLM that is ready to meet the demands of the future. The triumph of this project hinges on the ongoing involvement of stakeholders, meticulous testing, and an unyielding dedication to innovation. While this journey may present its share of challenges, the presence of a well-defined roadmap and a visionary outlook assures that the destination holds the potential for significant transformation.

Contents

List of Figures

List of Tables

1 Introduction

The Joint Interdisciplinary Project (JIP) is a TU Delft program that allows Master students to cooperate with firms and work on real-world issues. KLM and the JIP are collaborating on this project to conduct company-oriented research.

KLM was founded in 1919 and is the world's oldest airline that still operates under its original name. KLM operates three core activities from its hub at Schiphol Airport: Network (passenger and cargo activity), Engineering & Maintenance, and Leisure.KLM, KLM Cityhopper, and Transavia fly leisure and business travelers to 168 European destinations and 80 Intercontinental locations. KLM is also a SkyTeam alliance member, serving 1,062 destinations in 170 countries. AIR FRANCE KLM Martinair operates the Cargo activity, while Engineering & Maintenance provides services to KLM and third parties. In 2004, KLM combined with Air France to form the AIR FRANCE KLM Group. (KLM, 2022). (CFReport - KLM Annual Report 2022)

KLM, a 104-year-old aviation firm, has emerged as an industry pioneer in digital innovation and IT. Despite its long history, KLM has remained committed to exploring new frontiers. KLM's BlueLabs section has been exploring the intriguing areas of Virtual Reality (VR) and Augmented Reality (AR) since its first flight in 1920. This modernization push entails integrating cutting-edge technology such as AR and VR to improve the airline's products and services (KLM). KLM's innovation laboratories are leading the way in altering the aviation industry. KLM is pushing industry change with projects such as automated 3D cargo scanning, autonomous robotic operations, and the development of smart workspaces. Individuals inside the organization are urged to serve as the guiding radar, imagining the future and contributing to the Minimum Viable Products (MVPs) that pave the path for that future. One significant use of VR is in training programs, which provide an invaluable experience that cannot be replicated in real life. KLM's cabin crew, for example, employs virtual reality to rehearse fire safety procedures in a simulated steel aircraft kitchen, demonstrating the airline's commitment to cutting-edge training methods. More information can be found here.

The company is currently exploring additional VR applications, with a specific emphasis on the potential benefits of VR in the realm of employee training. This report presents a feasibility study that investigates the potential of using adaptive AI within a phone applications to train sales professionals. Presently, sales professionals in their training often face a lack of enthusiasm and motivation stemming from the predictability and monotony of their current phone-based training methods. The primary objective of this research is to formulate an alternative training approach that enhances the confidence, motivation, and overall competence of these sales professionals, ultimately leading to improved sales interactions with B2B consumers. The goal of this feasibility study is to demonstrate the viability of achieving this with the integration of adaptive AI. The incorporation of the unpredictable adaptive AI integrated into the phone application is expected to boost the confidence of sales professionals when confronted with unforeseen queries during the process of selling tickets to business partners.

1.1. The Interdisciplinary Team

The interdisciplinary team comprising of TU Delft Master students is working with KLM's IT services team and the Sales team to develop a feasibility assessment for the Learning and Development team's theory.The Joint Interdisciplinary Project (JIP) of project 1.4.1 team consists of 6 team members as follows:

Figure 1.1: Kshitij S R - MSc. Management of Technology

Figure 1.2: Ujjayan Dhar - MSc. Strategic Product Design

Figure 1.3: Ramon Mozaffarian - MSc. Complex Engineering and Management

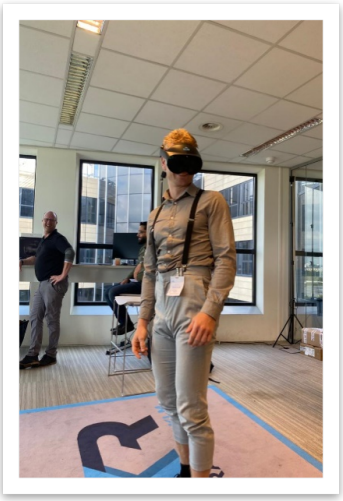


Figure 1.4: Henk Jekel - MSc. Robotics

Figure 1.5: Ehsan Sadegh - MSc. Complex Engineering and Management

Figure 1.6: Deepak Sogasu - MSc. Aerospace Engineering

Kshitij, Ujjayan, and Ramon, with their expertise in project management, concentrated their efforts on the literature study and the qualitative and quantitative analysis of the project. Ehsan and Henk, possessing strong coding skills, dedicated their time to the product development. Deepak played a pivotal role in both groups, acting as a communication bridge to ensure that each group was well-informed about their respective roles.

1.2. Objective and Purpose of Study

The objective of this research is to assess the feasibility and potential impact of incorporating adaptive AI into KLM's existing training programs. The aim is to enhance the quality of training for sales representatives and, concurrently, to boost employee confidence and motivation for their real-world responsibilities. Upon establishing feasibility, the intention is to extend this concept to other KLM departments, including the cabin crew.

KLM encompasses various sectors that necessitate specialized training programs, spanning sales, maintenance, and innovation. This initiative is centered on the "Learning and Development" segment within the sales department, with the aspiration to broaden its reach to encompass other sectors in the future. The initial phase of the research, as documented in this report, delves into the examination of the feasibility of implementing an adaptive AI phone application to replace the existing sales training program.

In future research phases, the foundational components of the application can be repurposed and integrated into other training applications or support tools, such as a personal desk AI assistant. This assistant, featuring an emotionally expressive avatar, has the potential for greater immersive capabilities when compared to the phone application. Potential use cases encompass its utilization as a training tool, as well as a more intuitive search engine for desk employees. This allows them to use speech and visual cues for information retrieval, enhancing their overall user experience. Another pathway for expansion, capitalizing on the foundational elements of the phone application, revolves around creating pioneering VR training scenarios for departments like the cabin crew. This initiative aims to bolster their readiness in this particular domain.

Our primary objective is to evaluate the feasibility of an intuitive phone application for sales training, powered by adaptive AI. If feasibility is confirmed, it is anticipated that KLM will undertake further research, resulting in an application that significantly enhances the motivation and preparedness of sales employees. This, in turn, is expected to lead to increased customer satisfaction within the sales department and large cost savings for training. Once the concept is proven in the sales department, the application is anticipated to be extended to all sectors within KLM, leading to increased customer satisfaction and business growth across all departments among KLM employees.

1.3. Problem Statement

In a fast-changing airline industry, effective employee training is key to success. Traditional methods are costly and limited in scope. This JIP team collaborates with KLM to modernize training by blending adaptive AI into virtual programs. This approach aims for more unpredictable scenario training, improving real-world preparedness while lowering costs.

The problem defined for this research is how to enhance employee training programs at KLM, starting with the Learning and Development sector within sales. The objective is to investigate the potential of integrating adaptive AI into the current standardized and predictable training methods with the aim of enhancing training quality through increased immersion and unpredictability. The aim is to create intuitive, easy-to-set-up training modules that incorporate unpredictable scenarios, enhancing employee motivation and confidence, while maintaining the believability of the training.

1.4. Stakeholders Analysis

With a proof of concept of the phone application in the sales department, the phone application can be deconstructed, and its fundamental building blocks can be integrated into 3D tablet applications, such as an AI assistant on a personal desk. This assistant can assist in delivering training and information to employees with the help of an emotionally expressive avatar, thereby constructing a simulation closely mirroring reality. Additionally, it can be employed within the customer segment to aid customers in finding the information they require. The fundamental building blocks can also be leveraged to develop an innovative VR application that simulates training scenarios for the cabin crew department, enhancing their immersion, believability, unpredictability in training, and their preparedness for real-life situations.

- **KLM Research Department:** Responsible for providing critical data, insights, and validation for the project. They serve as an essential bridge between theoretical concepts and practical implementation in the VR-AI training system.
- **KLM IT Department:** Subdivided into VR team, Development team, and AI team, they are responsible for the technical architecture and development of the VR-AI integrated training system.
- **L&D Department:** Oversee the training curriculum and are key in integrating the VR-AI system into the existing training programs.
- **KLM Operational Employees:** Including cabin crew and sales crew, these are the end-users of the VR-AI system whose feedback is crucial for system optimization.
- **KLM Higher Management:** Decision-makers in terms of budget allocation and overall project approval.
- **Customer or Clients:** Include B2B partners, travel agencies (B2T), and individual passengers (B2C), who are the ultimate recipients of the enhanced service quality that the training aims to achieve.
- **JIP Team,** is the bridging student group between TU Delft and the KLM environment:They collaborate with a company coach from KLM along with their services to successfully complete the project.
- **TU Delft,** the institution supporting JIP team and maintains contact with the company throughout the project period.

As KLM Royal Dutch Airlines continues its mission to revolutionize employee training through virtual reality and adaptive AI, a range of stakeholders are collaborating to shape its trajectory. Leading this initiative, KLM's research department serves as the intellectual backbone, providing essential data and insights, while bridging the gap between innovative ideas and their tangible realization. KLM's diverse IT teams provide the technological power, the VR team ensures great experiences, while the development and AI teams design and fine-tune the system's backbone. L&D plays a key role in seamlessly integrating the new VR-AI system into the existing training structure, ensuring that content remains relevant and effective. On the forefront, KLM operations staff, including cabin crew and sales staff, interact directly with this advanced training platform. Their valuable opinions act as a compass for system improvement. KLM senior management members guide the

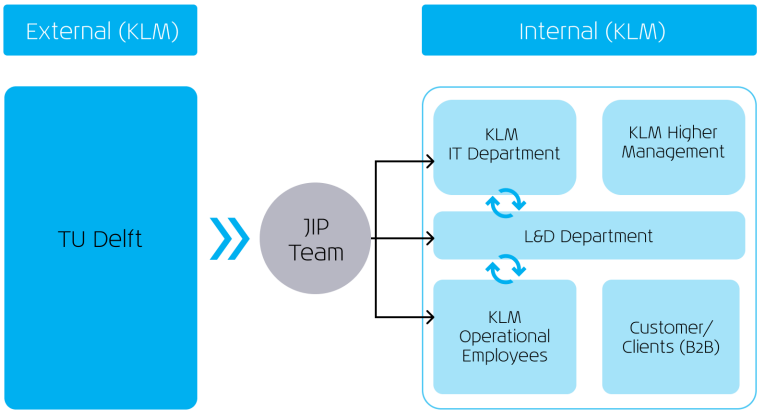


Figure 1.7: Stakeholders involved in the project.

strategic and financial direction of the project, with their decisions shaping the course of the initiative and ensuring alignment with organizational goals. Ultimately, the beneficiary of this transformative training approach is KLM's broad customer base, including B2B partners, travel agencies (B2T) and individual passengers (B2C). The potentially improved service experience speaks to the success of the project and highlights the importance of comprehensive stakeholder engagement.

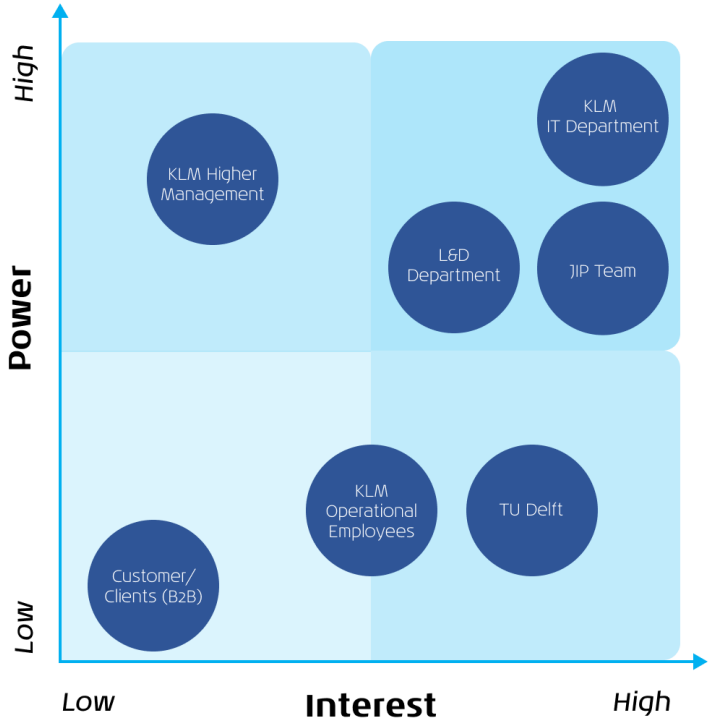


Figure 1.8: Stakeholder power grid

1.5. Project Plan

The main goal is to design, develop & deliver an adaptive AI phone application to create randomized, realistic scenarios for sales representatives in a B2B environment. Employees can interact with AI-driven personas and scenarios to prepare for their actual sales conversations, enhancing the training's believability, thus enhancing skills and confidence in real-world situations. The resulting conversation with the AI-bot provides a highly realistic and immersive training experience.

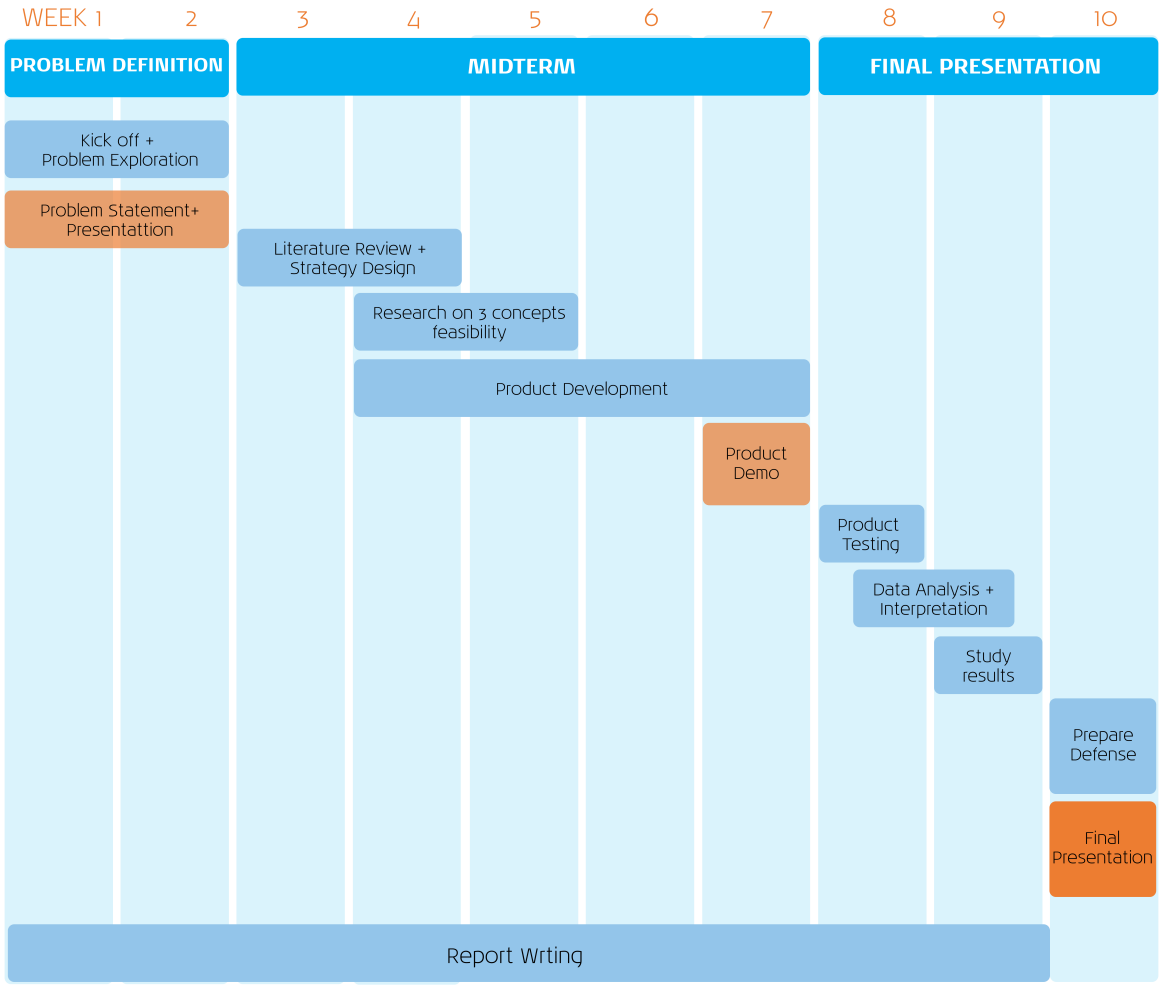


Figure 1.9: Project Planning Timeline

This concept involves integrating AI algorithms that adjust scenarios based on user interactions, providing personalized, realistic, and unpredictable training experiences tailored to user interactions. This fosters deeper engagement and skill retention.The initiative aims to optimize training effectiveness, increase employee engagement and encourage innovative approaches in training techniques.

- **Phase 1: Groundwork and Initial Development:** During the foundational phase, our focus was on rigorous research and analysis. A deep dive into the literature surrounding AI and VR provided insights into potential applications. Immersion in VR experiences further shaped our understanding. A cross-functional team of programmers, UX experts, data analysts, and other relevant experts was assembled. Their

collaborative approach drove the initial development of the product, ensuring it exposed core functionality and reduced target latency. At the same time, stakeholder engagement is paramount, with regular team meetings and feedback sessions ensuring alignment with the project vision.

- **Phase 2: Enhancement and Integration:** As we move into the second phase, we placed more emphasis on accelerating development. We strived to refine AI algorithms, further reduce response latency, and integrate the AI bot seamlessly into phone calling features. A unique feature under consideration was the establishment of a phone pipeline, facilitating users to interact with the AI bot via voice calls. In tandem with technical advancements, feedback and refinements were crucial. Post-training sessions involved comprehensive feedback collection, ensuring the product evolved based on user insights. Engaging internal and external stakeholders also played a pivotal role, ensuring that product advancements aligned with industry expectations and user demands.

- **Phase 3: Finalization and Reporting:** As we move into the final phase, we focused upon finalizing the product and thoroughly documenting it. The AI system has been fully integrated with phone call app, providing a comprehensive training experience. A comprehensive report detailing the project journey, its innovations, challenges and outcomes has been prepared. This report demonstrates our commitment to transparency, detailing every step of our journey. Stakeholder engagement remains essential, ensuring ongoing communication and addressing any concerns. In addition, post-project evaluation and feedback mechanisms by use of qualitative & quantitative analysis that included questionnaries, surveys, & so other methods assessed the effectiveness and future scalability of the project.

Expected results: At the culmination of these stages, we envisioned a cutting-edge training solution that aligned perfectly with KLM's overall goals. Through this effort, we anticipate increased employee engagement in training, which will translate into tangible organizational growth. Equally important, we aim to strengthen relationships with all stakeholders, appreciating their ongoing engagement and feedback.

Conclusion: This roadmap summarizes our vision, strategy and commitment to providing outstanding training solutions for KLM. By embracing innovation, fostering collaboration and valuing stakeholder engagement, we are not only poised to achieve our project goals but also redefine industry standards for training solutions.

2 Literature Review

2.1. State of the art: VR

VR is broadly described as an environment in which individuals may naturally accept and respond to artificial inputs. Other definitions of VR include the human-machine interface, which lets users "project" themselves into a computer-generated world in order to achieve specific goals. (Zhang, 2014). Depending on the configuration of the human-machine interface, the hardware components, and the number of real-world images placed into the virtual world, a user's experience will differ between the various types of mixed reality such as augmented reality (AR), augmented virtuality (AV), mirror reality (MR), and virtual reality (VR) (Cochrane et al., 2017), (Tacgin & Arslan, 2017), (J. C. Brendan & Mary, 2019).

Kirner and Siscoutto defined Virtual Reality (VR) as an improved user interface for accessing computer-based applications, providing real-time display, handling, and user interaction in three-dimensional computer-generated worlds. The sense of sight is typically the primary sense in virtual reality applications, but additional senses such as touch and sound, among others, can also be used to improve the user experience. (Frigo and Barbosa, 2016)

Advanced lighting and shadowing techniques are used to improve the user's spatial awareness within the virtual environment, as well as a complete dynamics simulation facility that governs the interaction of all bodies within the environment, including sensor-attached anthropomorphic parts. The modeling of maintenance chores in a virtual world is primarily reliant on people gaining environment awareness through intuitive and natural interaction (Abate et al., 2009).

2.2. Impacts

2.2.1. Medical

Surgical simulators based on AR and MR have been developed in addition to VR surgical simulators. Advanced techniques and algorithms have enabled surgeons to perform correct surgical procedures, intraoperative control, and postoperative tracking. However, AR and MR are clearly preferred for surgical planning and intraoperative assistance.

VR, on the other hand, is employed for its immersive nature and is hence more suited for training reasons. The connection between VR, AR, and MR is that all three technologies employ virtual data to affect the physical world around the user. (R. Agha, 2015) further emphasize that simulators allow trainees to improve their skill set more sequentially and at a rate that is uniquely adjusted, which would not be achievable with a real patient.

2.2.2. Design, Manufacturing and Maintenance of Products

Industry 4.0 began as a concept suggested by the German government, academic institutions, and private businesses; it aspires to leverage sophisticated technology and concepts to transition industrial manufacturing into a new stage known as smart manufacturing. Its goal is to incorporate innovative technology to improve the manufacturing industry's quality, productivity, and effectiveness.

If a designer can intuitively feel and even interact with a virtual prototype of a product during the early design stage, rather than using his or her imagination or an animation, the productivity of the maintainability design process would substantially improve. (Guo et al., 2020).

As per Gong et al., 2019, Virtual reality (VR) technology that is both affordable and powerful is gaining acceptance in the manufacturing business, notably in factory layout planning (FLP). Layouts that are well-designed can cut operational expenses by up to 50%. Virtual reality-supported layout planning (VLP) is gaining popularity because it allows designers to readily examine "what if" scenarios in a virtual environment, boosting efficiency and product complexity management.

2.2.3. Mining Industry

Traditionally, training has relied mostly on video and handbooks as media. Mine managers will organize novices to view instructional videos and master handbooks before practicing in a real underground environment. Simulated environment: VR-based safety training can provide trainees with an immersive virtual world with physical qualities, and the simulation technology's instantaneous and physics feedbacks make interactive learning more realistic.The use of adaptive technology allows the training difficulty level to be dynamically modified based on the user's performance, avoiding learner abandoning of the training work due to high difficulty (Z. Liang & Gao, 2019).

2.2.4. Aerospace

Aerospace design is a potentially attractive application where technologies that offer faster design cycle times with potentially greater efficiency improvements can be game changers. However, because the aerospace sector typically works with cutting-edge computational tools and sophisticated computer-aided design software, there are significant barriers to getting the community to accept VR (S. K. Tadeja & Kristensson, 2019). As per Pradhan et al., 2023, the rapid expansion of the Urban Air Mobility (UAM) business needs the introduction of Virtual Reality (VR) because to the rising interest in technology. As a result, instead of dealing with physical controls, everything will be digital and fictitious, and the flight mechanics and dynamics of the aircraft will be displayed on the screen using Human-Machine Interfaces and Interactions (HMI2) concepts via symbology design while projecting the orientation and position of the UAM aircraft. Furthermore, when compared to earlier technologies, this new technology will significantly shorten the time required to teach UAM pilots because the integrated virtual system can be used to train the pilots from start to finish.

2.2.5. Military

Military training utilizing virtual reality was featured in a list of seven critical technologies that will assure the US army's dominance in the twenty-first century. Furthermore, the United States Defense Department warned that virtual reality (VR) technology might transform the notions and methods of military utility. Virtual reality (VR) is used in various domains in the military, including virtual training, virtual weapon manufacturing and design, virtual combat exercises, marine applications, warfare, and engineering design. The precision and striking quality of those re-enactment frameworks have been greatly increased, and the level of recreation of the picture has also been nearly indistinguishable from the genuine (Ahir et al., 2020).

2.2.6. Entertainment

Zyda, 2005 mentions Beginning in 1997, with the release of an NRC paper titled "Modeling and Simulation–Linking Entertainment and Defense,"2 the video gaming community has moved into previously uncharted territory for the VR community. Clearly, the VR sector is shifting into employment that is influenced by video games and consequently influences that industry as well. Because much of the research and development being done in the gaming community tracks the work of the VR community, it has the potential to touch a larger audience. With the current surge in virtual reality (VR) content, the creation and consumption of virtual avatars may become mainstream in the near future. Embodied virtual avatars are a visual depiction of your own body viewed through an immersive head-mounted display (HMD) that moves in rhythm with your own body motions and occupies the same space as your own body. These findings demonstrate how virtual avatars can influence users' perceptions and behaviors (Maloney et al., 2019).

2.2.7. Training Challenges in the Airline Industry

This research looks on the aviation industry, which is at the forefront of augmented and virtual reality programs in corporate training. Flight attendants are the people in charge of passenger safety and security on commercial aircraft. In this study, in-flight virtual reality training is compared against in-flight real-life simulation training to determine whether virtual reality has a beneficial or negative impact on knowledge transfer.

Air New Zealand proposed utilizing an augmented reality device to recognize a key on-board passenger and then retrieve and augment their information for the flight crew. (Bellamy, 2017). The suggested application is Inflight, which was created for the HoloLens to offer critical passenger information to crew members such as flight attendants.

Unlike in the 2018 pilot study, flight attendant speed is not monitored in the behavior drills. The behavior exercises assess the correctness of occupational skills such as door security and emergency equipment management. While the pilot study monitored change in response time effectively, it did so while the pilot was in the training setting. In the flight attendant study, the behaviors observed after training were measured rather than the behaviors observed during training (Rosellini, 2023).

2.2.8. Integration of Adaptive AI in VR Training

Data from planes is beginning to influence how pilots are trained. For example, decisions about pitch and throttle settings during takeoff and climb can have a significant impact on overall fuel usage, which affects both the environment and the airline's bottom line. Because everyone learns at a different rate and has varied experiences and backgrounds, we call this approach adaptable training. Previously, a one-size-fits-all strategy was the norm, but technology is enabling more individualized ways. And the advantages might be substantial. AI-powered adaptive training will make the sector safer while also making training more affordable. The expenditures made by airlines and training institutions in this technology will pay off in the long run, since training demands will be cut by 30-50% (Klassen, 2021).

As the aviation industry experiences rapid technological advancement, new opportunities for improving training processes and optimizing performance results emerge. This in-depth examination investigates the impact of technology integration on the efficacy of aviation instruction, learner involvement, and skill retention. It investigates a wide range of research, focusing on major topics such as immersive technologies, artificial intelligence, game-based education, remote and cooperative training, and the long-term effects of technology-enhanced training. These technologies might assess students' performance and provide real-time feedback, adjusting training content and difficulty levels as needed. Future study should investigate the possible benefits and obstacles of integrating AI-driven instruction systems in aviation, as well as their effectiveness (DNÇER, 2023).

Novel techniques for algorithm verification and validation, as well as usability testing with typical direct users and indirect stakeholders, are required for AI and machine learning algorithms included in Human-centered AI (HCAI) systems. The goal is to increase the likelihood that the HCAI system will perform as expected by users while decreasing the risk of unexpectedly damaging effects. Civil aviation provides strong models for new design approval, rigorous verification and validation testing during early and ongoing use, and pilot certification testing (Shneiderman, 2020).

2.2.9. Best Practices for Integration

As per European Union Aviation Safety Agency (EASA) Operational domain in which an ML-based system could be implemented/used. ATM/ANS, air operations, flight crew training, environmental protection, and aerodromes are some of the topics covered by this guideline. Crew training is another critical element for ensuring safe operations. The application of AI results in adaptive training solutions, where machine learning (ML) can improve the effectiveness of training activities by using the enormous quantity of data produced during training and operations (EASA, 2023).(EASA Artificial Intelligence Roadmap 2.0 - A Human-centric approach to AI in Aviation | EASA, 2023)

DNÇER, 2023 also mentions, it is critical to establish best practices and guidelines for trainers, educators, and policymakers to enable the effective and efficient integration of technology in aviation teaching. Furthermore, addressing ethical and privacy considerations connected to data collecting, storage, and analysis is critical to ensuring data protection and responsible technology use in aviation instruction. By addressing these essential concerns, the aviation industry can continue to leverage technology's promise to improve the quality and efficacy of training programs.

Novel techniques for algorithm verification and validation, as well as usability testing with typical direct users and indirect stakeholders, are required for AI and machine learning algorithms included in Human-centered AI (HCAI) systems. The goal is to increase the likelihood that the HCAI system will perform as expected by users while decreasing the risk of unexpectedly damaging effects. Civil aviation provides strong models for new design approval, rigorous verification and validation testing during early and ongoing use, and pilot certification testing (Shneiderman, 2020).

Maloney et al., 2019 on the other hand mentions Virtual avatars can be incredibly powerful; for example, embodying a first-person virtual avatar in immersive VR can significantly influence users' perceptions, thinking, and behavior. Observing an avatar visibly lose weight while exercising, for example, boosts the amount of time and effort spent exercising, ac-

cording to research. These findings demonstrate how virtual avatars can influence users' perceptions and behaviors. However, virtual avatars may not always have a good impact on users; for example, inhabiting a black avatar in particular contexts has been proven to enhance unconscious racial bias. There has been very little investigation into the possibility for harmful or unintended alterations generated by virtual avatars. Given that avatars can influence our thinking and conduct, academics must begin defining ethical criteria for the usage of virtual avatars in commercial applications.

2.2.10. Employee Attitude and Acceptance

The role of VR expands beyond simple pilot training programs. Normally, more than one student is required to work in the virtual condition asking increasingly difficult settings commonly known as cutting edge manufactured situations. It becomes necessary to create highly authentic air-fighting reenactments (Ahir et al., 2020).

VR affects the four perception layer variables of perceived interactivity, perceived usefulness, perceived ease of use, and perceived interest, influencing workers' attitudes toward using and, ultimately, their willingness to participate in Safety Education and Training (SET). On the perceptual level, the positive feelings generated by VR, such as a positive interactive experience, perceived ease of use, utility, and curiosity, enable construction workers to fully participate in SET. On the one hand, maybe as a result of the low frequency of VR application in SET, construction workers are usually apprehensive during the process, and so they rarely experience the fun of VR. On the other hand, VR is currently focusing on SET material, with less emphasis given to form and interest (Y. Zhou & Yan, 2022).

Flight simulators provide a safe, regulated environment for pilots to hone their abilities while limiting the risks associated with real-world flight circumstances. Pilots can get expertise in managing various circumstances, such as crises or bad weather conditions, by simulating varied flight conditions, which adds to increased safety and performance. Serious games and game-based learning are potential methods for increasing learner engagement, motivation, and retention in aviation teaching. Future research should look into the design and development of game-based learning apps that are specifically designed for aviation teaching, as well as their effectiveness in fostering skill advancement and knowledge retention (DNÇER, 2023).

The participants were divided into two groups of 30 students each from 75 third-year Airline Business students in the experimental and control groups. The experimental group attended a virtual aviation safety session in which they interacted with the virtual replicated interior environment of an Airbus A320 series aircraft utilizing virtual reality goggles and controllers. The control group attended a traditional lecture-based class in a physical classroom setting. The data also revealed that immersive technology assisted in achieving a sense of reality that pushed students to participate in previously dulled aircraft safety procedures owing to loss of participation during traditional lecture class. Overall, the results showed that Airline Business students construct greater knowledge by using virtual technology that maintains their rather than passively absorbing information through lectures in a real classroom environment (Laovoravit et al., 2023).

Over 12,000 flight attendant behaviors were analyzed after training in a follow-up study utilizing an experimental design that compared real-life simulations with VR technology simulations. From the first pilot study, some changes were made to the technique. Flight at-

tendants in the first group received training that included both virtual reality and real-world hardware simulations. The purpose was to assess how successfully students transferred their information when these two training methods were used. Following their training, both groups took part in training drills. The group that received training in both real-life and VR technology simulations outperformed the group that received only real-life simulation training by 5.73%. Companies must investigate the order in which they incorporate AR/VR technology in training to determine how it affects knowledge transfer (Rosellini, 2023).

2.2.11. Realism in VR Training

There are numerous trainers and simulators in the aviation business, including door trainers, evacuation trainers, cabin service trainers, and cockpit trainers, to name a few. Some of these simulators are extremely realistic and provide tactile input. However, it is critical for flight crews to have real-world experience. Aside from flight attendants, many additional experts are involved in assuring the safety and comfort of passengers, including customer service representatives, flight dispatchers, ground crew technicians, and maintenance workers. They are in charge of doing safety checks on equipment such as life vests, flashlights, and fire extinguishers. They also monitor passenger conduct, demonstrate safety, and prepare for eventualities such as hijackings, water landings, and medical issues. Trainees can immerse themselves in numerous cabin settings in VR training for flight attendants to prepare for emotional obstacles such as dealing with a passenger's medical emergency or a security threat (Seedhouse et al., 2020).

Although there is a focus on high visual fidelity with close copies of the environment, it was discovered that interaction with digital items is not realistic, and is not identical to the real physical objects and instruments. This is due to the experience's design and the use of generic VR controllers, which do not represent the dimensions, weight, and other physical qualities of the objects and instruments being educated to use. A "break in immersion" can be defined as any factor that causes the user to lose their perceived experience of immersion in the virtual environment.Communication between an immersed VR student and another learner located outside of the Immersive Virtual Reality (IVR) environment (who observes the IVR scene via an iPad). One could argue that this interaction could disrupt immersion, affecting training performance. The usage of commercial HMDs, the requirement for multi-user interactions, immersive features, realism in object interaction, and immersion (U. Radhakrishnan & Koumaditis, 2021).

3 Strategic Planning Framework

3.1. PESTLE

To identify the political, economic, social, technological, legal, and environmental variables that influence the adoption and effective implementation of a concept or technology, a PESTLE study followed by a SWOT analysis is utilized (Christodoulou & Cullinane, 2019). PESTLE Analysis helps to analyze. This analysis helps in comprehending the external landscape and context in which KLM operates. It can reveal industrial and larger environmental trends, developments, and challenges. In the instance of JIP, we conducted a PESTEL analysis in order to apply Adaptive AI in virtual systems for the Learning and Development team to train sales reps, employees, and, eventually, cabin crew. The analysis is as follows :

1. **Political Factors:**
   (a) Regulatory Compliance Assurance: Ensure strict compliance with aviation and technology regulations within KLM.
   (b) Government Partnerships: Collaborate with government agencies to develop standardized training modules.
   (c) Impact of Bilateral Agreements: Assess the impact of international agreements on cross-border training.
   (d) Influence of Organizational Policies: KLM's internal policies and decision-making processes play a significant role in technology adoption.
2. **Economic Factors:**
   (a) Budget Allocation for Technology Updates: Allocate funds for VR and AI technology updates.
   (b) Cost-Benefit Analysis: Determine the cost-effectiveness of VR and AI training.
   (c) Exchange Rate Fluctuations: Monitor currency fluctuations when procuring technology.
   (d) Economic Trends: Consider broader economic trends that can affect the need for well-trained staff.
3. **Social Factors:**
   (a) Employee Engagement Metrics: Measure employee satisfaction with new training methods.
   (b) Diversity and Inclusion Initiatives: Ensure inclusivity in training programs.
   (c) Changing Customer Expectations: Adapt training content to evolving customer needs.
   (d) Employee Acceptance Rate: Monitor employee adaptation to new training technologies.
4. **Technological Factors:**
   (a) Compatibility Testing: Ensure seamless integration of VR and AI within KLM's existing technology.
   (b) Data Management Protocols: Establish secure data protocols.

(c) Technological Expertise: Assess KLM's internal capabilities in AI and VR development.
(d) Scalability Evaluation: Determine how well AI and VR solutions integrate with existing infrastructure.

5. **Environmental Factors:**

(a) Carbon Emission Reduction Targets: Set specific targets for reducing carbon footprint through VR adoption.
(b) Sustainability Reporting: Transparently report sustainability initiatives to stakeholders.
(c) Compliance with Environmental Regulations: Ensure adherence to environmental regulations.
(d) Impact on Carbon Emissions: Evaluate actual emission reductions achieved through AI and VR training.

6. **Legal Factors:**

(a) Data Privacy Protocols: Develop and implement robust data privacy protocols for employee data.
(b) Intellectual Property Agreements: Secure content and software used in training.
(c) Aviation Safety Regulations: Stay up to date with safety regulations affecting training.
(d) Global Data Protection Standards: Comply with international data protection when handling employee data in cross-border training programs.

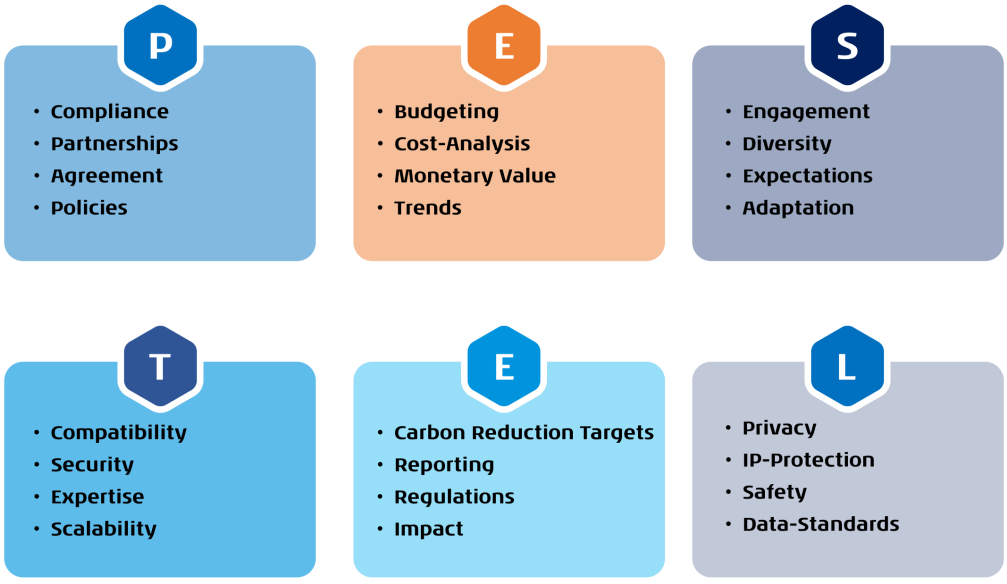


Figure 3.1: PESTEL

Considering the elements we've explored, KLM's VR and AI training program benefits from solid internal support, notably from the company itself. This backing provides a strong foundation for the initiative. Budget allocation and cost-benefit analysis are crucial for project sustainability. Social factors such as employee involvement, diversity, and customer adoption are critical to the program's success. Compatibility testing and data standards are critical on the technology front, and KLM's internal technological skills can be game changers. Environmentally, the key considerations are lowering carbon emissions and harmonizing with sustainability goals. Legal considerations include data privacy, intellectual property agreements, aircraft safety, and worldwide data protection regulations.

These insights will form the basis of a SWOT analysis, which will assist KLM in identifying internal strengths and weaknesses while also taking into account the external opportunities and dangers identified in the PESTEL research. This integrated information will help KLM make strategic decisions and develop action plans for its training activities.

3.2. SWOT

The PESTEL analysis reveals crucial external and internal factors influencing the VR and AI training project. It underscores the need to navigate a complex landscape of political, economic, social, technological, environmental, and legal considerations. This includes ensuring political stability and regulatory compliance, adapting to economic trends, focusing on employee engagement and customer expectations, addressing technological compatibility and scalability, reducing carbon emissions, and complying with data privacy and intellectual property requirements. These insights help in developing the SWOT framework as follows.

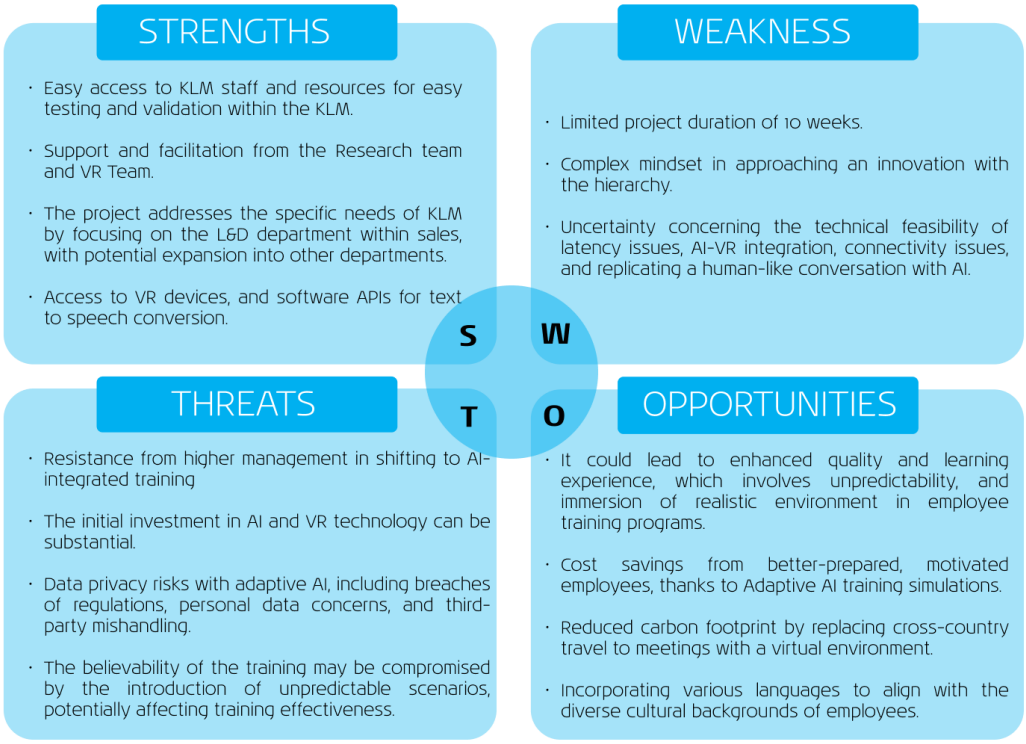


Figure 3.2: SWOT Analysis

These insights serve as the foundation for generating a SWOT analysis. The SWOT analysis for KLM's VR and AI training program highlights key internal strengths, such as easy access to company resources and strong support from internal teams. However, it also uncovers weaknesses, including the challenge of promoting innovation within the organizational hierarchy and uncertainties regarding technical feasibility. It identifies opportunities for enhanced quality and cost savings, driven by better-prepared employees and a reduced carbon footprint. On the flip side, the SWOT analysis also outlines potential threats, such as resistance from higher management and data privacy risks. Collectively, these insights empower KLM to make informed strategic decisions and develop action plans for its training initiatives.

4 Generation and Selection

4.1. Collaboration and Communication

In our ambitious journey to integrate adaptive AI to improve our training programs at KLM, collaboration and communication have proven to be two pillars of our success. It was clear from the beginning that the complexity of the project required expertise from a variety of disciplines. This insight led to the creation of a cross-functional team of skilled programmers, data analysts, and user experience researchers. Each member brought unique skills and knowledge, fostering an environment full of innovation and interdisciplinary synergy.

During the development of KLM's innovative AI training programs, the synergy between diverse departments and the emphasis on robust communication channels became paramount. The project, focusing on the integration of AI in phone conversations and the subsequent advanced training, demanded the collaboration of several KLM departments and external entities. At the center of this collaborative spirit was the JIP team, which interfaced primarily with KLM's IT department, especially its XR Lab. This unit delves into the potentialities of extended reality within KLM's operational sphere. Concurrently, the team liaised with KLM, Learning & development team (L & D) responsible for training of sales employees. This department was crucial in bridging our project's connection with the problems & challenges faced by the KLM's sales representatives in their training phase like the accessibility, timing & costs involved in travelling to the KLM's training centres from one place to another and the cultural differences that includes language barrier led to affecting the training process of the sales representatives who are primarily engaged to KLM's B2B clients.

Given the project's nature, our team undertook experiments , ensuring the training phases for employees are at ease and comfort during the phone conversations by integrating adaptive AI chatbot responses that helps in sale of tickets to KLM's B2B clients . They were thoroughly briefed about the experiment's objectives, their rights, and the safety protocols, as evidenced in the consent forms detailed in Appendix (A). Although the JIP team operated somewhat independently, channels were maintained with KLM executives through company coaches, ensuring strategic alignment. Additionally, academic touchpoints with TU Delft were established, ensuring that the project met all academic criteria. Specialized AI experts from TU Delft were also consulted to fine-tune the study.

Reflecting on the journey, two aspects stood out that is ,the synergy achieved by merging diverse expertise and the overarching emphasis on transparent, empathetic, and effective communication. These elements not only fostered a conducive environment for technological advancements but also emphasized the human-centric approach to the project.

4.2. Concept Selection

In the rapidly evolving aviation industry, KLM identified a pivotal need to revamp its traditional training modules, especially for sales representatives. With customer expectations shifting and the demand for more personalized services rising, KLM felt the urge to move away from conventional training tactics to a more immersive and adaptive model. The core objective was to ensure that their sales representatives not only meet but exceed these evolving customer expectations (Fernandes and Oliveira, 2021).

KLM has multiple departments that operate independently. Throughout the project, the JIP team collaborated closely with the Information Technologies (IT) department. The IT department has a development section with an XR Lab, which is used to investigate the possibilities of extended reality within the company. Among the myriad of technologies available, Adaptive AI stood out to be transformative potential. The choice wasn't just about embracing technology, it was about selecting tools that could provide tangible improvements in the training process. Adaptive AI, with its capability to offer real-time feedback, promised to enhance the responsiveness of sales representatives during phone interactions. On the other hand, VR, known for creating immersive environments, has the potential to bridge the gap between theoretical knowledge and practical application, but this integration with VR technology is not our focus of study during the JIP project, but the results obtained from designing a product that enhances the training of sales representatives through phone interactions will be used as a base model to further build upon technologies by integrating VR and training of KLM sales representatives.

In today's digital age, phone interactions form the bedrock of customer relationships (Osbourne and Arora, 2022). Recognizing this, KLM envisioned a system where adaptive AI would be seamlessly integrated into these conversations. Such an integration would allow the AI to analyze conversational dynamics on the go, offering representatives instant feedback and suggestions. This real-time feedback loop ensures that representatives are constantly refining their approach, making every conversation an opportunity for learning and improvement.

To bridge the gap that exists in training due to lack of time, cost of travel, adaptive AI was identified as the perfect tool, to effectively train the sales representatives and the team focussed on building the product that involves adaptive AI to effectively train sales representatives via phone conversations. The idea of "Digital training" was incorporated to resonate the training experience of the sales team at KLM, allowing representatives to engage with products and services in a digital realm, thus deepening their understanding and passion and provides elevated customer experience by ensuring that sales representatives are trained using the latest technologies, KLM aims to offer unparalleled customer service, seeks to differentiate itself in the crowded market by staying ahead in terms of training techniques and offering digital training methods, like adaptive AI , for easy scalability and consistent training quality to all sections of the people with varying cultural backgrounds.

4.3. Rationale for the Chosen Concept

In the vast spectrum of technological innovations, KLM's decision to adopt the "Adaptive AI-Enhanced Existing Training Programs" as a core training mechanism stands out. The choice was derived after meticulously weighing the feasibility, effectiveness, and alignment with KLM's broader mission and goals("Analytical study on use of AI techniques in tourism sector for smarter customer experience management", 2020).

Realism & Immersive Training: The immersive nature of VR, when combined with believability of Adaptive AI, offers an unprecedented degree of realism in training sessions. KLM's sales representatives can interact with AI-driven personas in environments that mirror real-world scenarios. Such realistic and engaging training prepares them better, gives them more confidence and sharpens their skills.

Flexibility & customization: Different roles within KLM pose unique challenges. The beauty of the chosen idea lies in its adaptability and feasability. The flexible and unpredictable nature of adaptive AI, ensures that training sessions remain relevant, meeting the changing needs and unique requirements of each role of sales trainee.

Holistic Integration & Interdisciplinary Approach: KLM's decision also signifies its capability to bridge gaps between various disciplines like the domains of AI Development, Learning & Development, and Sales, KLM promises a training experience that's both comprehensive and novel. This holistic approach underscores KLM's futuristic vision and its dedication to leading the technological curve.

Robust Technical Framework: KLM's journey into this new training paradigm isn't just a whimsical decision; it's rooted in sound scientific principles. Leveraging mature AI technologies and recent advancements in integrating adaptive AI with assistant bots, KLM ensures that its training mechanism is both innovative and technically robust.

Sustainability and profitability: KLM's importance to sustainable development is demonstrated through the choice of this training model. By reducing the need for physical training infrastructures and the associated travels, the company not only achieves significant cost savings but also takes a substantial step in decreasing its carbon footprint. The eco-friendly undertones of this decision resonate with KLM's sustainability vision.

Alignment with Broader Goals: Beyond the immediate benefits of enhanced training, the incorporation of Adaptive AI aligns with KLM's broader ethical and sustainability objectives. The concept particularly resonates with the **People, Planet, and Profit (PPP)** ethos. By minimizing its physical and carbon footprint, KLM solidifies its commitment to environmental and social responsibilities.

Hence, KLM's choice of the Adaptive AI-Enhanced Existing Training Programs isn't just a technological leap, it embodies the company's vision for the future. A harmonious blend of advanced technology, commitment to sustainability, and a pursuit of excellence defines this strategic direction, setting KLM on a trajectory that promises unparalleled training experiences and customer interactions.

4.4. Safety and Ethical considerations

In today's rapidly evolving digital landscape, companies are continuously seeking innovative approaches to train their workforce. At the heart of KLM Royal Dutch Airlines' approach is the integration of Adaptive AI in improving the training of sales representatives, a fusion that promises unparalleled training experiences.This venture isn't just about technological advancements, it's a concerted effort to ensure that the training is ethically administered, data privacy is upheld, and the content remains inclusive and culturally sensitive. Our exploration delves into how KLM, through inter-departmental collaboration and continuous feedback loops, achieves a harmonious balance between technological advancement and ethical considerations in training sales representatives (Pflanzer et al., 2022).

Adaptive AI & Effective Training: One of the unique facets of integrating AI technology is the adaptability it brings to training scenarios. At KLM, our focus is on employing AI to enhance phone conversation training for sales representatives. This adaptive AI framework allows these training modules to evolve based on user interactions and feedback, making the learning curve more tailored and effective. The KLM Research and L&D Departments work in tandem to continuously refine this adaptive approach, ensuring that the scenarios presented remain relevant, challenging, yet achievable. This is especially crucial for sales representatives, where effective communication is paramount.

Robust Data Security Measures: As KLM delves deeper into AI-enhanced training, the collection of individualized data becomes inevitable. This brings forth concerns regarding data privacy and protection. KLM's IT Department, with its specialized Development and AI units, is at the forefront of ensuring rigorous data protection. By employing cutting-edge encryption techniques and data anonymization protocols, they guarantee the utmost confidentiality of personal data. KLM's unwavering dedication to ethical data practices is further fortified by the vigilant oversight of its senior management, ensuring that these high standards are always maintained.

Stakeholder Collaboration for Ethical Deployment: The use of AI isn't just about leveraging advanced technology, it's about deploying it ethically and responsibly. The collaborative framework between various departments ensures that participants' well-being is always a priority. The feedback from KLM Sales employees, who are at the frontline of using these training modules, is pivotal. Their insights help us gauge the system's psychological impact and make necessary adjustments.

Commitment to Inclusive Training: KLM places a high premium on inclusivity in its training programs. The close collaboration between its research and development units guarantees that the training modules are tailored to cater to a diverse range of learning requirements. The IT department's AI specialists ensure that the technical elements of the training are universally accessible. Beyond just technical accessibility, there's a deep-rooted commitment to cultural inclusivity. Given the rich cultural tapestry of KLM's sales staff, their input is invaluable in creating training content that's not only culturally sensitive but also devoid of biases and stereotypes and ensure that the training content respects cultural nuances and promotes an inclusive learning environment for everyone.

A notable concern in this experiment arises from the collaboration between the research team and KLM. Given this partnership, there's a potential perception among employees

that the company is closely monitoring their responses. This perception might inadvertently skew their answers on the questionnaires. To mitigate this risk, we've taken rigorous measures like any data collected, including questionnaires and video recordings, are strictly kept away from KLM's management and coordinators. Instead, this data is securely stored on JIP servers and is slated for deletion post the completion of the final report, guaranteeing participants' anonymity. While the JIP team undertakes a comprehensive analysis of the data, only the conclusive report will be shared with KLM's higher-ups, ensuring participant confidentiality ("Privacy Issues in Voice Assistant Ecosystems", 2020).

Moreover, we have implemented stringent data collection guidelines. Only information pertinent to the study's objectives is procured, ensuring that no direct identifiable personal data gets archived. However, there's the collection of certain indirectly identifiable data, for which we assign each participant a unique identification code to preserve their anonymity.

Prior to the experiment, we obtained informed consent from all participants. This process was carried out in alignment with the guidelines of TU Delft's ethical committee, and all associated documentation can be found in Appendix (A) for reference.

The journey of KLM Royal Dutch Airlines in integrating AI for training purposes is a testament to how technology can be harmoniously blended with ethical considerations to create effective and responsible learning environments. Through complex interactions between different departments and valuable feedback from executive staff, KLM sets a model for how companies can and should approach modern training. As we move into the digital age, KLM's model serves as an example, demonstrating that technological prowess, when combined with ethical responsibility and inclusivity , can lead to transformative results and act according to the GDPR guidelines for data and privacy issues.

4.5. Sustainability Impacts

In KLM's pursuit of leveraging adaptive AI , there's a distinct alignment with several of the UN's sustainable development goals, notably:

- **Goal 7:** Ensure **access to affordable and clean energy for all**. Promoting sustainable energy solutions, our initiatives use cutting-edge AI technology to improve energy efficiency and environmental management in KLM training operations and improving productivity of sales representatives.
- **Goal 8:** Promote sustained, inclusive and sustainable **economic growth**, full and productive employment and **decent work for all**. Ensuring sustained economic growth and offering productive, fulfilling employment experiences. The project's potential to reintegrate employees across industries embodies this goal.
- **Goal 9:** Build resilient infrastructure, promote inclusive and sustainable industrialisation and **foster innovation**. By embedding innovative adaptive AI technology in effective training of sales representatives & workforce reintegration, KLM exemplifies building resilient infrastructure and fostering innovation.

Central to this innovative endeavor is KLM's vision of harnessing AI to refine phone-based conversations, training sales representatives more effectively. In subsequent phases, the ambition is to heighten conversational efficacy through immersive VR experiences.

KLM's strategic deployment of AI not only promises enhanced training experiences but also champions sustainability. By minimizing the need for physical resources and travel, there's a significant reduction in carbon footprints. This proactive shift, encouraged by KLM's top leadership, accentuates KLM's dedication to environmental stewardship, an ethos heartily endorsed by KLM's customer base.

Yet, the sustainability conversation transcends just environmental conservation. Efficiency is key, and the IT team at KLM ensures that despite the complexity of AI modules, energy consumption remains minimal. Regular oversight from the KLM Research Department ensures that this technological journey remains efficient in terms of resource utilization.

KLM's adaptive AI training initiative emphasizes sustainability, notably through reducing carbon footprints by lessening the demand for physical resources and travel. This sustainable ethos, fervently endorsed by KLM's top management, not only highlights the company's environmental dedication but also aligns with the preferences of eco-conscious customers. Beyond environmental conservation, the KLM IT team ensures the AI modules are energy-efficient, while the Research Department oversees resource optimization. The combined efforts of the L&D and IT Departments guarantee that the AI training is both innovative and economically sustainable, with long-term cost-saving benefits.A pivotal advantage of this project is its long-term economic viability. With the synergy between the Learning & Development and IT Departments, these AI modules are designed to be adaptable and scalable. Such foresight, backed by KLM's visionary leadership, positions KLM's training initiatives at the intersection of innovation and fiscal prudence, ensuring sustainability in the broader sense.

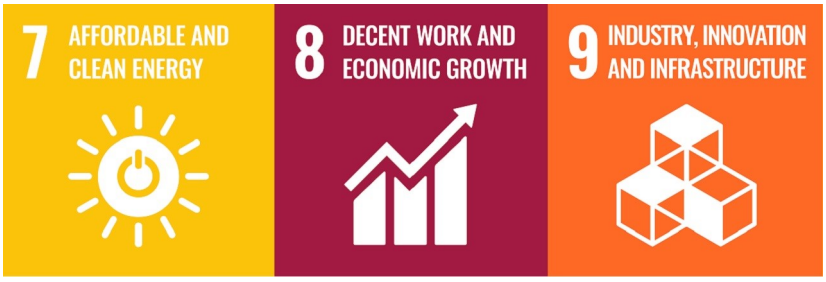


Figure 4.1: The sustainable development goals relevant for this project

4.6. Risk Assessment

KLM Royal Dutch Airlines is venturing into an ambitious integration of adaptive AI technologies, aiming to revolutionize training methodologies, especially for sales representatives. However, implementing such technology doesn't come without challenges (Klein et al., 2021).

The Research Department at KLM, which bridges academic theory with tangible application, faces the possible pitfall of creating training modules that might not fully capture real-world dynamics. To counteract this, regular engagements with the IT and L&D departments are paramount. This collaboration ensures that the devised training methodologies are not only technically feasible but are also aligned with the current global trends in adaptive AI . Synchronized sessions between these teams are essential to maintain alignment, thereby fostering a synergistic environment. Given the nature of technology, occasional problems are inevitable. Therefore, a rigorous testing phase becomes necessary to identify and resolve issues upfront, thereby ensuring that the integration of AI is made seamless. Additionally, as technology advances rapidly, it is important to stay up to date with global trends in AI to keep the curriculum relevant.

Meanwhile, L&D departments, which typically oversee training programs, face the challenge of seamlessly combining conventional training techniques with this advanced AI system. To avoid a fragmented learning experience, consistent dialogue with research and IT teams is critical. This will allow for a balanced mix of conventional and advanced training tools. Diving into the technical realm, IT department at KLM faces its own challenges when entering the technical field. Split between VR, learning and development, and AI teams can lead to inconsistencies in goals and hinder progress. Regular sync sessions between these teams can foster a collaborative environment. We cannot ignore the fact that training can be interrupted due to technical glitches. Therefore, a robust testing phase is essential to proactively identify and resolve technical issues to ensure seamless integration of AI into training modules.

KLM training staff are at the heart of this approach, their participation and comments are essential to the success of the system. Resistance to new technology is a natural human tendency. Therefore, introducing them to the system through practical workshops can reduce their apprehension. Additionally, by promoting an environment in which they can express their opinions, the system remains employee-centric.

For KLM's senior management, who are overseeing the project's trajectory, the challenge is to ensure that the significant investment in this new AI initiative generates tangible returns. Consistent updates, clear communication of expected Return on Investments(ROI) and ensuring alignment with KLM's broader objectives will provide clarity and direction.

Ultimately, the real challenge lies with KLM's customer base. The ultimate goal of all training is to improve service delivery. Without a clear improvement in service, it risks losing favor with this important group. Establishing robust feedback mechanisms, both from frontline staff and through direct customer surveys, will be essential to ensure that service standards not only meet but exceed expectations.

4.7. Responsible Innovation:

KLM Royal Dutch Airlines has firmly anchored the principles of responsible innovation in their strategic approach. Embracing technology trends, especially in the realm of adaptive AI and virtual reality, they're charting a course towards a future that's responsive, inclusive, and forward-thinking. By incorporating AI into phone conversations, KLM is revolutionizing how sales representatives are trained. This initiative isn't merely about improving efficiency but aims to create a genuine, personalized experience for every customer. As they transition into the later stages of training, virtual reality further augments this process, providing representatives with a near-real world environment to hone their skills.

For KLM Royal Dutch Airlines, responsible innovation isn't just a catchphrase, it's an ethos. The Research department demonstrates this by continuously refining their methods to reflect current technology trends and training needs. IT departments focus on collaborative innovation and ensure that the technology developed is both cutting-edge and relevant. The L&D department's focus on balancing technology with personal interactions underscores KLM's commitment to retaining the human touch in their services. KLM's approach to integrating these technologies has been meticulous. Cognizant of the disruptive nature of VR, as pointed out by (Stahl, 2022), the airline has been careful in its application, particularly concerning the reintegration process of employees. This consideration ensures that both the privacy and health of participants are safeguarded.

Furthermore, the airline's commitment to inclusivity is evident in their demographic approach. They understand that a truly responsible innovation is one that caters to all. This philosophy extends to the sales team, cabin crew, where they're actively researching the myriad reasons for absences to provide comprehensive solutions. Encouraging the Operational staff to share insights ensures innovation is based on real-world needs. The role of senior management is to constantly ensure that every step of innovation is consistent with KLM's ethics and sustainability vision. By prioritizing customer feedback, KLM Royal Dutch Airlines ensures that innovation is always customer-centric and drives both satisfaction and loyalty.

By targeting the United Nations' Sustainable Development Goals, KLM's project underscores its intent to enhance the quality of life for future generations of flight attendants. They've left no stone unturned, even analyzing the content of consent forms to guarantee participant safety and project reliability. Assessments have been conducted on various fronts from the feasibility of the adaptive AI technology to its desirability among sales representatives (Owen et al., 2013) remind us that responsible innovation is best judged by the future implications of its results and conclusions. KLM's approach to integrating AI and is a testament to this, balancing business viability with ethical considerations.

In essence, KLM's focus on responsible innovation emphasizes its commitment to both technological advancement and human-centric values. The airline's research, IT, and L&D departments work in tandem to create solutions that are not only state-of-the-art but also deeply rooted in real-world needs. By valuing insights from operational staff and continuously seeking customer feedback, KLM ensures its innovations are both impactful and aligned with its ethical and sustainability vision. This integrated approach reinforces KLM's reputation as a leader in customer satisfaction, loyalty, and responsible innovation.

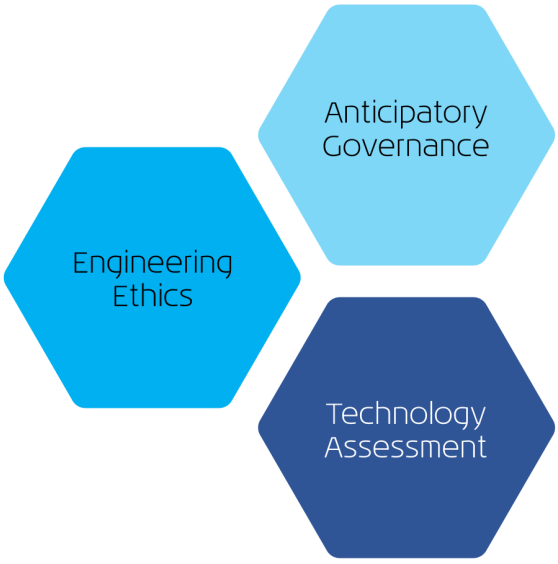


Figure 4.2: Responsible Research and Innovation

4.8. Financial Perspective

In the dynamic world of aviation, effective training is the cornerstone of operational excellence and customer satisfaction. KLM, a recognized leader in the industry, invests heavily in its training initiatives, underscoring its commitment to unparalleled service standards. Yet, with substantial expenditures, especially in hiring actors for training simulations, there's an evident gap in optimizing return on investment. Our groundbreaking product, poised to transform KLM's training landscape, designed to significantly boost retention rates and offer sizable cost savings. Our solution is more than just an innovative tool, it's a strategic asset that promises to redefine training efficacy and financial prudence for KLM.

KLM's Current Training Investments- A Deep Dive: KLM's commitment to training is evident in its sizable annual investment of 58 million euros. This substantial outlay reflects the airline's dedication to maintaining high standards of service and operations. A significant fraction of this, 1.3 million euros, is earmarked for hiring actors for role-playing simulations, underscoring the interactive and realistic nature of their training programs. However, with a retention rate standing at just 25%, there's a clear indication of inefficiencies. This figure suggests that for every four euros spent on training, only one euro translates to lasting knowledge or skills.

About Our Product: A Revolution in Training Efficiency: Our innovative product, designed to pivot KLM's training approach. It promises to not just incrementally improve, but to substantially revolutionize the efficacy of these programs. A leap from a 25% retention rate to a staggering 75% is not just an impressive statistic on paper but translates to three out of every four euros spent on training being effectively utilized.

Eliminating the need for actors in training scenarios is a game-changer. An annual savings of 1 million euros from this factor alone is a compelling argument. But the financial benefits don't stop there. The ripple effect of using our product means fewer logistical costs. Fewer actors mean less travel, accommodation, and subsistence expenses, leading to indirect savings.

The Broader Implications for KLM: The advantages of our product are multifaceted. On the one hand, it ensures employees are equipped with superior skills and knowledge, enhancing their performance, and elevating the customer experience. This, in turn, can lead to increased customer loyalty, potentially boosting revenues and brand reputation.

On the financial side, cost savings are just the tip of the iceberg. A more effective training program could reduce errors and inefficiencies in operations, leading to further savings. Moreover, a well-trained workforce can streamline operations, leading to faster turnaround times and increased flight punctuality.

Conclusion : Our product offers KLM an opportunity to be at the forefront of training innovation. By embracing this solution, KLM can set industry benchmarks, ensuring its workforce is among the best trained while optimizing costs. This isn't just about immediate savings, it's about long-term growth, brand reputation, and setting new standards in aviation training. With our product, KLM has the potential to lead, inspire, and ascend to new heights in the industry.

5 Product Design and Implementation

5.1. Functional architecture

In this chapter, we employ a systems engineering approach to elucidate the functional architecture of the adaptive AI phone application software. The software's design structure is based on Chapter 9 of the book "Practical Model-Based Systems Engineering." We initiate this exploration by narrating a brief scenario that illustrates the intended usage of the application. This usage scenario is subsequently divided into two main phases: the deployment phase, focusing on setting up the phone application for training, and the operational phase, which revolves around conducting the training.

From these usage scenarios, we extract the needs and functional requirements of the phone application. These functional requirements serve as the foundation for generating functional specifications, which, in turn, are utilized to construct a functional hierarchy tree. This hierarchy tree provides an organized breakdown of the phone application into parent and child functions.

Ultimately, an activity diagram offers an overview of how the application is employed during the setup phase of a training session.

5.1.1. Usage Scenarios

The purpose of the usage scenario is to steer the software design process, providing developers with valuable insights into the application's needs and functional requirements. This approach allows us to progress from abstract concepts to specific details, aiming to reduce the need for rework.

Simple usage scenario

This scenario provides an illustrative glimpse into the envisioned utilization of the phone application by the sales department.

"At 8:45 am, I strolled into the office, gearing up for what promised to be a pivotal sales call with one of our esteemed Indian clients. Rajeev Malhotra, the CEO of Malhotra Tech Innovations, was on the other end of the line, and I needed to make this conversation count.

I reached for the AI-Flight-Innovators app on my smartphone, the secret weapon that had transformed the way we prepare for our sales calls at KLM. As I opened the app, it prompted me to input essential information about the impending call with Rajeev. I filled in the details, making sure to include key notes gathered from our database. It painted a vivid picture of Rajeev: a man renowned for his unwavering commitment to business ethics, cultural values, and sustainability.

My goal for this call was clear:

- Negotiate a ticket price not lower than 500 euros.
- Ensure to convince Rajeev that our partnership resonates with his cultural values and norms.
- Define and agree upon sustainability targets.

With the server fired up, I knew I was ready to dive in. I initiated the call with Rajeev, and from the very beginning, it was evident that he wanted to discuss those very objectives. We delved into KLM's sustainability targets, ensuring that every detail was to his satisfaction. I emphasized our strong commitment to cultural values and norms, and, of course, skillfully negotiated a ticket price that would benefit both parties.

What made this experience truly remarkable was the unpredictability of it all. Rajeev posed questions and provided responses that I hadn't anticipated. It was like a high-stakes chess game, and I was playing to win. The AI-Flight-Innovators app had superbly prepared me for this moment, and I felt confident and in control as I navigated the conversation with Rajeev.

In the end, I was more than ready for my real call with Rajeev later in the day. The AI-powered guidance had elevated my abilities and set me on the path to success, making every sales call a thrilling journey."

This brief narrative underscores the enhanced confidence and motivation derived from employing the phone application, which effectively readies the employee for the unpredictability of real-life phone calls.

Usage Scenarios

The usage scenarios can be categorized into two distinct phases: the Deployment Phase and the Operational Phase. In the Deployment Phase, the focus is on configuring the application to equip the sales employee for a sales negotiation with a particular customer. This process is repeated each time an employee needs to prepare for a different customer and encompasses tasks such as setting up the server and inputting the relevant customer data into the application.

On the other hand, the Operational Phase involves the actual phone call itself. This phase can be reiterated for a specific customer until the sales employee is sufficiently confident to initiate the real-life call with the customer.

In the progression from the abstract to the specific, the next phase involves delineating the usage scenarios for both the deployment and operational phases as represented in table 5.1.

Deployment phase	
S1 Load customer database	The sales employee collects customer data and provides it to the application for setup and configuration.
S2 Application Initialization	The server is launched and made accessible on the internet, primed for incoming calls.
Operational phase	
S3 Call commencement	The application's phone number is dialed, and the call is redirected to the application's webhook.
S4 Welcome Message Playback	As the call commences, the application initiates the conversation by playing a welcome message.
S5 Audio Input Reception	The application captures the caller's audio input.
S6 Speech-to-Text Conversion	The application transcribes the user's voice input.
S7 Response Generation	The transcription is sent to the adaptive AI for generating a response.
S8 Text-to-Speech Conversion	The generated response is transformed from text into audible speech.
S9 Call Termination	The caller ends the call.

Table 5.1: Usage Scenarios

5.1.2. Extended operational scenarios

Continuing the transition from the abstract to the specific, this section extends the foundation laid in Section 5.1.1. Instead of merely outlining general operational contexts for the application, the focus shifts to detailed decomposition, resulting in well-defined, specific scenarios that encompass pre- and post-conditions as well as the events that trigger them. Expanding on these operational scenarios becomes crucial to ensure that the system or solution meets all necessary criteria for effective operation during its use.

Before any call can take place, the application must be prepared and configured. This includes the collection of customer data by the sales employee and providing this information to the application. The extended deployment scenario S1, as described in Table 5.2, outlines this essential setup phase.

Table 5.2: Extended Deployment Scenario S1 - Load Customer Database

Deployment Scenario	*S1 - Load Customer Database*
Preconditions	• The sales employee collects customer data and provides it to the application for setup and configuration.
Triggering Event	The process is initiated when the employee needs to prepare for a specific customer call.
Description	The sales employee loads the customer's data into the application, ensuring it's ready for the upcoming call.
Postconditions	• Customer data is successfully loaded into the application for call preparation.

The application, including its server, needs to be initialized and made accessible on the internet, ready to receive incoming calls. This is a crucial step in the deployment phase, as it ensures the application is operational. The extended deployment scenario S2, detailed in Table 5.3, defines the process for launching the application.

Table 5.3: Extended Deployment Scenario S2 - Application Initialization

Deployment Scenario	*S2 - Application Initialization*
Preconditions	• The server is configured and ready to accept incoming calls.
Triggering Event	The application is launched, and the server is exposed to the internet, prepared for usage.
Description	The process begins with the initialization of the application, making it accessible on the internet for incoming calls.
Postconditions	• The application and server are successfully initialized for operation.

When a caller dials the application's phone number, it triggers the initiation of a call. The application then efficiently redirects the call to its webhook for further processing. This process is defined in the extended operational scenario S3, as presented in Table 5.4.

Table 5.4: Extended Operational Scenario S3 - Call Commencement

Operational Scenario	*S3 - Call Commencement*
Preconditions	• The application's phone number is dialed, and the call is redirected to the application's webhook.
Triggering Event	The initiation of the call occurs when the caller dials the application's phone number.
Description	The application receives the call and redirects it to the application's webhook for handling.
Postconditions	• The call is successfully redirected to the application for further processing.

Upon the commencement of a call, the application plays a welcome message to initiate the conversation. This step is essential to create a positive and engaging caller experience. The extended operational scenario S4, detailed in Table 5.5, defines the process for playing the welcome message.

Table 5.5: Extended Operational Scenario S4 - Welcome Message Playback

Operational Scenario	*S4 - Welcome Message Playback*
Preconditions	• The call is initiated and connected.
Triggering Event	The application plays a welcome message at the beginning of the call.
Description	Upon the commencement of the call, the application initiates the conversation by delivering a welcome message to the caller.
Postconditions	• The welcome message is successfully played, and the conversation progresses.

The application is designed to capture the caller's audio input during the call. This reception is a critical step in understanding the caller's needs and queries. The extended operational scenario S5, as presented in Table 5.6, outlines the process of audio input reception.

Table 5.6: Extended Operational Scenario S5 - Audio Input Reception

Operational Scenario	*S5 - Audio Input Reception*
Preconditions	• The welcome message has been played, and the call is ongoing.
Triggering Event	The application is ready to capture and process the caller's audio input.
Description	The application records and processes the caller's audio input as they speak.
Postconditions	• The audio input is successfully captured and ready for transcription.

After receiving the caller's audio input, the application performs speech-to-text conversion to transcribe the spoken words into text format. This textual representation is then used for further analysis and response generation. The process is defined in the extended operational scenario S6, detailed in Table 5.7.

Table 5.7: Extended Operational Scenario S6 - Speech-to-Text Conversion

Operational Scenario	*S6 - Speech-to-Text Conversion*
Preconditions	• The audio input has been received.
Triggering Event	The application processes the audio input to convert it into text.
Description	The application performs speech-to-text conversion on the caller's audio input.
Postconditions	• The audio input is successfully transcribed into text.

Once the audio input is transcribed into text, the application sends this text to an adaptive AI system for generating a response. This step is crucial for providing meaningful and context-aware replies to the caller's queries. The extended operational scenario S7, as presented in Table 5.8, outlines the process for generating a response.

Table 5.8: Extended Operational Scenario S7 - Response Generation

Operational Scenario	*S7 - Response Generation*
Preconditions	• The transcribed text from the caller's input is available.
Triggering Event	The application sends the transcribed text to the adaptive AI for generating a response.
Description	The application forwards the transcribed text to the adaptive AI, which generates a response.
Postconditions	• A response is successfully generated by the adaptive AI.

After the adaptive AI generates a response in textual form, the application proceeds to convert this text into audible speech. The resulting voice response is then played back to the caller for a seamless conversation. The process is detailed in the extended operational scenario S8, as outlined in Table 5.9.

Table 5.9: Extended Operational Scenario S8 - Text-to-Speech Conversion

Operational Scenario	*S8 - Text-to-Speech Conversion*
Preconditions	• A response has been generated in text format.
Triggering Event	The application transforms the generated text response into audible speech.
Description	The application converts the generated text response into speech format.
Postconditions	• The response is successfully converted into audible speech.

When the caller decides to end the call, they initiate the process of call termination. The application recognizes this triggering event and promptly terminates the call, ensuring a smooth and efficient user experience. The extended operational scenario S9, presented in Table 5.10, defines the process for call termination.

Table 5.10: Extended Operational Scenario S9 - Call Termination

Operational Scenario	*S9 - Call Termination*
Preconditions	• The call is ongoing, and responses have been exchanged.
Triggering Event	The caller decides to end the call.
Description	The caller initiates the termination of the call.
Postconditions	• The call is successfully terminated by the caller.

5.1.3. Needs and functional requirements

The development of any sophisticated system requires a clear understanding of stakeholder needs and precise definition of functional requirements. These requirements are fundamental in shaping the design, development, and successful deployment of the system. Just as a finely tuned machine operates with precision, stakeholder needs and functional requirements are the gears and cogs that propel the system toward achieving its objectives.

In the context of the following scenarios, the functional requirements are a direct reflection of stakeholder needs. These needs are classified into two key categories: deployment needs, which pertain to preparing and launching the application, and operational needs, which relate to the application's functioning during active use.

The tables in this section provide a detailed breakdown of the deployment and operational needs and functional requirements. Each table corresponds to a specific scenario and outlines the unique needs and requirements associated with that scenario. This granular approach ensures that every aspect of the application's design, deployment, and operation is methodically addressed.

These needs and requirements serve as a guiding light, ensuring that the application's design and operation are meticulously aligned with its intended purpose. Whether it's handling incoming calls, playing welcome messages, converting speech to text, or generating

meaningful responses, these tables offer a comprehensive blueprint for the application's development and utilization, underpinning the mission to enhance customer satisfaction and streamline operational processes.

Table 5.11: Deployment needs and functional requirements for S1 - Load Customer Database

Deployment needs	*S1 - Load Customer Database*
DN_S1_1	The sales employee collects customer data and provides it to the application for setup and configuration.
DN_S1_2	The application needs to prepare for a specific customer call.
Functional requirements	*S1 - Load Customer Database*
FR_S1_1	The application must have the capability to receive and store customer data provided by the sales employee.
FR_S1_2	The application should initiate the loading process when needed for a specific customer call.

Table 5.12: Deployment needs and functional requirements for S2 - Application Initialization

Deployment needs	*S2 - Application Initialization*
DN_S2_1	The server is configured and ready to accept incoming calls.
DN_S2_2	The application needs to be launched and made accessible on the internet.
Functional requirements	*S2 - Application Initialization*
FR_S2_1	The application must have the capability to configure the server to accept incoming calls.
FR_S2_2	The application should be able to launch and make itself accessible on the internet.

Table 5.13: Operational needs and functional requirements for S3 - Call Commencement

Operational needs	*S3 - Call Commencement*
ON_S3_1	The application's phone number is dialed, and the call is redirected to the application's webhook.
ON_S3_2	The initiation of the call occurs when the caller dials the application's phone number.
Functional requirements	*S3 - Call Commencement*
FR_S3_1	The application should have the capability to receive incoming calls when its phone number is dialed and redirect them to its webhook for handling.
FR_S3_2	The application should initiate the call when the caller dials its phone number.

Table 5.14: Operational needs and functional requirements for S4 - Welcome Message Playback

Operational needs	*S4 - Welcome Message Playback*
ON_S4_1	The call is initiated and connected.
Functional requirements	*S4 - Welcome Message Playback*
FR_S4_1	The application must play a welcome message at the start of each call.

Table 5.15: Operational needs and functional requirements for S5 - Audio Input Reception

Operational needs	*S5 - Audio Input Reception*
ON_S5_1	The application must have access to the caller's audio input during the call.
ON_S5_2	The application should capture the caller's speech for further processing.
Functional requirements	*S5 - Audio Input Reception*
FR_S5_1	The application must have the capability to record the caller's audio input.
FR_S5_2	The application should be able to process the recorded audio input for analysis.

Table 5.16: Operational needs and functional requirements for S6 - Speech-to-Text Conversion

Operational needs	*S6 - Speech-to-Text Conversion*
ON_S6_1	The application must have recorded audio input to convert to text.
ON_S6_2	The application should transcribe the spoken words into text format.
Functional requirements	*S6 - Speech-to-Text Conversion*
FR_S6_1	The application must have the capability to convert recorded audio into text.
FR_S6_2	The application should accurately transcribe the spoken words into textual format.

Table 5.17: Operational needs and functional requirements for S7 - Response Generation

Operational needs	*S7 - Response Generation*
ON_S7_1	The application must have transcribed text from the caller's input.
ON_S7_2	The application should forward the transcribed text to an adaptive AI system.
Functional requirements	***S7 - Response Generation***
FR_S7_1	The application must have the capability to send transcribed text to an adaptive AI system.
FR_S7_2	The application should receive meaningful and context-aware responses from the adaptive AI system.

Table 5.18: Operational needs and functional requirements for S8 - Text-to-Speech Conversion

Operational needs	*S8 - Text-to-Speech Conversion*
ON_S8_1	The application must have generated text responses.
ON_S8_2	The application should convert the text responses into audible speech.
Functional requirements	***S8 - Text-to-Speech Conversion***
FR_S8_1	The application must have the capability to transform text responses into audible speech.
FR_S8_2	The application should play the audible speech responses to the caller.

Table 5.19: Operational needs and functional requirements for S9 - Call Termination

Operational needs	*S9 - Call Termination*
ON_S9_1	The application must have ongoing calls with responses exchanged.
ON_S9_2	The application should recognize when the caller decides to end the call.
Functional requirements	***S9 - Call Termination***
FR_S9_1	The application must have the capability to monitor ongoing calls and responses.
FR_S9_2	The application should detect when the caller decides to terminate the call and respond accordingly.

5.1.4. Functional hierarchy tree

This section offers a thoroughly organized overview of the functional components and their intricate relationships within the phone application. These components have been meticulously designed to align with the established needs and functional requirements,

as previously outlined in Section 5.1.3. The structural hierarchy is illustrated as a tree diagram, effectively categorizing these elements into two primary phases: Deployment and Operational. Each phase, in turn, is further subdivided into multiple functionalities, providing detailed insights into the precise processes and interactions that unfold during the deployment and operational phases of the application.

Within the Deployment Phase, these functionalities are meticulously structured in a hierarchical format, offering a comprehensive view of the sequence from loading the customer database to initializing the application. These actions collectively lay the groundwork for the application to effectively accommodate incoming calls. In essence, this phase prepares the application to cater to the unique requirements of each customer, ensuring that it is finely tuned for specific use cases.

The Operational Phase is characterized by a sequence of interactions that are recurrent in nature, serving to prepare for sales conversations with specific customers. Due to page size constraints, the hierarchy tree for the Operational Phase is divided into multiple figures. Part 1 of the Operational Phase revolves around activities linked to call commencement. These activities include the reception of incoming calls, their redirection to the appropriate processes, and the playback of a welcoming message.

Part 2 delves into the intricate processes of speech-to-text conversion and response generation, highlighting the progression from audio conversion to meaningful response reception.

Part 3 encapsulates the final stages of operational activities, encompassing text-to-speech conversion, call termination procedures, and the identification of the caller's intent to conclude the call.

Each functionality is accompanied by a succinct overview of its purpose, input prerequisites, and anticipated outputs. This comprehensive documentation provides an all-encompassing comprehension of the application's intricate functional structure.

Deployment Phase Functionalities

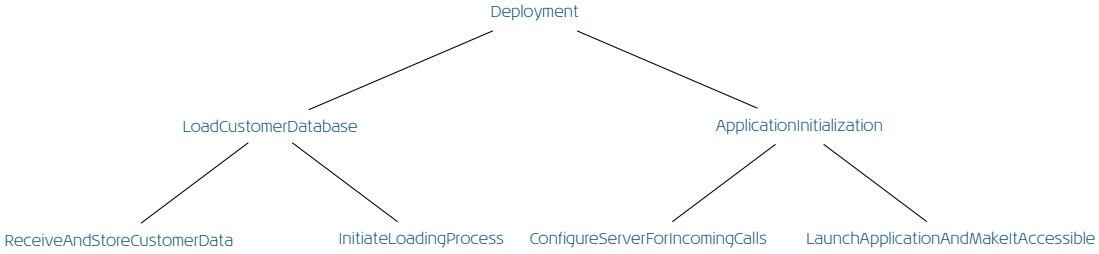


Figure 5.1: Deployment Phase Functional Hierarchy Tree (Landscape)

Table 5.20: Function Description 1.1

Function name:	Load Customer Database
Label:	1.1
Description:	Load Customer Database Functionality Description
Inputs:	The input for this functionality is a request to load the customer database, which includes customer data and specific parameters for personalization.
Outputs:	The output of this functionality is a successfully loaded customer database, ensuring that the application has access to customer-specific information.

Table 5.21: Function Description 1.1.1

Function name:	Receive and Store Customer Data
Label:	1.1.1
Description:	Receive and Store Customer Data Functionality Description
Inputs:	This sub-functionality receives incoming customer data, which may include personal information, preferences, and conversation history.
Outputs:	It outputs the stored customer data in a format accessible by the application, allowing for tailored interactions.

Table 5.22: Function Description 1.1.2

Function name:	Initiate Loading Process
Label:	1.1.2
Description:	Initiate Loading Process Functionality Description
Inputs:	The initiation of the loading process requires parameters related to the data to be loaded and triggers the loading sequence.
Outputs:	The loading process results in a fully loaded customer database ready for use by the application.

Table 5.23: Function Description 1.2

Function name:	Application Initialization
Label:	1.2
Description:	Application Initialization Functionality Description
Inputs:	Application initialization involves configuring the server for incoming calls and launching the application for customer interactions.
Outputs:	The initialized application is made accessible, prepared to handle incoming calls.

Table 5.24: Function Description 1.2.1

Function name:	Configure Server for Incoming Calls
Label:	1.2.1
Description:	Configure Server for Incoming Calls Functionality Description
Inputs:	The input specifies the configuration settings for the server, ensuring it is ready to receive incoming calls.
Outputs:	The server is configured to handle incoming calls efficiently.

Table 5.25: Function Description 1.2.2

Function name:	Launch Application and Make it Accessible
Label:	1.2.2
Description:	Launch Application and Make it Accessible Functionality Description
Inputs:	The launch request requires information about the application's setup and access parameters.
Outputs:	The application is successfully launched and made accessible for users.

Operational Phase Functionalities (Part 1)

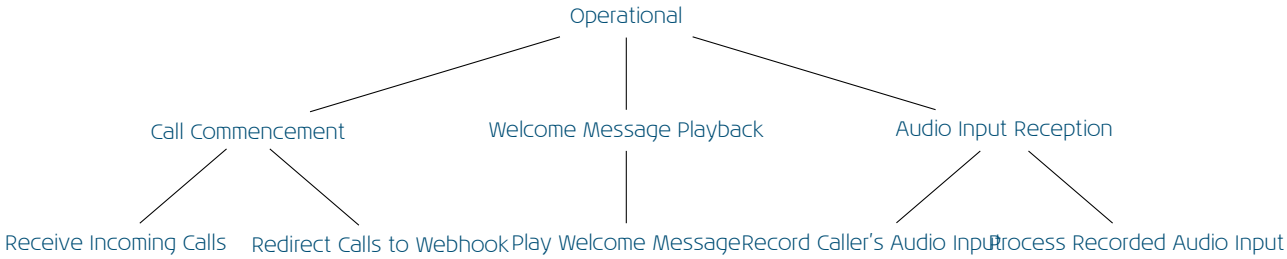


Figure 5.2: Operational Phase Functional Hierarchy Tree (Part 1)

Table 5.26: Function Description 2.1

Function name:	Call Commencement
Label:	2.1
Description:	Call Commencement Functionality Description
Inputs:	Call commencement begins with incoming call information, including the caller's identity and request.
Outputs:	The functionality outputs a call that is ready to progress into the conversation.

Table 5.27: Function Description 2.1.1

Function name:	Receive Incoming Calls
Label:	2.1.1
Description:	Receive Incoming Calls Functionality Description
Inputs:	This sub-functionality requires call-related data to be received and processed, ensuring it is ready for redirection.
Outputs:	The functionality outputs an incoming call that is prepared for further processing.

Table 5.28: Function Description 2.1.2

Function name:	Redirect Calls to Webhook
Label:	2.1.2
Description:	Redirect Calls to Webhook Functionality Description
Inputs:	The call redirection process requires details about the destination and the specific call to be redirected.
Outputs:	It outputs the redirected call to the specified destination, such as a webhook.

Table 5.29: Function Description 2.2

Function name:	Welcome Message Playback
Label:	2.2
Description:	Welcome Message Playback Functionality Description
Inputs:	Playback of a welcome message involves the message content and the target caller.
Outputs:	The functionality outputs the played welcome message, initiating the conversation.

Table 5.30: Function Description 2.2.1

Function name:	Play Welcome Message
Label:	2.2.1
Description:	Play Welcome Message Functionality Description
Inputs:	Playing the welcome message requires the message's content and the recipient's details.
Outputs:	It outputs the welcome message played to the caller, setting the tone for the conversation.

Table 5.31: Function Description 2.3

Function name:	Audio Input Reception
Label:	2.3
Description:	Audio Input Reception Functionality Description
Inputs:	Audio input reception involves recording the caller's speech and processing the audio data.
Outputs:	The output includes recorded audio data, ready for further processing.

Table 5.32: Function Description 2.3.1

Function name:	Record Caller's Audio Input
Label:	2.3.1
Description:	Record Caller's Audio Input Functionality Description
Inputs:	This sub-functionality requires capturing and storing the caller's audio input.
Outputs:	It outputs the recorded audio data, preserving the caller's speech.

Table 5.33: Function Description 2.3.2

Function name:	Process Recorded Audio Input
Label:	2.3.2
Description:	Process Recorded Audio Input Functionality Description
Inputs:	The processed audio input requires details about the format and characteristics of the audio data.
Outputs:	It outputs the audio input in a suitable format for further conversion.

Operational Phase Functionalities (Part 2)

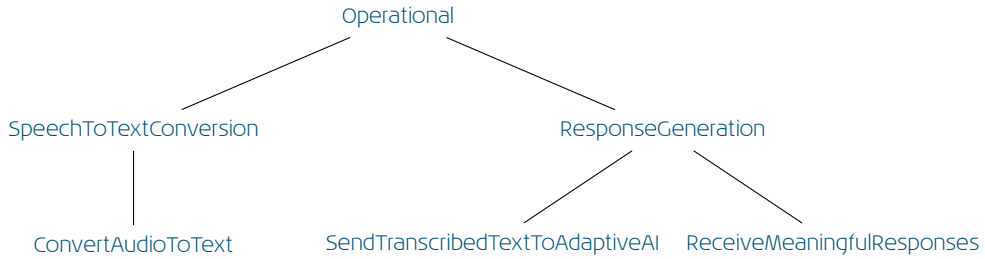


Figure 5.3: Operational Phase Functional Hierarchy Tree (Part 2)

Table 5.34: Function Description 2.4

Function name:	Speech-to-Text Conversion
Label:	2.4
Description:	Converts audio to text
Inputs:	Audio input from the caller
Outputs:	Textual representation of the caller's speech

Table 5.35: Function Description 2.4.1

Function name:	Convert Audio to Text
Label:	2.4.1
Description:	Transcribes audio data
Inputs:	Processed audio input
Outputs:	Transcribed text data

Table 5.36: Function Description 2.5

Function name:	Response Generation
Label:	2.5
Description:	Generates meaningful responses
Inputs:	Transcribed text and response generation instructions
Outputs:	Meaningful responses in textual format

Table 5.37: Function Description 2.5.1

Function name:	Send Transcribed Text to Adaptive AI
Label:	2.5.1
Description:	Transmits transcribed text
Inputs:	Transcribed text and AI interaction instructions
Outputs:	Transcribed text sent for response generation

Table 5.38: Function Description 2.5.2

Function name:	Receive Meaningful Responses
Label:	2.5.2
Description:	Receives and analyzes AI responses
Inputs:	AI-generated responses and response analysis settings
Outputs:	Meaningful responses ready for further processing

Operational Phase Functionalities (Part 3)

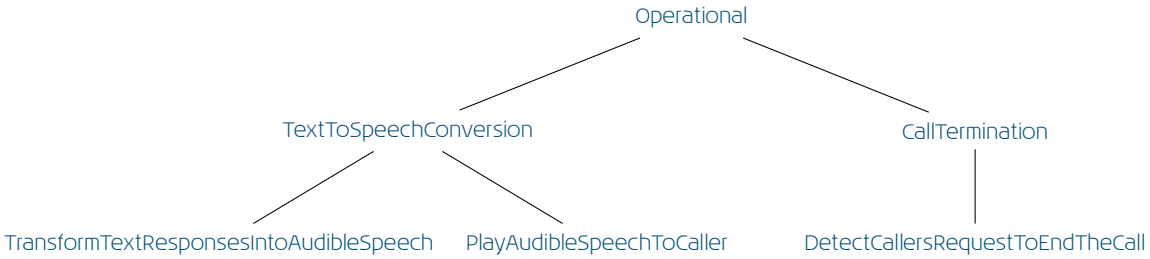


Figure 5.4: Operational Phase Functional Hierarchy Tree (Part 3)

Table 5.39: Function Description 2.6

Function name:	Text-to-Speech Conversion
Label:	2.6
Description:	Converts textual responses to audible speech
Inputs:	Textual responses
Outputs:	Audible speech responses

Table 5.40: Function Description 2.6.1

Function name:	Transform Text Responses into Audible Speech
Label:	2.6.1
Description:	Converts textual responses into audible speech
Inputs:	Textual responses
Outputs:	Audible speech data

Table 5.41: Function Description 2.6.2

Function name:	Play Audible Speech to Caller
Label:	2.6.2
Description:	Plays audible speech to the caller
Inputs:	Audible speech data
Outputs:	Audible speech played to the caller

Table 5.42: Function Description 2.7

Function name:	Call Termination
Label:	2.7
Description:	Ends the call
Inputs:	Caller's request or predefined conditions
Outputs:	Call termination confirmation

Table 5.43: Function Description 2.7.1

Function name:	Detect Caller's Request to End the Call
Label:	2.7.1
Description:	Identifies the caller's request to end the call
Inputs:	Caller's request signals
Outputs:	Confirmation of the caller's request

5.1.5. Activity diagram

The diagram in figure 5.5 illustrates the activity flow for the requested phone application incorporating all main functionalities as defined in section 5.1.4. It is divided into two main phases: Deployment and Operational. The Deployment Phase outlines the setup process, while the Operational Phase details the sequence of actions during an incoming call. The diagram visually represents the various steps involved, from the initialization of the application to the call termination, providing an overview of the entire process.

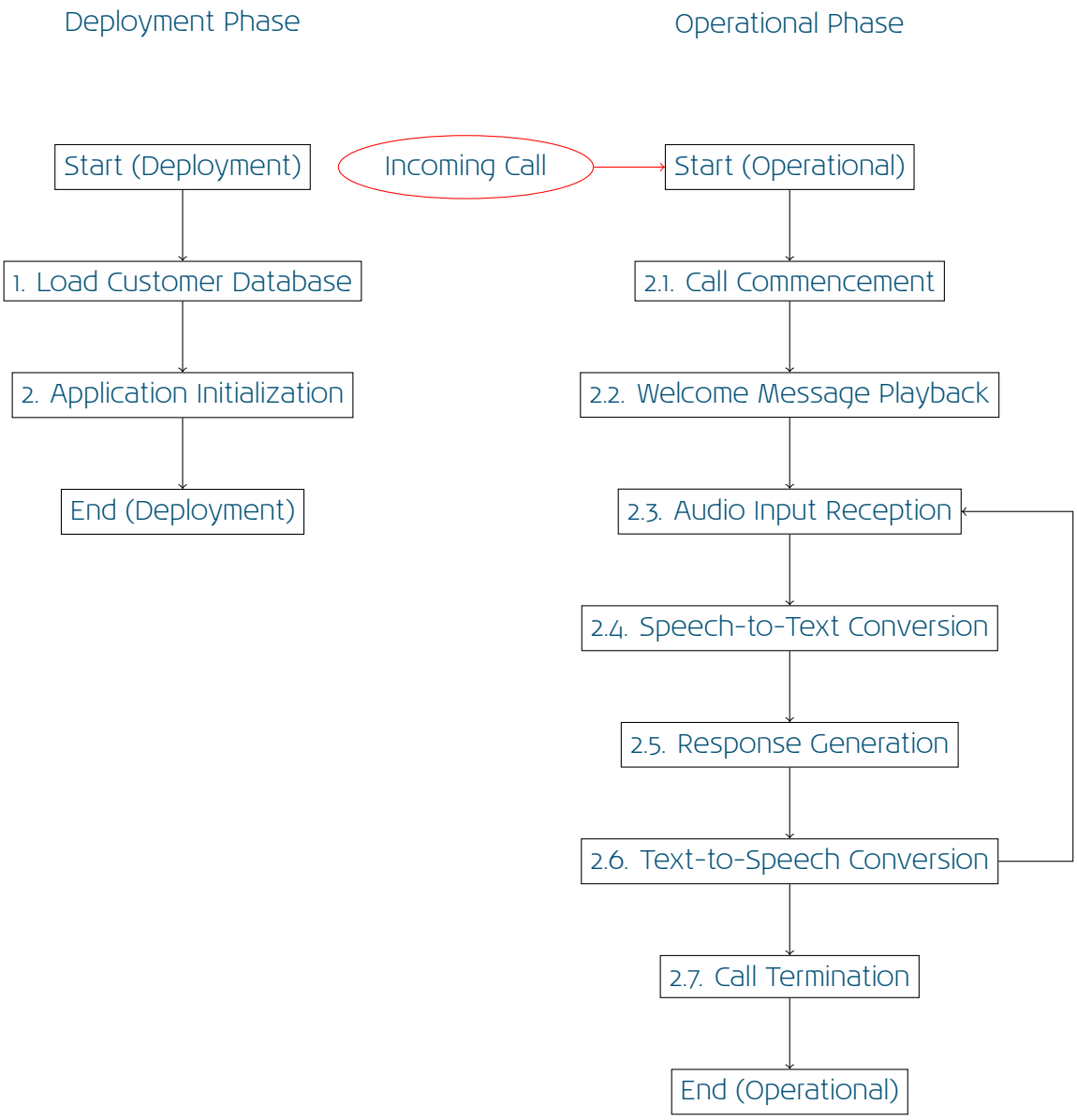


Figure 5.5: Activity Diagram

5.2. Full Functional description of the software

Our developed software exemplifies a fusion of artificial intelligence and business, providing a platform for advanced AI voice communication. It seamlessly integrates cutting-edge features such as speech recognition, real-time processing, and interactive voice response (IVR). This subchapter is dedicated to elucidating the software, offering insights into its components, processes, user experience, and the underlying rationale.

At the core of our software lies the concept of Interactive Voice Response (IVR), a technology that enables computer-human interactions through voice [source]. However, our system differentiates itself from traditional IVR systems by harnessing the power of artificial intelligence to facilitate natural language processing, comprehend intent, and provide responses that closely resemble human communication.

Section 5.1 has guided us through a transformative journey, starting from an abstract

usage scenario and culminating in the detailed definition of the phone application's essential functionalities. This section delves into our iterative development process, outlining the evolution that led to our successful product. It begins with an overview of the three primary iterations and concludes with a comprehensive exploration of the final software product.

5.2.1. Iteration 1

Before embarking on a new project from scratch, it's always prudent to scour the internet for existing software solutions that might fulfill all the functionalities outlined in Section 5.1. We diligently followed this approach and discovered that no prior applications provided the comprehensive feature set we had defined. However, during our research, we did come across information related to the development of non-real-time web applications capable of engaging in meaningful conversations through voice. This material covered functions 5.31 through 5.39 and introduced us to OpenAI's API for response generation. It also introduced us to the concept of prompt engineering, which allowed us to integrate our customer specific database, as mentioned in 5.20, into OpenAI's API conversation. This approach essentially enabled us to instruct the chatbot to emulate specific individuals with their distinct characteristics. To achieve human-like speech, we leveraged elevenlabs for speech generation, producing a voice that sounded remarkably authentic.

The activity diagram for this initial pipeline is presented in Figure 5.6. Following the DevOps methodology, we set up the web application for testing and had KLM employees give it a try. They were pleasantly surprised by the rapid development of our first functional pipeline. The voice quality was excellent, the conversational depth was satisfactory, and the adaptability to simulate different customers was commendable. However, it was evident that there were essential areas for improvement, including the integration of phone capabilities instead of relying on a web application and addressing latency issues, as prolonged pauses before playing speech to the caller affected the overall immersion.

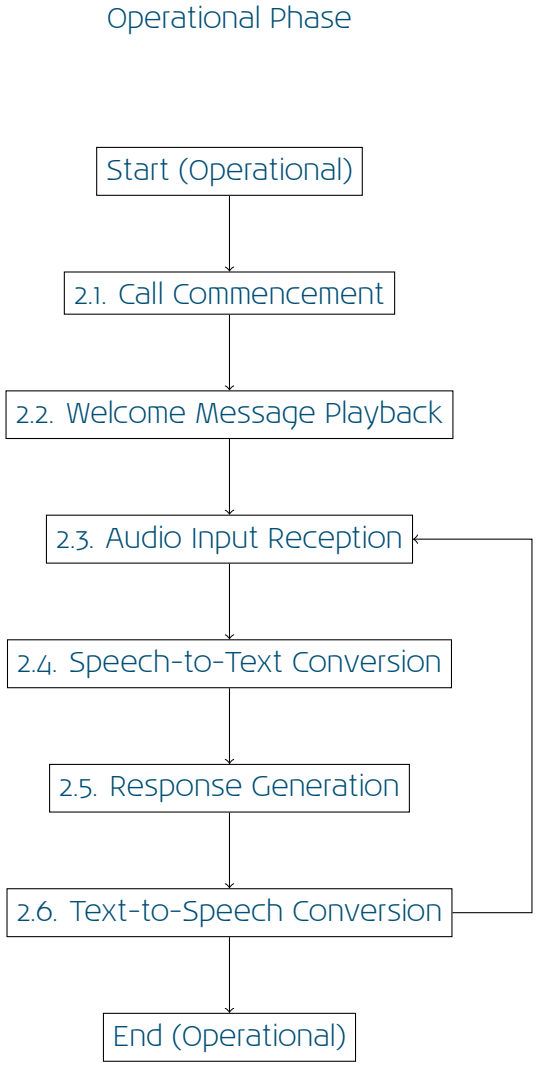


Figure 5.6: Activity diagram of the operational phase of the first iteration

5.2.2. Iteration 2

The second iteration primarily focused on transitioning from the conceptual stage to a tangible product. This transformation involved integrating an Interactive Voice Response (IVR) system with the backend of the web application, as initially outlined in Section 5.2.1. To achieve this, we began by exploring the feasibility of utilizing open-source software known as fonoster. The integration process with fonoster showed promise, and we made significant progress in linking our backend with this system. However, we encountered a major setback when the fonoster API unexpectedly went offline towards the final stages of the integration. This incident served as a valuable lesson, emphasizing the importance of prioritizing stable software solutions, especially when operating within tight project timelines.

In response to the API issue, we transitioned to Twilio's voice API, marking the initiation of our second product iteration. This iteration successfully encompassed all the functionalities detailed in Section 5.1, and the activity diagram was now complete, as illustrated in Figure 5.7. We introduced this product to KLM, and the initial response from KLM employees was positive. However, we still faced a significant challenge in the form of latency.

Within the second iteration, we implemented a Python Flask app as the product server. This server was made accessible via the internet through a webhook generated by ngrok, facilitating Twilio's access to handle incoming calls. The Flask app featured both Post and Get endpoints. Reviewing Figure 5.7, audio input reception involved recording the caller's audio and saving it in a file. This file was subsequently transcribed during the speech-to-text conversion. Once the text conversion process was complete, the text was sent to OpenAI for response generation. When the entire response was generated, it was forwarded to elevenlabs for text-to-speech conversion. We would wait for elevenlabs to finish the entire conversion and save the audio in a file. Finally, the audio was played back to the caller. With some available time remaining, we proceeded with a final iteration to transform the application into a real-time phone system.

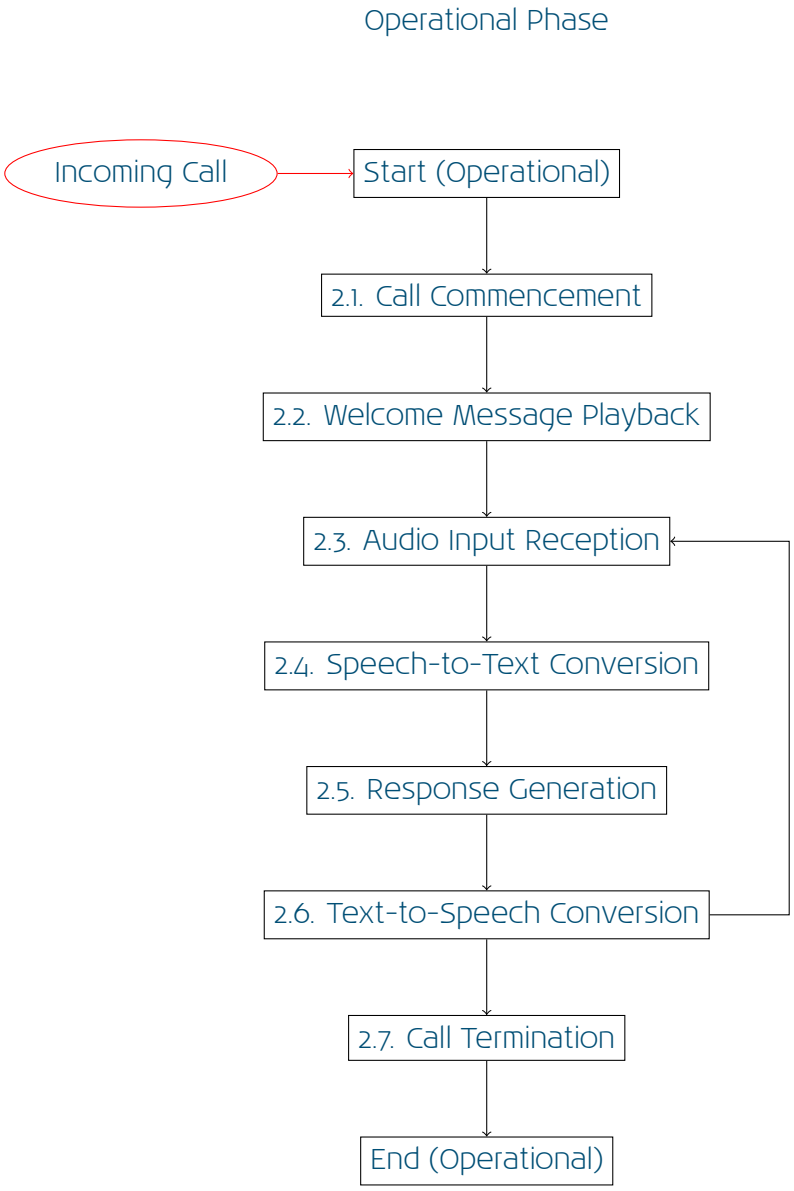


Figure 5.7: Activity diagram of the operational phase of the second iteration

5.2.3. Iteration 3

The third iteration represented the apex of our software development journey, culminating in the final product. This phase signified a remarkable shift towards achieving near-real-time latency by implementing streaming technology, in contrast to the previous reliance on traditional HTTP methods such as post and get.

Revisiting the operational part of the activity diagram in Figure 5.7, the 3rd iteration leveraged streaming technology at three pivotal junctures. Firstly, the speech-to-text conversion process was adapted to a streaming format, ensuring that transcription occurred simultaneously with the caller's speech. By the time the caller concluded their speech, the transcription was nearly complete, resulting in a significant reduction in latency compared to the audio file transcription method used in the second iteration.

As in the previous iteration described in Section 5.2.2, the entire transcription was sent to OpenAI for response generation. However, since the OpenAI response was generated in chunks, we implemented streaming once more, this time between the response generation and the speech generator. This arrangement enabled a continuous flow of text chunks, empowering elevenlabs to commence the synthesis of human speech while OpenAI was still in the process of generating the response. This strategic implementation further contributed to a significant reduction in latency.

Finally, the third streaming connection bridged the two sub-functions of the text-to-speech conversion process. This stream traveled from subfunction 2.6.1, responsible for "Transforming Text Responses into Audible Speech," to subfunction 2.6.2, tasked with "Playing Audible Speech to the Caller," as detailed in Tables 5.20, 5.21, and 5.22. With this configuration, as the "Transform Text Responses into Audible Speech" function generated speech, it was instantly streamed to the caller through the "Play Audible Speech to Caller" functionality. This streamlined approach ensured that as soon as an audio chunk was ready, it was promptly delivered to the caller without the need to wait for the entire adaptive AI response to be fully generated in advance, further minimizing latency.

The introduction of the streaming functionality significantly reduced latency, but the application was still not entirely real-time. Reflecting on human voice communication, we realized that even in human conversations, there are pauses before we speak. During these pauses, we often express our thought process with disfluencies such as "uhmm," "aah," "I see," and others. This observation led to the final feature of our product, where we integrated a database of pregenerated disfluencies. These disfluencies could be played when the first speech chunk was still being processed and before being transmitted to the caller. With this enhancement, we are proud to announce the development of a nearly real-time phone application.

5.2.4. Final software product

The architecture of the software is modular, consisting of several key components that work in unison to deliver a seamless experience. These include the Voice Interface, Speech Recognizer, Conversation Manager, Response Generator, and Audio Processor. Each component plays a pivotal role in the overall functionality of the system.

Voice Interface: The Voice Interface is the first point of contact between the caller and the software. It works just like a business client who answers the phone and points it in the right direction. Regardless of caller experience with technology, the interface is meant to be friendly and lifelike, giving the impression that a real call has started with a client.

Speech Recognizer: The Speech Recognizer is the component that carefully analyzes the spoken words from the call and transcribing them into text. It performs this task using specialized software designed for the purpose, which we'll delve into in a later chapter. The reason for using text is because we create a format that an advanced AI Language Model can interpret and respond to. It's important for this process to be accurate because the quality of the conversation with the chatbot depends on its ability to understand the transcribed text as if it were a message typed by a human. This step is the bridge that connects human speech to the understanding of the chatbot. We will look deeper into the reasons behind choosing which AI language model and other technologies for our software and the benefits it offers to our software in a later section.

AI Language Model: The AI Language Models acts as our conversation manager, which is the brain of the conversation, interpreting the transcribed text and maintaining the context of the conversation. It keeps track of the dialogue flow, manages the timing of responses, and makes sure that the conversation progresses logically.

Response Manager: By using realistic text-to-speech and voice synthesis services, our software constructs replies that are coherent, adjustable, and formulated in a conversational style. This component is what enables the software to simulate human-like interactions. The response manager is responsible for converting the generated text responses from the language model back into audio. It employs the mentioned text-to-speech and voice synthesis technology to create a voice that is clear, natural and (near) realistic.The processor also ensures that the audio is optimized for transmission over phone lines, maintaining the quality of the conversation.

Workflow of the System The workflow of the system is a sequence of events, it starts by an outgoing call from the caller that initiates a conversation between the software and the caller. The process proceeds as follows:

1. Call Initiation: The system receives an incoming call from the caller and initiates the Voice Interface.
2. Greeting: The caller is greeted by a pre-recorded welcome message.
3. Speech Capture: As the caller speaks, their voice is captured by the system in real-time.
4. Transcription: The captured speech is transcribed into text by the Speech Recognizer.
5. Context Analysis: The Conversation Manager analyzes the intent behind the transcribed text.
6. Response Formulation: The Response Generator formulates an appropriate response based on the analyzed intent.
7. Audio Conversion: The formulated response is converted back into audio by the Audio Processor.
8. Delivery: The response is delivered to the caller, completing the interactive cycle.

5.2.5. Comparative analysis for Speech-To-Text technologies

In the course of our research, we found the study titled ***A Comparative Analysis of Real Time Open-Source Speech Recognition Tools for Social Robots*** by (Pande et al., 2023) to be a leading work in the field. We followed the insights of this recent study.

For the initial evaluation of the speech recognizer, we explored various speech recognition software solutions, rating and selecting them based on specific criteria. We will start

with the details of the criteria before presenting a table that summarizes the evaluation metrics used to compare the different software options.

When assessing speech recognition software, several key factors are critical to ensure that the selected system aligns with the application's requirements. The following were the criteria considered:

- **Accuracy:** Measured by the metric Word Error Rate (WER).
- **Latency:** The time required for the speech recognizer to process spoken input and return the text output.
- **Language Support:** The number of languages the speech recognition software can understand and process.
- **Offline Capability:** Whether the speech recognition can function without an internet connection.
- **Ease of Integration:** How simply the software can be incorporated into existing systems.
- **Resource Requirements:** The computational power and memory needed for the software to operate effectively.
- **Cost:** The financial implications of using the software, including initial setup and ongoing usage fees.

Given these criteria, five open-source speech recognition tools were examined: Google Speech Recognition, Sphinx, DeepSpeech, Vosk, and WhisperAI.

Criteria/ Software	Google Speech	Sphinx	Deep-Speech	Vosk	WhisperAI
Word Error Rate (WER)	42.3%	High	High	51.7%	9%
Latency	Low	High	Medium	Low	Medium
Language Support	Extensive	Limited	Good	Good	Good
Offline Capability	No	Yes	Yes	Yes	Yes
Ease of Integration	Moderate	Moderate	Moderate	Easy	Moderate
Resource Requirements	Moderate	Low	High	Low	Moderate
Cost	Free	Free	Free	Free	Free

Table 5.44: Evaluation of Speech Recognition Software

Based on the evaluation, selecting the software involved a balance between accuracy, functionality, and operational constraints. WhisperAI, with the lowest WER, was exceptional in terms of accuracy. However, Vosk was chosen for its offline capability and fastest execution time, despite its higher WER compared to Google Speech Recognition and WhisperAI. As mentioned previously, our focus was on optimizing latency. During the initial development phase, we utilized WhisperAI, but it was slower than Vosk. The minimal resource requirements and straightforward integration of Vosk were also pivotal in its selection, making it an excellent fit for real-time applications where rapid response and offline capability are of the essence.

This evaluation exemplifies that the superior software is not determined solely by accuracy metrics but also by how well it fits the specific constraints and aims of the project.

5.2.6. Comparative Analysis for AI Language Models

This subchapter delves into the comparative analysis of some of the most advanced AI language models: ChatGPT 3.5, ChatGPT 4, Bard, LLaMA, and Claude, which act as our

conversation manager. This component is the brain of the conversation, interpreting the transcribed text and maintaining the context of the conversation.
The following criteria have been established to evaluate and compare the performance and capabilities of the different AI language models:

- **Accuracy:** The exactness of the model's responses to factual information and user prompts.
- **Ease of Use:** The user's ability to interact with and integrate the model into various platforms.
- **Cost:** The economic considerations involved in deploying and maintaining the model.
- **Integration:** The model's compatibility with other software and systems.
- **Response Time:** The delay between the user's request and the model's reply.
- **User Experience:** The overall satisfaction of users based on the model's interaction quality.
- **Creativity:** The model's capacity for producing innovative and imaginative responses.
- **Personalization:** The extent to which the model can customize responses to individual user contexts.
- **Scalability:** The model's ability to maintain performance as demand or load increases.

Criteria	Tortoise	Amazon Polly	Google TTS	ElevenLabs
Naturalness of Voice	Moderate	Good	Good	Excellent
Expressiveness	Limited	Moderate	Moderate	Superior
Filler Sound Naturalness	Poor	Moderate	Moderate	Excellent
Diversity of Voices	Moderate	Extensive	Extensive	Extensive
Real-time Streaming	No	Limited	Limited	Yes
Ease of Integration	Good	Good	Good	Very Good
Cost Efficiency	Good	Moderate	Moderate	Good

Table 5.45: Evaluation of TTS Providers

To asses the models with the criteria, the studies of (Hadi et al., 2023) and (Calonge et al., 2023) were used and studied for critical evaluation. Considering the specific requirements of our real-time conversational system as evidenced in our code, ChatGPT 4 proved to be the best choice among the AI language models we evaluated. Its performance metrics, particularly in accuracy and the generation of contextually relevant responses, aligned very well with our system's architecture. The human-like responses and creativity of ChatGPT 4's interactions were especially crucial for the dynamic and varied queries our application is designed to handle. Furthermore, its scalability ensures that as our system grows and the complexity of user interactions increases, ChatGPT 4 will continue to provide intelligent conversational capabilities.

5.2.7. Evaluation of Text-to-Speech Providers

During the project, we undertook an independent research initiative to identify a text-to-speech (TTS) provider capable of delivering a seamless and realistic conversational experience. Due to the limited research available, we opted to conduct our own comprehensive testing of various TTS providers. Our evaluation was driven by the need to find a service that could overcome the uncanny valley often associated with synthetic voices,

particularly when producing filler sounds and expressing emotions such as "uhm," "ehm," "ah," and laughter.

We examined leading TTS providers, including Tortoise, Amazon Polly, Elevenlabs, and Google Text-to-Speech, by subjecting them to rigorous testing across several critical parameters. These included the naturalness of the voice, expressiveness, ability to handle filler sounds naturally, diversity of voice options, capability for real-time streaming, ease of integration into our systems, and cost efficiency. The following table summarizes the results of our comparative analysis, with ElevenLabs standing out due to its superior performance in key areas, making it the most suitable option for our system's requirements.

Criteria	Tortoise	Amazon Polly	Google TTS	ElevenLabs
Naturalness of Voice	Moderate	Good	Good	Excellent
Expressiveness	Limited	Moderate	Moderate	Superior
Filler Sound Naturalness	Poor	Moderate	Moderate	Excellent
Diversity of Voices	Moderate	Extensive	Extensive	Extensive
Real-time Streaming	No	Limited	Limited	Yes
Ease of Integration	Good	Good	Good	Very Good
Cost Efficiency	Good	Moderate	Moderate	Good

Table 5.46: Evaluation of TTS Providers

Despite the advancements in technology made by each provider, a robotic undertone was still present, particularly with the automated human-like sounds. This deviated from the conversational realism we strived to achieve. ElevenLabs, however, provided a breakthrough with its extensive library of voices that closely emulate human tonality and expressiveness. The customizability offered by Elevenlabs allowed us to tailor the voice to the unique personality of our AI. Furthermore, ElevenLabs' capability for real-time streaming was crucial for our needs, ensuring that our AI could deliver responses with minimised latency, thereby maintaining the natural flow of dialogue.

5.2.8. Integration of Phone Call

During the development of our product, we made use of Twilio's Voice API to provide the voice interaction through a phone call with an AI. Our prototype was engineered for one-on-one user interaction, prioritizing immediate and clear voice communication. The stability and reliability provided by Twilio's network were pivotal in our testing phase, ensuring uninterrupted voice interaction for a comprehensive assessment of the prototype's functionality. Although our prototype did not utilize the full spectrum of Twilio's features, such as advanced security protocols or high-volume call handling, these aspects underscore the platform's capacity for future scalability and enhanced security if required for subsequent iterations. At this juncture, the principal benefit derived from Twilio's Voice API was its dependable performance and the superior quality of voice exchanges, which facilitated an effective evaluation of our prototype's conversational performance (Twilio, 2023).

5.2.9. Architecture

In summary, the design of our software skillfully combines advanced technologies to create a user-friendly and efficient system. Vosk was selected for its ability to quickly and accurately recognize speech, even without an internet connection. When it comes to converting text into speech, ElevenLabs proved superior, offering a voice that sounds

natural and can express a wide range of human sounds and emotions.

At the heart of our language processing, we have ChatGPT 4, which excels in understanding and responding to complex user interactions. Its ability to scale means it can handle increasingly intricate conversations as our user base grows.

Twilio's Voice API acts as the bridge, connecting all these elements together into a seamless voice communication platform. It's a flexible and robust system that adapts and grows with the needs of our users.

Overall, this system stands out for its integration of leading-edge components, ensuring that our software not only meets the immediate needs of our users but is also well-prepared to incorporate future advancements in the relevant technologies. All the components come together in the architecture below.

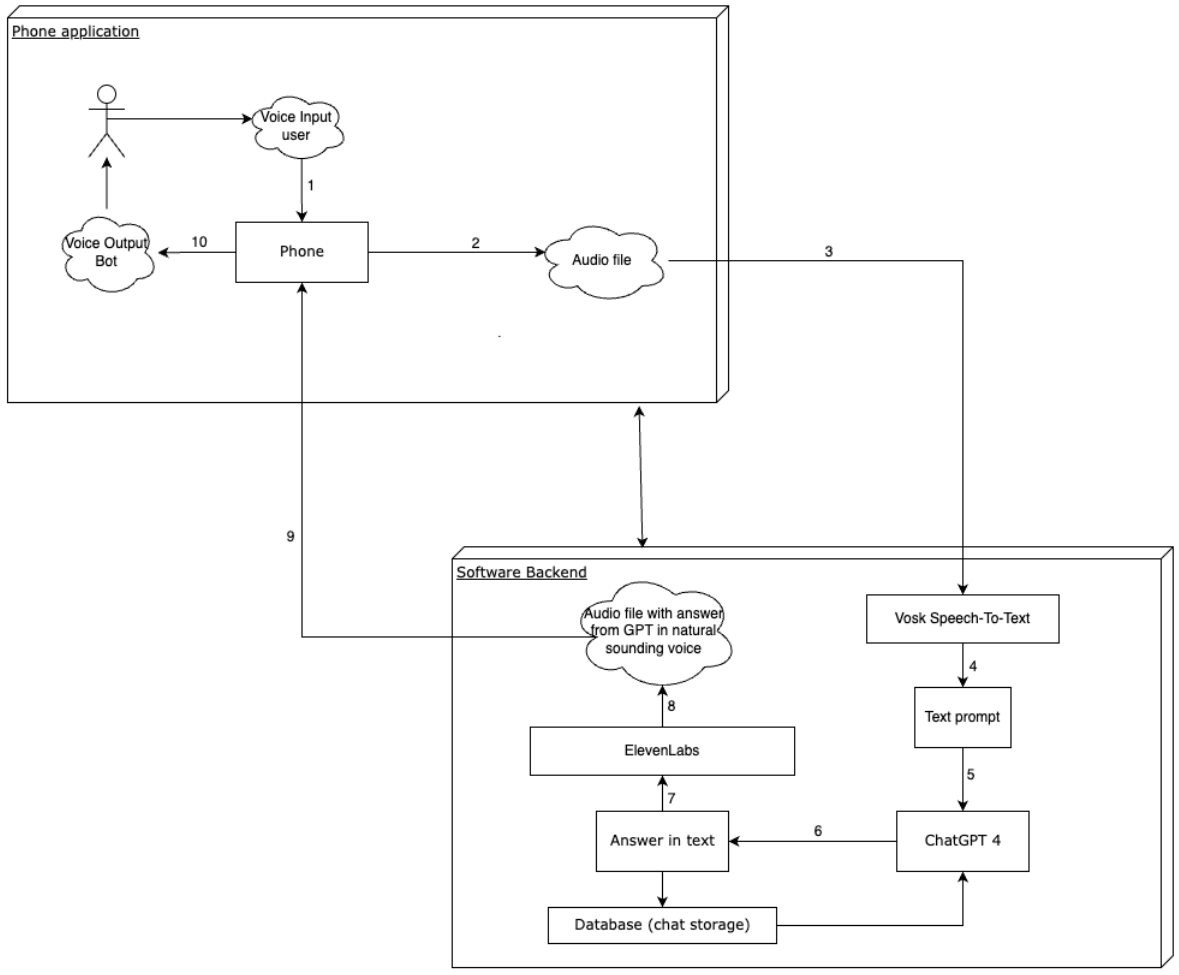


Figure 5.8: System Architecture

5.3. Research Methodologies

5.3.1. Global Setup

To address the aim of this explorative study which was a feasibility study to test the AI communication training toolkit using Adaptive AI , we divided the explorative problem statement into two different feasibility analysis a)Quantitative Analysis and b) Qualitative Analysis, which in turn justify the explorative search study problem statement as given in the introduction. The subsections of this analysis are:

5.3.2. Data Collection

For data collection, we chose to perform multiple triangulation as we conducted a qualitative survey, did a literature review study and structured one-on-one interviews. By using three different methods and data sources, both methodological and data triangulation were intended to be achieved. (A. Bans-Akutey, 2021). If done right, validity is established if the conclusions drawn from the findings of the various methods are consistent. (Franklin et al., 2001).

5.3.3. Quantitative Analysis

In order to measure the impact of the newly developed application, an experiment was set up and prepared for the participants who volunteered in testing the new employee sales application. The primary aim of the experiment was to engage employees who were well-versed with the existing employee trainings and to have them assess the quality of the training and the learning experience of the newly developed AI training. This assessment encompassed key factors outlined in our problem statement, including unpredictability, believability, and intuitiveness. The data collection process encompassed the following phases- experiment was set up in 5 stages, over a period of two days, namely:

- A quantitative assessment of the existing employee training programs through the first survey.
- The examination of the newly developed AI employee training program
- A quantitative assessment of the newly developed AI employee training program through the second survey, conducted immediately after participants engaged with the new training.
- Structured interviews that were conducted with participants to gain insightful feedback on the new training program
- A final quantitative assessment of the newly developed AI employee training program through the third survey, carried out one day after participants experienced the training to mitigate the potential influence of enthusiasm bias.

All surveys and interview can be found in Appendix C.

Participants

The pool of participants for this experiment comprises of KLM employees, mostly from the IT department. As illustrated in Table 5.48, a total of 11 respondents completed all three surveys and tested the product. As can be seen from the table, all the respondents are

active in the IT and the XR Centre of Excellence department of KLM, which is also a department within IT, focusing on augmented reality (AR), virtual reality (VR), and mixed reality (MR).

Upon closer examination of the table, it becomes evident that all respondents hail from the IT department within KLM and, some of them more specifically from the XR Centre of Excellence department within IT, which specializes in augmented reality (AR), virtual reality (VR), and mixed reality (MR).

While the initial intention was to arrange sales employees for the experiment, It proved out to be challenging since a significant number of them were on vacation the week prior to the experiment. Consequently, the research necessitated the inclusion of the directly available IT employees, albeit an imperfect substitution. This can be considered a form of non-response bias, because the people who were intended to be part of the study do not participate, which can lead to a bias in the study's findings.

Moreover, a sample size of 11 is relatively small, and as such, it poses limitations. These limitations include reduced generalizability for instance, but also an increased risk of Type II errors, which occur when failing to detect a true effect due to the small sample size (Harmon & Losos, 2005). Furthermore, there is an elevated potential for selection bias.

During the course of the experiment, the IT team members were instructed to assume the roles of sales employees tasked with selling tickets to B2B clients. This is of course an inherent limitation and potential bias that must be considered when interpreting the study's findings. The complete instruction of the experiment can be found in the appendix.

Questionnaires

As previously mentioned, there were three types of surveys conducted. In the first survey, respondents were asked about their personal characteristics, including gender, as well as their department and position within the company. Furthermore, each participant was assigned a unique identification number to maintain data anonymity and facilitate the recognition of individuals and their prior responses in subsequent surveys and during the conducted interview.

Subsequently, the primary measurement objectives of the surveys can be derived from the problem statement:

"The problem is how to enhance the existing standardized and predictable employee training programs at KLM, starting with the Learning and Development sector within sales, by incorporating adaptive AI. The aim is to create intuitive, easy-to-set-up training modules that incorporate unpredictable scenarios, enhancing employee motivation and confidence, while maintaining the believability of the training."

Thus, the survey aimed to measure whether there is a (positive) difference between the new training and the old training in terms of the following five main factors:

- Overall training quality
- Level of intuitiveness
- Level of unpredictability

- Level of difficulty
- Level of believability

The ratings for these factors were measured on a continuous scale ranging from 0 to 10, with a precision of one decimal place. To enhance the respondent's experience and make the ranking process more intuitive, a colored slider was employed. This slider was chosen to evoke a particular sentiment from the respondent based on their experience. However, as shown in Figure 5.9, when respondents moved the slider, they could also see the corresponding numerical value on the screen, allowing them to compare it with their previous responses in the subsequent surveys."

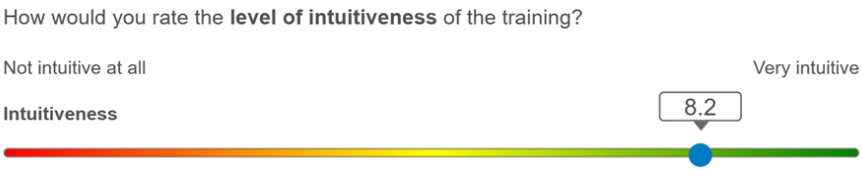


Figure 5.9: Quantitative Analysis Slider

Subsequently, a set of in-depth questions was posed, and respondents provided their answers using an ordinal Likert scale ranging from 1 to 5. These four questions were related to the following aspects:

- The training adequately prepares me and provides enough confidence for communicating with sales customers.
- The training adequately prepares and motivates me for communicating with sales customers.
- The training incorporates unpredictable scenarios that simulate real-world unpredictability.
- The training enhances cross-cultural communication skills, deepens understanding of cultural differences, and enables us to better meet the diverse customer needs.

In the second and third surveys, the respondents were also asked about the AI bot of the newly developed training, specifically whether they thought the bot's voice and response time felt realistic, whether the conversation with the AI bot had sufficient depth, and how feasible and effective they thought such a product would be within their department as part of the training programs. These were also based on a Likert scale ranging from 1 to 5. All 9 statements can be seen in (5.47).

Finally, respondents were also asked to describe the training in three words to provide a comprehensive understanding of their perception. The complete surveys can be referenced in the appendix B.

Table 5.47: All Statements

Statement	Description
Statement 1	The training adequately prepares me and provides enough confidence for communicating with sales customers
Statement 2	The training adequately prepares and motivates me for communicating with sales customers
Statement 3	The training incorporates unpredictable scenarios that simulate real-world unpredictability
Statement 4	The training enhances cross-cultural communication skills, deepens understanding of cultural differences, and enables us to better meet the diverse customer needs
Statement 5	The AI bot's response time in a typical conversation feels realistic
Statement 6	The AI bot has a realistic voice
Statement 7	The conversation with the AI bot had sufficient depth regarding the content
Statement 8	The introduction of the AI bot would be feasible in our department
Statement 9	The introduction of the AI bot would enhance the effectiveness of our current employee training programs

5.3.4. Qualitative Analysis

The data from our quantitative study was analyzed by the team and helped in forming the qualitative questions. Herein, demographic information and unpredictability touch points were conducted. Eventually resulting in a set of probes which we used to validate and sensitize our participants in the one-on-one structured interview. This Interview would be conducted after the participants would be sensitized with the demonstration of the current AI assistant or a AI Communication Training tool which is a proof of concept where a human interacts with a specific AI persona, thus providing a mock training assistant for Sales representatives at KLM.

The set of questions are followed by the guidelines by Developing qualitative research questions: a reflective process by Agee, 2009 and Qualitative interviewing and grounded theory analysis by Charmaz, 2002, which focuses on six types of questionnaires which are as follows 1) Demographic questions 2) Quality of training questions 3) Intuitiveness of training questions 4) Unpredictability of training questions 5) Believability of training questions 6) Feasibility of training questions .By combining the 6 types of questions with the mostly 3 different tenses (present, past, future) in few cases it can provide a holistic

approach towards answering the Problem Statement and making the existing proof of concept feasible, viable and desirable.

Structured one-on-one interviews

Structured one-on-one interviews were conducted to gather qualitative data from technical experts and project managers who had experienced an immersive experience of VR and Adaptive AI at Extended Reality labs in KLM. This interview aimed to allow participants to express their behaviors, feelings, knowledge, background and views towards the use of Adaptive AI during communication training programs intended for Sales representatives. The main topics in the interview were:

- **Role and Work Context:** Understanding participants' roles, departments, and positions, including factors influencing their work.
- **Training Quality:** Reflecting on past experiences' influence on perceptions and identifying key factors affecting AI training quality.
- **Intuitiveness:** Exploring elements that enhance or hinder the intuitiveness of AI training and suggestions for improvement.
- **Unpredictability:** Delving into specific scenarios showcasing training unpredictability and its impact.
- **Believability:** Analyzing factors contributing to the believability of AI training, such as response time, realism, and depth.
- **Feasibility:** Investigating the factors considered when assessing the feasibility of AI integration within the department.

To collect detailed insights into these topics, 10 interview questions were developed based on information gathered from our literature review and the probes based on the Quantitative survey analysis. We conducted one pilot survey interview with Participant P11 to gather test results for Quantitative survey and adjust the Qualitative interview if needed, due to which Qualitative interview wasn't performed on Participant P11. A structured approach was chosen to improve comparability, as 3 researchers conducted the interview individually.

Acquiring as many participants was not the focus of this interview study (table 1); rather, the aim was to reach the point of saturation where the collected data began to provide little new information (Nielsen & Landauer, 1993). Before the interview, participants were provided with a participant Survey sheet after the product demonstration which acted as a sensitizing scenario to get a general idea of what Adaptive AI is and its impact on the Communication training program.

Table 5.48: Composition of final sample of interview participants

Participant ID	Department and Role	Familiarity with AI
P1	XR Lab - Lead Developer	High
P2	VR Gaussian splitting KLM Intern	Low
P3	XR Lab - Intern	Moderate
P4	Innovation track team - External VR Developer	Low
P5	XR Centre of Excellence - External Developer	Low
P6	XR Technical Specialist	Moderate
P7	Visual Designer at KLM	Low
P8	Senior Researcher for new technologies	High
P9	Programmer	High
P10	Programmer	High

5.4. Data Analysis

5.4.1. Quantitative Analysis derived from three surveys

Selecting the correct data analysis method is of paramount importance. Failing to do so may result in unreliable measurements and the unintended capture of data. When choosing the appropriate method, considerations should include:

- The aim of the measurement: What precisely needs to be assessed?
- Scale of the dependent variable: Nominal, ordinal, interval, or ratio?
- Data distribution: Does the data follow a normal distribution, or another distribution, making a non-parametric test more suitable?
- The presence of outliers: Outliers can have a big effect on the outcomes of many statistics tests.

Table 5.49: Paired Sample Statistics

Pair		Mean	Standard Deviation
Pair 1	Overall quality (1)	4.309	1.8289
	Overall quality (2)	7.055	1.9781
Pair 2	Intuitiveness (1)	4.882	1.8367
	Intuitiveness (2)	7.645	1.4439
Pair 3	Unpredictability (1)	1.964	1.8991
	Unpredictability (2)	6.491	2.3645
Pair 4	Difficulty (1)	4.155	2.3847
	Difficulty (2)	3.918	2.1465
Pair 5	Believability (1)	3.227	2.8443
	Believability (2)	7.082	1.7122

In a Paired Samples T Test, the observations are defined as the differences between two sets of values. Paired samples can appear in different designs, one of them being when the same participants are measured at two points in time, which is the case for our experiment, since the KLM Employees assessed the training based on the five factors mentioned in 5.3.3, before and after having tested the new AI training (Rietveld & van Hout, 2017).

Therefore, a Paired Samples T-Test is suitable for this experiment, and it will allow us to measure whether there is a significant difference between the means for the scores on the five factors of the old employee training compared to the new AI training program. In addition, a Paired Samples T-Test is appropriate due to the data being of the interval type. Furthermore, normality assumptions were tested and found to be met.

The first observation in 5.49 pertains to the mean values of the five crucial factors, with a distinction made between the old employee training (1) and the new AI employee training (2). Notably, we observe that the "Overall Quality" of the AI training, with an average score of approximately 7.0, stands out as it is, on average, 2.7 points higher than the old employee training, which received an average rating of 4.3. This same difference is evident for "Intuitiveness" in the new training, which scores, on average, 2.8 points higher than the current training.

The most significant contrast, however, can be seen in "Unpredictability," with the new training receiving an average rating of 6.5, compared to only 2.0 for the current training. Additionally, "Believability" of the new training far surpasses that of the old training, with an average score of 7.1 in contrast to 3.2.

Conversely, the "Difficulty" of the new training does not score higher, as it has an average score of 3.9, compared to 4.2 for the current training. This is noteworthy given the significantly higher unpredictability of the new training, where one might expect that increased unpredictability would lead to a higher difficulty rating. One potential explanation could be how respondents interpreted the question. Interviews revealed that the question about difficulty, shown in figure 5.10 was highly open to interpretation, and people interpreted it differently. Some respondents had difficulty choosing an interpretation, which suggests that the value of this question for the results should be adjusted. In future research, a more specific formulation of the question is therefore preferred.

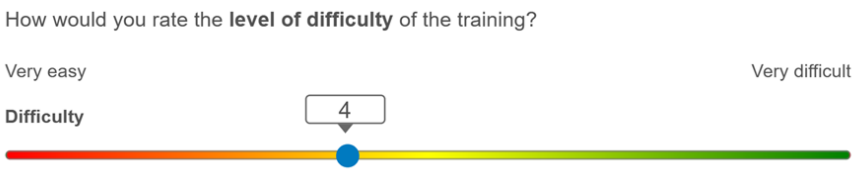


Figure 5.10: Quantitative Analysis

The standard deviation serves as a measure of the spread and assesses how individual ratings vary from the mean (Cleophas & Zwinderman, 2016). A higher standard deviation indicates greater variability in scores. Notably, for the "Believability" factor, the standard deviation is 1.7 for the new training compared to 2.8 for the old training. This suggests that there was less variability in scores after the new training, signifying greater agreement among respondents regarding the average "Believability" score of the new training. For the other factors, the differences in standard deviation are closer to each other.

Table 5.50: Paired Samples Test

Pair	Paired Differences	Mean	95% Confidence Interval		Two-Sided p
			Lower	Upper	
1	Overal quality (1) - Overall quality (2)	-2.7455	-4.5696	-0.9213	0.007
2	Intuitiveness (1) - Intuitiveness (2)	-2.7636	-4.4098	-1.1174	0.004
3	Unpredictability (1) - Unpredictability (2)	-4.5273	-6.8389	-2.2156	0.001
4	Difficulty (1) - Difficulty (2)	0.2364	-1.4611	1.9339	0.763
5	Believability (1) - Believability (2)	-3.8545	-6.0495	-1.6596	0.003

In the preceding Table 5.49, it appeared that there were differences between the scores before and after participants engaged with the new training. Nevertheless, an important

question remains: Is this difference statistically significant? It is important to note that a conventional confidence level of 95% has been employed, which implies that variables are not considered significant if the p-value is less than 0.05.

When assessing the significance for the first pair, "Overall quality," it becomes apparent that the two-sided p-value is 0.007. This p-value represents the likelihood that the observed difference arose purely by chance. Expressed as a percentage, we can assert that there is a mere 0.7% chance that the observed disparity in scores for the "Overall quality" factor, as applied to the old and new training, is a result of random chance. As previously mentioned, the test has been conducted at a 5% alpha level; thus, a p-value less than 0.05 is sought. In this instance, that condition is met. Consequently, this means that statistical evidence has been acquired indicating that the AI training surpasses the current standardized employee training in terms of overall quality, implying that the new AI training has had a significant influence.

Similarly, it can be asserted that the observed differences for the factors of "Intuitiveness," "Unpredictability," and "Believability" are also statistically significant. Therefore, the higher mean scores that were observed earlier in the new training compared to the old training can be attributed to the fact that the new training has had a significant impact, leading to superior ratings in terms of overall quality, intuitiveness, unpredictability, and believability.

Conversely, the same cannot be said for the "Difficulty" factor. The p-value for "Difficulty" is 0.76, which exceeds the chosen significance level (e.g., 0.05). This p-value of 0.76 indicates that the observed difference in mean scores for "Difficulty" is likely to have occurred due to random chance alone. In other words, there is insufficient statistical evidence to conclude that a significant difference exists in "Difficulty" between the two datasets. As mentioned earlier, a potential explanation may lie in the lack of specificity in the wording of the question related to difficulty, which allowed for interpretation.

A Wilcoxon Signed Rank Test will be performed to examine whether there are significant differences in the ratings of the first four statements from Table 5.47. These statements relate to employees' confidence and motivation in communicating with customers, the incorporation of unpredictable scenarios in the training, and the promotion of cross-cultural communication skills. Since the claims were rated on a Likert range from 1 to 5, the Wilcoxon Signed Rank Test is a good choice for this analysis. This test works like a Paired Samples T Test, but it's designed to work with non-parametric data like ordinal factors, which is befitting this dataset.

Table 5.51 indicates that the p-value for the difference in assessments of Statements 1, 2, and 4 in the first two surveys is below 0.05, signifying statistical significance in this Wilcoxon Signed Rank Test. This implies the rejection of the null hypothesis, providing evidence of a statistically significant difference for these statements. Looking at Table 5.52 below, this suggests that the AI bot better prepares KLM employees, instilling them with more confidence in communicating with sales customers, and motivates them more for such interactions. Furthermore, the AI bot contributes to the enhancement of cross-cultural communication skills and a deeper understanding of cultural differences.

In contrast, for Statement 3, the p-value of 0.098 is not statistically significant, indicating that, based on this research, it cannot be concluded that the AI training effectively

Table 5.51: Wilcoxon Signed Rank Test

Null Hypothesis	P-value
The median of differences between the assessment of Statement 1 in Surveys 1 and 2 equals 0. Related-Samples Wilcoxon Signed Rank Test	,031
The median of differences between the assessment of Statement 2 in Surveys 1 and 2 equals 0. Related-Samples Wilcoxon Signed Rank Test	,098
The median of differences between the assessment of Statement 3 in Surveys 1 and 2 equals 0. Related-Samples Wilcoxon Signed Rank Test	,032
The median of differences between the assessment of Statement 4 in Surveys 1 and 2 equals 0. Related-Samples Wilcoxon Signed Rank Test	,018

incorporates unpredictable scenarios simulating real-world unpredictability. From participant feedback it already became clear that the detailed experiment instructions, as found in the appendix, may have contributed to this result, as participants were aware of what to expect, diminishing the surprise element. One participant even said after the experiment, "The predictability is quite high since I knew what I was tasked with to do." Future research could involve refining the AI bot and conducting a new experiment with less detailed instructions to preserve this the element of surprise, allowing for a more unpredictable training.

Table 5.52: Mean Analysis

Pair		Mean	Standard Deviation
Statement 1	Survey (1)	2.55	1.214
	Survey (2)	4.00	0.894
Statement 2	Survey (1)	3.00	1.183
	Survey (2)	4.00	1.000
Statement 3	Survey (1)	2.27	1.009
	Survey (2)	3.64	1.027
Statement 4	Survey (1)	2.73	1.009
	Survey (2)	3.91	0.701

Now that survey 1 and 2 have been compared, the same analysis will be performed for surveys 2 and 3.

As observed in Table 5.53, the differences in the average scores from Survey 2 compared to Survey 3 for the AI training on various factors are not notably substantial. This suggests that participants' opinions did not significantly change one day after the experiment in relation to the AI bot. This observation appears to eliminate the presence of

Table 5.53: Paired Sample Statistics

Pair		Mean	Standard Deviation
Pair 1	Overall quality (2)	7.060	2.0850
	Overall quality (3)	7.020	2.1170
Pair 2	Intuitiveness (2)	7.710	1.5051
	Intuitiveness (3)	7.710	1.0418
Pair 3	Unpredictability (2)	6.360	2.4500
	Unpredictability (3)	6.080	2.5728
Pair 4	Difficulty (2)	3.990	2.2487
	Difficulty (3)	4.720	1.9263
Pair 5	Believability (2)	7.130	1.7969
	Believability (3)	6.980	2.0826

enthusiasm bias.

Table 5.54 confirms this, as the p-values for all 5 pairs are well above 0.05. This leads to the acceptance of the null hypothesis that there is no significant difference in the average scores from Survey 2 compared to Survey 3 for the AI training on the factors of overall quality, intuitiveness, unpredictability, difficulty, and believability. Consequently, the presence of enthusiasm bias in the AI bot training can be ruled out, and it signifies that the results from Survey 2 have proven to be reliable.

Table 5.54: Paired Samples Test

Pair	Paired Differences	Mean	95% Confidence Interval		Two-Sided p
			Lower	Upper	
1	Overal quality (2) - Overall quality (3)	0.0400	-0.3960	0.4760	0.840
2	Intuitiveness (2) - Intuitiveness (3)	0	-0.9260	.9260	1.000
3	Unpredictability (2) - Unpredictability (3)	0.2800	-1.1379	1.6979	0.666
4	Difficulty (2) - Difficulty (3)	-0.7300	-2.5718	1.1118	0.393
5	Believability (2) - Believability (3)	0.1500	-1.1884	1.4884	0.806

To conclude the quantitative data analysis, 5.55 demonstrates that the participants' responses to Statements 5 through 9, rated on a Likert scale ranging from 1 to 5, align with the positive perception of the training identified in the earlier analyses. The only statement where the AI bot receives a slightly lower score is Statement 5, which relates to the perception of the response time in a typical conversation feeling realistic. However, for the remaining statements, the AI bot consistently receives a score of 4 or higher. These statements revolve around the AI bot having a realistic voice, providing sufficient depth in conversations, being feasible in the participants' departments, and enhancing the effectiveness of the current employee training programs through its introduction.

Table 5.55: Mean analysis

Statement	Mean
Statement 5	2.50
Statement 6	4.00
Statement 7	4.00
Statement 8	4.50
Statement 9	4.20

5.4.2. Qualitative analysis derived from interviews

The data from each of the 10 interviews was carefully collected and transcribed verbatim. Afterwards, they were uploaded to a qualitative data analysis software (ATLAS.ti 9). After transcription, the data were cleaned, labeled/coded, in which we followed a flexible coding approach (Deterding & Waters, 2018). The moment each interview was coded, the three coders selected and discussed and followed the following steps (see Appendix D):

- First, our codes (n=42) were divided based on their overarching subject. Five subjects that illustrated our explorative research were chosen: Unpredictability, Believability and Intuitiveness, Motivation, Immersive factors
- Second, within these three subjects, thematic analysis was performed together, coming up with sub themes covering the overarching ideas between a large number of codes (V. Braun, 2006).
- Third, we allowed sub themes to be connected between the different subjects, which finally resulted in 6 final themes or factors that allowed us to answer our problem statement and support the feasibility study.

5.5. Results

5.5.1. Interview Results

During the one-on-one structured interview participants expressed their own experiences, providing the quality of Adaptive AI communication training and their views on feasibility in the future trainings at KLM to discover how Adaptive AI in the virtual environment could impact the factor of unpredictability during sales calls or in the future training kits in for pilots at KLM.

We finalized 6 final themes based on quotation and code clustering, as the frequency of the labels were always on a higher level in the audio transcripts thus providing a rich data. The final factors which attest to the feasibility factors of the proof of concept are as follows.

1. **Experience of Unpredictability.**

   Participants in the discussion mentioned the factors of Unpredictability in the communication training with Adaptive AI like the conversation can change direction in ways they have never experienced before along with the reactions as P[1] states *"when you say like you can predict that what he's going to say next or how is the conversation going to I can't predict that Okay, I Have no idea what he's gonna react"*.

Furthermore participants felt the fun of talking to an AI Communication training bot and it responds the way we way to but there is also a sense of excitement where the participant felt they had control as there was no script involved but were given new conversation scenarios from the AI as P[8] mentions *"it's fun because you can ask the bot anything you want and you can again answer back. And you have also control, which is, and again, going back to the unpredictability, you can anticipate any way you want, then you're not restricted by a script"*.

2. **Believability.**

The participants shared their thoughts on the believability factor, emphasizing the extremely acceptable nature of the behavioral response in AI training, where interactions are similar to human-to-human exchanges as P[5] feels *"I think is very nice is answering in time, but he has behaves pretty human to say it like that, which I think was very nice"*.

An interesting finding was that a few participants felt that if the prompt was not provided to the participant during the sensitizing session and the mention of AI communication bot was removed, it would have built more excitement as realistic as the participant wouldn't have known whether it was a real person or AI, thus making the entire interaction very much believable. as P[1] states *"so if you should test this for instance by oh I've got a person on the line could you maybe try this call and not saying that it is a training bot how differently the conversation would be"*.

3. **Natural Conversation.**

Participants selected natural discussions in which they felt the replies were natural, and the feasibility factors increased because there was no actor involved, implying that there were no staged chats, making it more realistic as stated by P[10] *"Yeah, they could because they could really answer true and they could ask more. And it was more feasible because they didn't have to feel that there was an actor. They had to feel eventually that they were talking to something."*

There was an important finding where participants tried to play around with the provided prompt in the conversation, where participants tried to stick to the script or the goal of the conversation, and in between when the conversation was side-tracked, the AI conversation bot gave a natural answer which was highlighted. P[3] *"Even though you kind of you stayed within his what he wanted, but you asked him something about he just gave a natural answer."*

4. **Motivation.**

Few participants mentioned that motivation is increasing because the sales representative or anyone else who is undergoing this AI-based communication training is not developing communication skills in a scripted format, but rather an unscripted and unpredictable communication done in a natural sense by the tool. As highlighted by P[7] *"Yeah, there's just no script. So it becomes fully unpredictable and that's a motivating factor to look upto in a conversation."*

Also, participants believe that this can have a direct impact on staff competency

enhancement because it is unscripted and not a trick. P[10] states *"Yeah, I think it was very predictable, very intuitive actually. So people could learn a trick instead of learning a skill."*

5. **Intuitiveness in Interaction.**

There was a mixed reaction in this section, which exposes present feasibility factors and may give rise to suggestions to improve sensitization. Participants believed that intuitiveness was a little too strong in the instance where they knew the prompts and target to be completed with the AI communication bot, which caused them to approach with a direction, as stated in P[4]. *"I can say that I was giving the goals of Rajeeb. I was giving his goals beforehand so I knew a little bit about what he was going to tell me so that felt a little bit too intuitive for me because I knew what was going to happen."*

Few Participants found the instruction was incredibly helpful and intuitive, and the participant appreciated the flexibility it provided. Even if conversations did not go as planned, they noticed that the flexibility to instantly pivot and explore different topics made the experience adaptive and entertaining as P[7] states *"very intuitive because I could just Even if it didn't went the way I wanted it or that I expected you could always sidetrack and be like oh, okay Let's ask something else."*

6. **Quality of response.**

The training was generally accepted by participants, however some believed it might be improved. They specifically stated that an initial delay between when they finished talking and when the AI bot answered made the conversation feel rather robotic. Some participants recommended that changing the words used at the start of the conversation to reflect the context could improve the interaction's overall human-like quality as P[1] states *"was a realistic conversation but still in my mind I had I'm talking to a bot. Okay, so that was a bit that the combination of realistic and bots."*

Participants expressed optimism about the technology's possibilities, emphasizing its unique and natural way of facilitating talks. They saw great potential in these technologies, emphasizing their potential for transformative effects in communication and interaction. P[2] makes a note of *"Yeah, I chose good potential because I think this is a really interesting way of like having these conversations and it did feel pretty natural, so I do think there's a there's a like a big potential in these kinds of technologies."*

6 Conclusion

In the dynamic airline customer service sector, KLM is pioneering an innovative approach to improving the skills and performance of their sales representatives through the strategic integration of adaptive AI to enhance customer relationships. This Innovative initiative is a revolution in a way to train sales reps to meet future unpredictable challenges.

Our journey demonstrates the spirit of innovation behind KLM's adoption of new technologies to improve the skills of its sales force and customer service capabilities powered by adaptive AI to create the feasibility of training experiences. In summary, the software under discussion embodies a sophisticated AI-driven voice communication system. It features a well-organized architecture designed to deliver a seamless and human-like user interaction. The system's components, including the Voice Interface, Speech Recognizer, Conversation Manager, Response Generator, and the Advanced Audio Processor – synchronize perfectly to produce a fluid and intuitive user interaction.

At the heart of our training is adaptive AI, which personalizes the learning experience for each salesperson. The system's workflow, from the initiation of a call to the delivery of responses, is thoughtfully designed to ensure smooth and efficient conversations. Core functions, such as speech recognition, context analysis, and realistic text-to-speech conversion, collectively contribute to the system's capability to interpret and respond to user input accurately. Inspired by the studies of the literature review, which highlight the individuality of learning rates, adaptive AI will personalize training modules. It automatically adjusts the difficulty and complexity of interactions faced by a trainee, based on their progress and skills, ensuring optimal learning. Initially, sales staff will use the AI-based system during phone conversations. This adaptive AI does double duty, guiding sales reps through complex sales scenarios, while also analysing customer feedback to refine conversational techniques. This immediate feedback loop is designed to build confidence and improve performance in real-time.

Our initiative recognizes the important aspect of training in sales. This software's inventive approach serves as a bridge between AI and business, harnessing cutting-edge technologies to improve customer satisfaction and optimize operational processes. The comprehensive functional description provided in this document serves as a guiding framework for the software's development, ensuring precise alignment with stakeholder needs and functional requirements. The system's overarching goal is to create a user-friendly, natural, and effective voice communication platform adaptable to diverse business scenarios, ultimately enhancing the overall user experience, it will be designed to help sales representatives read and respond to different challenges faced during customer interactions and improve their empathy and efficiency. Also, considering the safety and ethical impact that the training can have, all training will take place in a support environment ensuring a safe and comfortable learning environment.

KLM's vision is to combine adaptive AI and virtual reality into a seamless training platform that will not only enhance the capabilities of the sales representatives but also pave the way for more complex interactions with clients. By harnessing the intelligence of AI, we are poised to redefine the standards of customer service and business efficiency. As

we perfect these technologies, we envision a future where every customer interaction is enhanced by unprecedented insights, expertise, and care, marking a new era for the growth of KLM's innovation strategy.

6.1. Recommendations and Future Implementations

The product described in this report shows certainty in its concept. However, it is still unrefined at this stage. The small delay in conversation must be reduced, along with uniformity in tone of voice. However to make it truly indistinguishable from a human, there are a few more areas of improvement and are as follows:

- **Emotional Intelligence:** Developing AI chatbots that can recognize and respond to users' emotions can make interactions more empathetic and personalized.
- **Enhanced Conversation Flow:** Improving the flow of conversations by making chatbots more skilled at maintaining context over long interactions and switching between topics seamlessly.
- **Increased Learning and Adaptation:** Giving chatbots the ability to learn from each conversation and adapt over time, allowing them to provide more accurate and relevant information.
- **Multi-Turn Conversations:** Enabling chatbots to engage in longer and more complex multi-turn conversations, including handling interruptions and resuming conversations gracefully.

Once the aforementioned conditions are met, this Adaptive AI bot can be embedded into a 3D tablet that can display a 3D image of an avatar that will serve as an assistant. As a result, the chatbot can be represented by an animation powered by the first phase's Adaptive AI model. This adds to the immersion by allowing you to view the assistant at the desk, who is powered by KLM's in-house digital twin and responds with unpredictability and a genuine flow of conversation. It can be configured to function as a personal assistant, easing corporate workflow.

This type of technology can also be used by KLM's ground crew and technicians, who rely on manuals for maintenance and repair procedures both throughout the manufacturing process and on airplanes. In the future, this application could be enhanced by combining Virtual reality (VR) to provide an immersive maintenance experience. This immersive method would improve their problem-solving efficiency, ultimately boosting their proficiency in aviation maintenance.

This can be integrated with VR, allowing KLM developers to create smart and visually engaging training programs. KLM already possesses cutting-edge technology in place, known as "In-World," where avatars interact with respective end users, such as flight attendants during training. These avatars can be enhanced by incorporating the current proof of concept of Adaptive AI, resulting in training based on different personas, fostering unpredictability in conversations. Additionally, they can be customized to offer region-specific training, promoting cultural awareness and empathy among stewardesses who frequently fly to different countries. This, in turn, boosts their confidence and motivation in their daily work.

7 Team Reflection

Over the past 10 weeks, our JIP project has been a vibrant journey of embracing new insights and phasing out old approaches, particularly in KLM's innovative venture into adaptive AI and VR to improve training for sales teams and enhance customer conversations.

We take great pride in the dynamic and harmonious interplay of our team's diverse educational and professional backgrounds, which fostered a spirit of collective enthusiasm and cohesiveness. Every member's input was valued and considered, helping us navigate through periods of feedback and various obstacles. This shared respect was the cornerstone for maintaining transparent communication, deepening team bonds, encouraging adaptability, and upholding a strong work ethic. With no rigid roles, our team members stepped up, stretched their potential, and propelled the project forward steadily. The 'Buddy Check' feedback process was instrumental in aligning our individual and collective expectations with the project's ambitious goals.

The project commenced with sharp, well-defined targets. We express our sincere thanks to KLM's coach and the managerial team for simplifying the integration of AI and VR with customer interaction into a manageable and strategic project outline. Conscious of our limited timeline, we organized our efforts strategically. The rhythm of our meetings, particularly intensified during the experimental stage, was essential in sharing ideas and crafting a unified path forward. Support from KLM's leadership and the unexpectedly smooth internal navigation through bureaucratic processes were key in starting our experiments on schedule. The solid background research and previous technological applications instilled confidence in our participants, though we did encountered some unforeseen conflicts towards the end. We managed to overcome it through effective communication.

For some, this project was their first professional foray. Nonetheless, the team as a whole remained dedicated to delivering results with precision and scientific rigor. We are indebted to our company coach for the opportunity to present our results to KLM's decision-makers. We also owe a debt of gratitude to the JIP Core Team for their all-encompassing support and the networking opportunities that have enriched our JIP journey immeasurably. In conclusion, diving into the pioneering realm of AI and VR at KLM has been an extraordinary chapter of growth and groundbreaking work for our entire team.

Bibliography

A. Bans-Akutey, B. T. (2021). Triangulation in research. https://doi.org/3392.10.20935/AL33922

Abate, A. F., Guida, M., Leoncini, P., Nappi, M., & Ricciardi, S. (2009). A haptic-based approach to virtual training for the aerospace industry. ***Journal of Visual Languages and Computing***, ***20***(5), 318–325. https://doi.org/10.1016/j.jvlc.2009.07.003

Agee, J. (2009). Developing qualitative research questions: A reflective process. ***International Journal of Qualitative Studies in Education***, ***22***(4), 431–447.

Ahir, K., Govani, K., Gajera, R., & Shah, M. (2020). Application of virtual reality for enhanced education learning, military training, and sports. ***Augmented Human Research***, ***5***(1), 1–9. https://doi.org/10.1007/s41133-019-0025-2

Analytical study on use of AI techniques in tourism sector for smarter customer experience management. (2020, March).

Bellamy, W. (2017). ***9 companies using augmented reality and virtual reality in aviation***. http://www.aviationtoday.com/2017/08/24/9-companies-using-augmented-virtual-reality-aviation/

Calonge, D. S., Smail, L., & Kamalov, F. (2023). Enough of the chit-chat: A comparative analysis of four ai chatbots for calculus and statistics. ***Journal of Applied Learning and Teaching***, ***6***(2).

Charmaz, K. (2002). ***Qualitative interviewing and grounded theory analysis. j. f. gubrium & j.a.holstein (eds.), handbook of interview research: Context & method***. Thousand Oaks, CA: Sage.

Christodoulou, A., & Cullinane, K. (2019). Identifying the main opportunities and challenges from the implementation of a port energy management system: A swot/pestle analysis. ***Sustainability***, ***11***(21), 6046. https://doi.org/10.3390/su11216046

Cleophas, T. J., & Zwinderman, A. H. (2016). Data spread: Standard deviations, one sample z-test, one sample binomial test. In ***Clinical data analysis on a pocket calculator: Understanding the scientific methods of statistical reasoning and hypothesis testing*** (pp. 201–205). Springer International Publishing. https://doi.org/10.1007/978-3-319-27104-0_35

Cochrane, T., Cook, S., Aiello, S., Christie, D., Sinfield, D., Steagall, M., & Aguayo, C. (2017). A dbr framework for designing mobile virtual reality learning environments. ***Australasian Journal of Educational Technology***, ***33***(2), 54–68. https://doi.org/10.14742/ajet.3613

Deterding, N., & Waters, M. (2018). Flexible coding of in-depth interviews: A twenty-first-century approach. ***Sociological Methods & Research***, ***50***(2), 708–739. https://doi.org/10.1177/0049124118799377

DNÇER, N. (2023). Elevating aviation education: A comprehensive examination of technology's role in modern flight training. ***Journal of Aviation.***, ***7***. https://doi.org/10.30518/jav.1279718

EASA. (2023). ***Easa artificial intelligence roadmap 2.0 - a human-centric approach to ai in aviation.*** https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

Fernandes, T., & Oliveira, E. (2021, January). Understanding consumers' acceptance of automated technologies in service encounters: Drivers of digital voice assistants adoption.

Frigo, E., Mauricio & Silva, & Barbosa, G. (2016). Augmented reality in aerospace manufacturing: A review. ***Journal of Industrial and Intelligent Information.*** https://doi.org/10.18178/jiii.4.2.125-130

Gong, L., Berglund, J., Fast-Berglund, A., Johansson, Wang, Z., & Börjesson, T. (2019). Development of virtual reality support for factory layout planning. ***International Journal on Interactive Design and Manufacturing (IJIDeM)***, ***13***(3), 935–945. https://doi.org/10.1007/s12008-019-00538-x

Guo, Z., Zhou, D., Zhou, Q., Zhang, X., Geng, J., Zeng, S., Lv, C., & Hao, A. (2020). Applications of virtual reality in maintenance during the industrial product lifecycle: A systematic review. ***Journal of Manufacturing Systems***, ***56***, 525–538. https://doi.org/10.1016/j.jmsy.2020.07.007

Hadi, M. U., Qureshi, R., Shah, A., Irfan, M., Zafar, A., Shaikh, M. B., & Mirjalili, S. (2023). Large language models: A comprehensive survey of its applications, challenges, limitations, and future prospects.

Harmon, L. J., & Losos, J. B. (2005). THE EFFECT OF INTRASPECIFIC SAMPLE SIZE ON TYPE I AND TYPE II ERROR RATES IN COMPARATIVE STUDIES. ***Evolution***, ***59***(12), 2705–2710. https://doi.org/10.1111/j.0014-3820.2005.tb00981.x

J. C. Brendan, E. S., & Mary, R. R. (2019). Head-mounted display virtual reality in post-secondary education and skill training. ***Educ.Inform. Technol.*** https://doi.org/https://doi.org/10.3389/feduc.2019.00080

Klassen, M. (2021). ***How using ai for pilot training makes aviation safer | paladin ai. medium.*** https://medium.com/paladin-ai/adaptive-pilot-training-makes-aviation-safer-and-cheaper-a8ce7d67f626

Klein, A. M., Rauschenberger, M., Thomaschewski, J., & Escalona, M. J. (2021). Comparing Voice Assistant Risks and Potential with Technology-Based Users: A Study from Germany and Spain. ***Journal of Web Engineering***.

KLM. (2022). Klm annual report 2022. https://view.publitas.com/cfreport/klm-annual-report-2022/page/1

Laovoravit, V., Manon, W., Inchamnan, W., & Swangarom, P. (2023). The learning achievement of airline business students on virtual experience during the pre-flight emergency equipment check procedures on the airbus a320 series aircraft. ***Asia Social Issues***, ***16***(4). https://doi.org/10.48048/asi.2023.255791

Maloney, D., Rajasabeson, S., Moore, A., Caldwell, J., Archer, J., & Robb, A. (2019). Ethical concerns of the use of virtual avatars in consumer entertainment. ***2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)***, 1489–1492. https://doi.org/10.1109/VR.2019.8797926

Nielsen, J., & Landauer, T. (1993). A mathematical model of the finding of usability problems. ***Proceedings of the Interact '93 and CHI '93 Conference on Human Factors in Computing Systems***, 206–213. https://doi.org/10.1145/169059.169166

Osbourne, E., & Arora, A. S. (2022, January). Artificial Intelligence (AI) in Marketing: How AI Supports Marketers Throughout the Consumer Journey.

Owen, R., Stilgoe, J., Macnaghten, P., Gorman, M., Fisher, E., & Guston, D. (2013). A Framework for Responsible Innovation. ***Responsible Innovation***, 27–50.

Pande, A., Shrestha, B., Rani, A., & Mishra, D. (2023). A comparative analysis of real time open-source speech recognition tools for social robots. ***International Conference on Human-Computer Interaction***, 355–365.

Pflanzer, M., Traylor, Z., Lyons, J. B., Dubljevi, V., & Nam, C. S. (2022, September). Ethics in human–AI teaming: Principles and perspectives.

Pradhan, P., Rostami, M., Kamoonpuri, J., & Chung, J. (2023). The state of augmented reality in aerospace navigation and engineering. ***IntechOpen***. https://doi.org/10.5772/intechopen.1002358

Privacy Issues in Voice Assistant Ecosystems. (2020, October).

R. Agha, A. F. (2015). The role and validity of surgical simulation. ***International Surgery***, ***100***(2). https://doi.org/https://doi.org/10.9738/intsurg-d-14-00004.1

Rietveld, T., & van Hout, R. (2017). The paired t test and beyond: Recommendations for testing the central tendencies of two paired samples in research on speech, language and hearing pathology. ***Journal of Communication Disorders***, ***69***, 44–57. https://doi.org/10.1016/j.jcomdis.2017.07.002

Rosellini, A. (2023). ***Measuring augmented reality and virtual reality trajectory in the training environment.*** https://doi.org/10.1007/978-3-031-43471-6_14

S. K. Tadeja, P. S., & Kristensson, P. O. (2019). Exploring aerospace design in virtual reality with dimension reduction. ***AIAA Scitech***. https://doi.org/https://doi.org/10.2514/6.2019-2206

Seedhouse, E., Brickhouse, A., Szathmary, K., Williams, E. D., & Seedhouse, P. E. (2020). ***Human factors in air transport: Understanding behavior and performance in aviation***. Springer International Publishing.

Shneiderman, B. (2020). Bridging the gap between ethics and practice: Guidelines for reliable, safe, and trustworthy human-centered ai systems. ***Acm Transactions on Interactive Intelligent Systems***, ***10***(4), 1–31. https://doi.org/https://doi.org/10.1145/3419764

Stahl, B. C. (2022, February). Responsible innovation ecosystems: Ethical implications of the application of the ecosystem concept to artificial intelligence.

Tacgin, Z., & Arslan, A. (2017). The perceptions of ceit postgraduate students regarding reality concepts: Augmented, virtual, mixed and mirror reality. ***Educ.Inform. Technol.***, ***22***(2), 1179–1194. https://doi.org/10.1007/s10639-016-9484-y

Twilio. (2023). Voice api capabilities [Available online: https://www.twilio.com/en-us/voice (accessed on October 30, 2023)].

U. Radhakrishnan, F. C., & Koumaditis, K. (2021). Immersive virtual reality training : Three cases from the danish industry. ***IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)***, 1–5. https://doi.org/10.1109/VRW52623.2021.00008

V. Braun, V. C. (2006). Using thematic analysis in psychology. ***Qualitative Research in Psychology***, ***3***(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Y. Zhou, T. L., & Yan, X. (2022). Virtual reality's influence on construction workers' willingness to participate in safety education and training in china. ***Journal of Management in Engineering***, ***38***(2). https://doi.org/https://doi.org/10.1061/(ASCE)ME.1943-5479.0001002

Z. Liang, K. Z., & Gao, K. (2019). Development of virtual reality serious game for underground rock-related hazards safety training. ***IEEE Access***, ***7***, 118639–118649. https://doi.org/https://doi.org/10.1109/ACCESS.2019.2934990

Zhang, W. (2014). ***On college oral english teaching in the base of virtual reality technology.*** 10.4028/www.scientific.net/AMM.687-691.2427

Zyda, M. (2005). From visual simulation to virtual reality to games. ***IEEE Computer***, ***38***(9), 25–32. https://doi.org/https://doi.org/10.1109/MC.2005.297

A Consent Form

CONSENT FORM FOR THE USE OF ADAPTIVE AI & VR TECHNOLOGY

INFORMED CONSENT FORM

Experiment: Adaptive AI & VR Integration for Enhanced Communication Training - A joint feasibility study of the application of AI & VR technology in training sales representatives.

Investigators: Students at TU Delft and KLM coaches - Joint Interdisciplinary Project (TU Delft and KLM)

Researchers: Deepak Sogasu, Ehsan Sadegh, Henk Jekel, Kshitij Sreenivasa Roopa, Ramon Mozaffarian, Ujjayan Dhar ,& Jae Maloney

Study Information: -10-2023

1. Introduction: This consent form pertains to the participation and use of adaptive Artificial Intelligence (AI) and Virtual Reality (VR) technology. The primary purpose of this initiative is to integrate AI technology with phone conversations to train sales representatives more effectively. In subsequent stages, we aim to enhance conversations using VR technology.

2. Purpose: This study, conducted in collaboration with KLM, aims to explore the potential benefits of integrating adaptive AI technology with phone conversations, focusing on training sales representatives. In the subsequent stages, the use of Virtual Reality (VR) will be harnessed to facilitate more effective and immersive conversations. Participants will be exposed to AI-enhanced phone interactions to understand the efficacy and reception of these technologies in a real-world setting. The study will involve an hour-long session, encompassing a pre-study questionnaire & AI interaction, and a post-study questionnaire. Participation is purely voluntary.

3. Benefits and Risks of Participating: Participants will experience firsthand the merger of AI technology with phone conversations for immersive communication. This is purely for training and development purposes and will be handled with the utmost confidentiality and respect for privacy.We anticipate minimal risks and discomforts associated with this technology. However, should you experience any unease or discomfort during the session, you are free to discontinue at any time.

4. Data Collection & Usage: Participants' data will be recorded and stored on secure devices, accessible only to the authorized researchers from this project, the university supervisor, and KLM coaches. This data will be stored for 10 weeks and includes questionnaire responses, observation notes. Personal identifiers will be limited, each participant will be associated with a unique ID number.

Participants can request access to or rectification of their data, transcripts, or recordings during these 10 weeks. After this period, all personal identifiers will be deleted, leaving only

the final report and any presentation material. This data will neither be sold nor shared with third parties for additional research.

The study's findings will be showcased in a report or final presentation at TU Delft and KLM, using anonymized data and charts. Any video content used for presentations will ensure participant's identities remain confidential through blurring or other methods.

During the training sessions, Phone conversations may be monitored, recorded, and analyzed by AI for training purposes. All personal identifiers will be removed, ensuring the anonymity of the participant. The data will be used exclusively for training and research and will not be shared with third parties without explicit consent.

5. Rights & Withdrawal: Participation is voluntary. You can withdraw consent and discontinue participation at any time without any adverse consequences.

6. Confidentiality: All data collected will be stored securely and used only for the purposes mentioned above. Your identity and personal details will remain confidential and will not be disclosed.

7. Queries & Concerns: If you have any questions about the study, please contact:

- Contact Name:
- Contact Email:
- Contact Phone Number:

Please tick the appropriate boxes: Taking part in the study
[ ] Yes [ ] No

By signing below, you acknowledge that you have read and understood the study information, or it has been read to me and consent to participate under the terms outlined. I have been able to ask questions about the study, and my questions have been answered to my satisfaction. I understand that I can refuse to answer questions, and I can withdraw from the study at any time, without having to give a reason.

Name of participant:

Signature:

Date: .10.2023

I have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

Student from JIP Team:

Signature:

Date: .10.2023

B Questionnarie

B.0.1. Quantitative Analysis: Survey 1- Pre-testing of Product, Survey 2 & 3 - Post-testing of Product

Dear participant,

Thank you for your cooperation in this study on the measurement of the impact of the **KLM AI Sales Employee training!** This questionnaire is part of the Joint Interdisciplinary Project (JIP), a master's course at the TU Delft, in which interdisciplinary project teams of MSc students from different faculties work together full-time for ten weeks alongside a technologically innovative company to address a business case and create innovative impact.

The aim of this research is to measure the impact of the developed AI employee training for the sales department and compare it to the existing employee training, assessing the quality of training and the learning experience for employees, considering factors like (un)predictability, believability, and intuitiveness.

Please note that this is the first of three surveys. This first survey will be conducted before testing the new AI Employee training, the second survey will be conducted right after testing the new training, and the third will be carried out at least a day after trying out the training.

Answering all questions takes about 5 minutes in total. By clicking further, you confirm that your participation is completely voluntary, and that you understand that you can withdraw from participation at any time without having to provide a reason. All your answers and data are processed completely anonymously. **Click the blue button below to start the survey!**

Figure B.1: Description of Survey 1

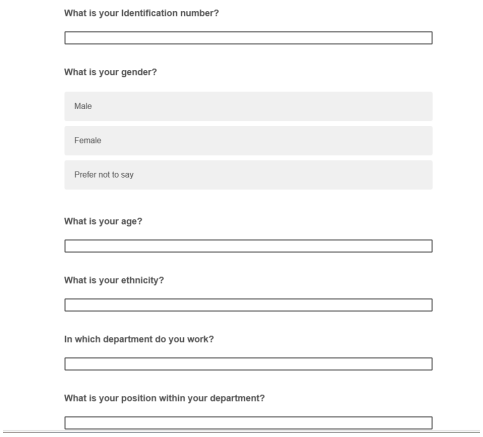
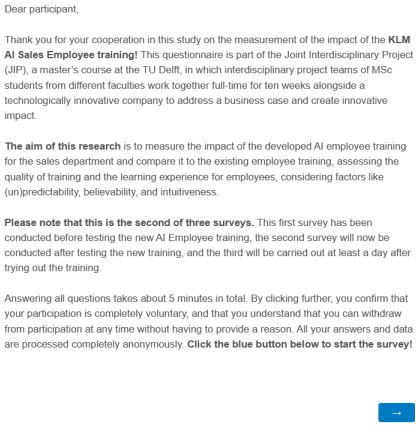


Figure B.2: Step1 of Survey1

Figure B.3: Step 2 of Survey 1

Figure B.4: Step 3 of Survey 1

Figure B.5: Step 4 of Survey 1


Dear participant,

Thank you for your cooperation in this study on the measurement of the impact of the **KLM AI Sales Employee training**! This questionnaire is part of the Joint Interdisciplinary Project (JIP), a master's course at the TU Delft, in which interdisciplinary project teams of MSc students from different faculties work together full-time for ten weeks alongside a technologically innovative company to address a business case and create innovative impact.

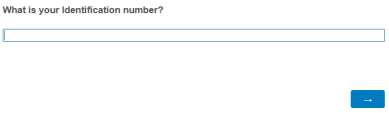
The aim of this research is to measure the impact of the developed AI employee training for the sales department and compare it to the existing employee training, assessing the quality of training and the learning experience for employees, considering factors like (un)predictability, believability, and intuitiveness.

Please note that this is the second of three surveys. This first survey has been conducted before testing the new AI Employee training, the second survey will now be conducted after testing the new training, and the third will be carried out at least a day after trying out the training.

Answering all questions takes about 5 minutes in total. By clicking further, you confirm that your participation is completely voluntary, and that you understand that you can withdraw from participation at any time without having to provide a reason. All your answers and data are processed completely anonymously. **Click the blue button below to start the survey!**

→

Figure B.6: Description of Survey 2


What is your Identification number?

→

Figure B.7: Step1 of Survey2


How would you rate the **overall quality** of the AI Sales Employee training?

Ranking AI Sales Employee training

How would you rate the **level of intuitiveness** of the training?

Not intuitive at all — Very intuitive

Intuitiveness

How would you rate the **level of unpredictability** of the training?

Very predictable — Very unpredictable

Unpredictability

How would you rate the **level of difficulty** of the training?

Very easy — Very difficult

Difficulty

How would you rate the **level of believability** of the training (Does it resemble a real-life situation)?

Not believable at all — Very believable

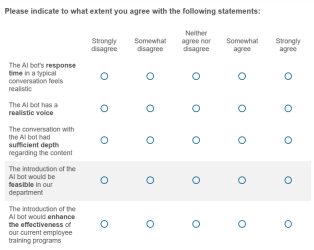
Believability

Figure B.8: Step 2 of Survey 2


Please indicate to what extent you agree with the following statements:

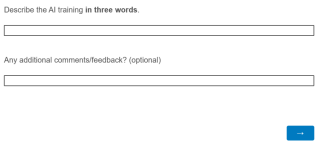
	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
The training adequately prepares me and provides enough **confidence** for communicating with sales customers.	○	○	○	○	○
The training adequately **prepares** and **motivates** me for communicating with sales customers.	○	○	○	○	○
The training incorporates **unpredictable scenarios** that simulate real-world unpredictability.	○	○	○	○	○
The training **enhances cross-cultural communication skills**, deepens understanding of cultural differences, and enables us to better meet the diverse customer needs.	○	○	○	○	○

Figure B.9: Step 3 of Survey 2


Please indicate to what extent you agree with the following statements:

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
The AI bot's **response time** in a typical conversation feels realistic	○	○	○	○	○
The AI bot has a **realistic voice**	○	○	○	○	○
The conversation with the AI bot had **sufficient depth** regarding the content	○	○	○	○	○
The introduction of the AI bot would be **feasible** in our department	○	○	○	○	○
The introduction of the AI bot would **enhance the effectiveness** of our current employee training programs	○	○	○	○	○

Figure B.10: Step 4 of Survey 2


Describe the AI training **in three words**.

Any additional comments/feedback? (optional)

→

Figure B.11: Step 5 of Survey 2

Dear participant,

Thank you for your cooperation in this study on the measurement of the impact of the **KLM AI Sales Employee training!** This questionnaire is part of the Joint Interdisciplinary Project (JIP), a master's course at the TU Delft, in which interdisciplinary project teams of MSc students from different faculties work together full-time for ten weeks alongside a technologically innovative company to address a business case and create innovative impact.

The aim of this research is to measure the impact of the developed AI employee training for the sales department and compare it to the existing employee training, assessing the quality of training and the learning experience for employees, considering factors like (un)predictability, believability, and intuitiveness.

Please note that this is the third of three surveys. This first survey has been conducted before testing the new AI Employee training, the second survey has been conducted after testing the new training, and the third will now be carried out at least a day after trying out the training.

Answering all questions takes about 5 minutes in total. By clicking further, you confirm that your participation is completely voluntary, and that you understand that you can withdraw from participation at any time without having to provide a reason. All your answers and data are processed completely anonymously. **Click the blue button below to start the survey!**

Figure B.12: Description of Survey 3

What is your Identification number?

→

Figure B.13: Step1 of Survey3

How would you rate the **overall quality** of the AI Sales Employee training?

Ranking AI Sales Employee training

How would you rate the **level of intuitiveness** of the training?

Not intuitive at all — Very intuitive

Intuitiveness

How would you rate the **level of unpredictability** of the training?

Very predictable — Very unpredictable

Unpredictability

How would you rate the **level of difficulty** of the training?

Very easy — Very difficult

Difficulty

How would you rate the **level of believability** of the training (Does it resemble a real-life situation)?

Not believable at all — Very believable

Believability

Figure B.14: Step 2 of Survey 3

Please indicate to what extent you agree with the following statements:

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
The training adequately prepares me and provides enough **confidence** for communicating with sales customers.	○	○	○	○	○
The training adequately **prepares** and **motivates** me for communicating with sales customers.	○	○	○	○	○
The training incorporates **unpredictable scenarios** that simulate real-world unpredictability.	○	○	○	○	○
The training **enhances cross-cultural communication skills**, deepens understanding of cultural differences, and enables us to better meet the diverse customer needs.	○	○	○	○	○

Figure B.15: Step 3 of Survey 3

Please indicate to what extent you agree with the following statements:

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
The AI bot's **response time** in a typical conversation feels realistic	○	○	○	○	○
The AI bot has a **realistic voice**	○	○	○	○	○
The conversation with the AI bot had **sufficient depth** regarding the content	○	○	○	○	○
The introduction of the AI bot would be **feasible** in our department	○	○	○	○	○
The introduction of the AI bot would **enhance the effectiveness** of our current employee training programs	○	○	○	○	○

Figure B.16: Step 4 of Survey 3

Has your opinion changed after some time has passed (at least 1 day) since trying out the training? If so how?

Any additional comments/feedback? (optional)

→

Figure B.17: Step 5 of Survey 3

B.1. Interview Guide

B.1.1. Demographics

- Could you share more about your role at KLM, your department and your position?

B.1.2. Quality of Training

- Past : The ratings for AI sales employee communication training may have been influenced by your experience. Can you reflect on how your experience impacted your perception of the training quality?
- Present : As per the AI sales employee communication training, what features, aspects or experiences influenced your overall quality rating ?

B.1.3. Intuitiveness of Training

- Past : What specific elements make the current communication training intuitive or less intuitive as your past experience?
- Present : What elements of the AI employee communication training contribute to the level of intuitiveness and do you think it should be improved ? If yes then how?

B.1.4. Unpredictability of Training

- When you talked about the unpredictability of the AI training, could you share specific instances or situations where this unpredictability was evident?

B.1.5. Believably of Training

- What aspects of the AI communication training make it believable or non-believable when compared to real life situations ?
- The factors of AI communication like response time, realism, and depth of conversation. Can you provide examples or describe situations where you felt these factors were notable?

B.1.6. Feasibility of Training

- When you expressed your views on the AI bot, especially in terms of its feasibility within your department, can you explain what factors you considered?

B.1.7. Elaboration of the words and emotions behind it

- You were asked to describe current employee training in three words. Could you explain why you chose those words and the emotions that drive in expressing those words?

C Experiment locations

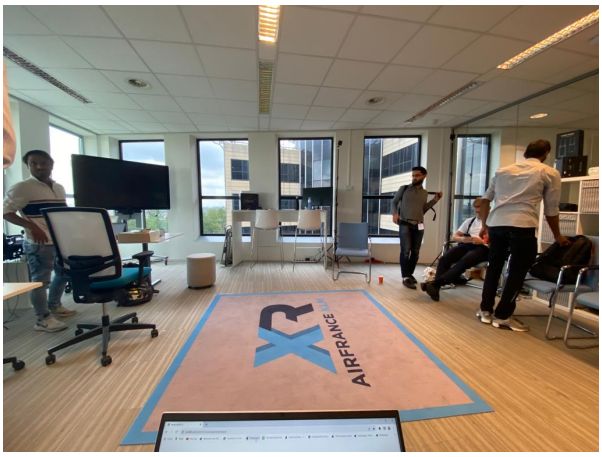


Figure C.1: XR lab

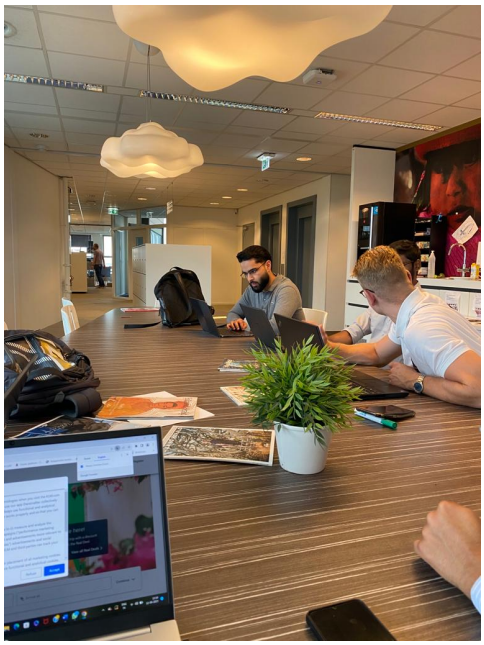


Figure C.2: Team Discussion & Collaboration

Figure C.3: Interactive Session with Participants

Figure C.4: Experimenting with VR

Figure C.5: VR Visuals

C.1. Interview transcripts

C.1.1. Participant - 1

Ujjayan : So, hello everyone. So, this is a part of a qualitative research for JIP KLM project 1 .4 .1 and I will be interviewing Participant-1 who is our participant number two who has gone through a sensitizing situation of going through the product demonstration and this interview will help us to know more about the feasibility of the product. So, welcome Participant-1. Hello. So, recently you had a demonstration of the product that we built over the past few weeks but before we jump directly into that could you share more about your role at KLM, your department and your position?

Participant-1 : Yeah, I'm working at KLM at XRCOE lab as you can call it and yeah my role is the lead developer of this team and we develop all kinds of VR training of XR trainings throughout the entire company from virtual cockpit to fire safety trainings all in VR.

Ujjayan: Yeah, what else can I say? That's really nice because you say about different types of genres of training. Yes. Especially to build upon the airline industry but since now you recently had a demo can you tell me the quality of demo that you took the quality of training that is being reflected in the demo how your past experience helped in the perception of the training quality that you just experienced?

Participant-1 : Yeah, what is very interesting is that I just had a phone call with a person at least it felt like that. There is an actual character a person you talk to and it feels very natural for to talk and very immersive to talk through a phone. And that was something that I didn't experience earlier before in this type of trainings because normally it was always those chat kind of stuff. I don't have that much experience with trainings with actors within KLM.

Mainly the trainings that I see or do is just following steps and procedures for a cockpit, so not really with a character. So for me this was also quite of an introduction to that. But what I noticed is it was quite very actually interesting to see that the audience was very much in the way Yeah, it's the believability that you talk to a person that is really good.

Ujjayan : Okay, that's nice. And when you say believability and these other things, it comes down to the question that I would ask the next, which is, in this employee communication training, were there any specific features or experience or aspects that you felt or that influenced your overall experience when you gave the rating under quality?

Participant-1 : No, I think the idea that you actually have to take a phone and that you have to call, that already adds so much to the immersion. I like the idea that you can respond in any way that you would like. So you can say whatever you want to say. That makes it also very immersive, you could say, because you are not strictly in the bounds. Like you often have games or anything like that, that you have to stay within these boundaries. But you can just go talk about any subject you want and it will divert back to the training. But it gives some form of freedom. It's really nice when you say the word bounce and then you also say the word freedom at the end. I was trying this to talk about a weird subject. Like when was the first time you actually flew an airplane? It was more or less testing how far can we go. But he actually deferred it in a quite naturally way back to the training. I think that is very powerful because for you it feels some sort of a

freedom. But you are still in a balance. for training that you can measure. But when you say freedom, one word that comes to my mind is unpredictability in the training.

Ujjayan: Yeah. Yeah. So when you talked about when you went through the demo and during the survey, what when you talked about the unpredictability of AI in the training, can you share a few instances that you felt unpredictability in the call that you felt?

Participant-1 : It was very evident. Yeah, More you, talk about that. Was it completely unpredictable? That maybe not. Because I got of course the I saw the description of the this call and you see that the three topics that he's in that's important for him and that is something he immediately starts with. And if you have, yeah. finished one I said for instance he wanted to have tickets below 500 dollar and I said okay I can give you for 400 dollar and he said okay check up next item and then Going to the next up it is that that yeah There was no discussion anymore or nothing special about that that first step Okay, so in that way it is Very predictable of what he's going to do next. Yeah And Yeah, if I think more big now also everyone I said from Talking about his youth or something you don't want to talk about of course But when did you fly for the first time or what kind of lunch did you had that kind of questions? It is not that he's going to answer those Things okay, and that makes it. Yeah, is that unpredictable or is it just stick to the Target, you know it's picking the boxes. Yeah But when you say like you can predict that what he's going to say next or how is the conversation going to I can't predict that Okay, I Have no idea what he's gonna react. Okay. Yeah, but that does make the training Believable or non -believable compared to the real -life discussions that you have Yeah, I think believable. Okay. Yeah, I Mean, I think if you have this same type of conversation maybe within With an actor yeah, I can also ask these questions and then they probably also gonna say from okay Let's go back to the main topic. So in that way, yeah Okay, it's I think quite But was there any factor like when you say believable like was the time pause the depth of conversation Realism I think I'm too biased maybe because I know these kind of bots that you can talk to. What I noticed now was that the response time, it is quite long, it was I think five or six seconds before he actually started to respond, that should be shorter. I think that already helps a lot if that is like two or three seconds to make it more believable. I noticed also that I like that he starts, the response with actually of hmm and that kind of things, but sometimes I was feeling that it doesn't fit with but when you say like you had experiences with different bots like this given your past experiences,

Ujjayan : what do you think like after today's experience and the past experiences you had the feasibility within the department when you say what factors you considered when you gave the rating for the feasibility?

Participant-1 : That you can... I think it is a very believable conversation. You can actually talk in a correct way and you can talk in a way for the training and I think it is very great to have that you can easily change it. I mean, within no problem you can change the training and that's what I think is a very powerful thing or you can change it into a different person or you can change it into a different scenario. I think that is a very very powerful... yeah, very powerful for this type of application. Okay, but at the end, so this is like the end of the interview conversation and

Ujjayan : we normally end the interview with one last question which you mentioned like the three important words that you felt in the entire thing which was realistic conver-

sation but can you say like why can you dive a bit?

Participant-1 :No, I type that because it is... what were the emotions? It was a realistic conversation but still in my mind I had I'm talking to a bolt. Okay, so that was a bit that the combination of realistic and bolts. So there was a thin line which you can... it was a thin line but I'm afraid that is mostly because I'm biased because I know I'm gonna talk to an AI that is what you guys said so then I'm gonna think okay I'm talking to a computer now. Yeah, so if you should test this for instance by oh I've got a person on the line could you maybe try this seal and not saying that it is a training bolt how differently the conversation would be. You understand what I mean?

Ujjayan : Yeah, I do get it. Okay, thanks for the insights Participant-1 and thanks for being a participant. Yeah, and a few more questions. I'm just right over there. Okay.

C.1.2. Participant - 2

Ujjayan : Hello, this is JIP KLM 1 .4 .1 and I am Ujjayan conducting the qualitative interview. We have participant Jordi. Welcome Jordi. Thank you. Yes, and Jordi has recently gone through communication training with our newly developed proof of concept and also provided the surveys. So, if you can share me more about your role at KLM, your department and your position.

Participant-2 :Right now I am an intern at KLM and I am currently working on experimenting with Gaussian splitting, implementing that in virtual reality, seeing if that can improve the immersion of the user.

Ujjayan : Okay, okay. That's nice, but just now when you had the training in the past experiences before you went through the proof of concept. Can you tell me how your experience impacted this current perception of the training quality that you just went through? Like your past experiences with the AI training bots and the communication bots impacted today to give you the rating of the current training that you went through. I haven't done much of the AI training yet from the existing one.

Participant-2 : But the product you just showed me did have some very good potential. Okay, okay. And what do you think about the quality? The voice was very nice. Okay. Only the delay was maybe a little too long. So when I stopped speaking, the delay before you started speaking, it might be a little too long still. And there was like this um, he said. And the tone of the um did not really fit the text he was saying afterwards. Okay, okay. But the um is a nice addition to like fill up that space. Okay, and so when you say um and the fear different tone and the depth of the voice

Ujjayan : Maybe you sound you want to say is like features in the training that so do you think that in this kind of communication training what features? Experiences they impact your overall quality. What are the other features that make it a good quality training or a bad quality training?

Participant-2 : I think Like repeating things we just talked about does help a lot So if I say our price must be something him like repeating that Does help a lot in that continuing

conversation and keeping the context So that does help a lot when you say repeating in a conversation.

Ujjayan : Why is it of importance? that's me might be more of a personal thing.

Participant-2 : Yeah, just keeps the context stronger and I'll explain it as well but just to Makes you feel that the person is listening. Yeah, exactly. Yeah, that's a nice way of saying it. Okay, so just Sometimes repeating what I said, maybe so we did listen to me seeing this and you understand that

Ujjayan : Okay, and based on your past experience and today's experience of the demo Do you think that there are specific elements in communication that makes the training? Less intuitive or more intuitive?

Participant-2 : Yeah, it's like a little less because that timing I talked about um I can't really think or something. Okay. Okay. Okay, but uh, do you think how This can be improved any other file any other room for development to improve the uh, intuitiveness um As you mentioned the time lag. Yeah, that's fine. Yeah, um Maybe adding instead of like like ah, yes or something more more than that instead of just More of a verbal Like confirmation, maybe like ah, yes or stuff like that. Okay.

Ujjayan : Okay. And uh When I talk about uh, the past trainings or the normal conversation that you have Or today when you had the conversation, do you think that there was unpredictability in the training? something That you can share about specific in the situation that was evident.

Participant-2 : I did after the first uh Like there were three goals first we talked about the price and then I skipped the second one and I started talking about the third one But then it just went to the second one again. So it didn't and go, yeah, like it went in order, but not the order I chose kind of. Okay. So it really wanted to go to step two instead of talk about step three first and then step two. So that was an unpredictable step for you. I told like, if I talk about this, then we're gonna talk about this. So the order, in my opinion, shouldn't matter, but he really wanted to go in order to finish the conversation.

Ujjayan : But in a communication training, do you think that's a good thing?
Participant-2 : Um, I think it would be nice to have it in like a specific order. Okay. Yeah. Okay. Which is that something that the Bulls they didn't go in order. But it's... Okay.

Ujjayan :But when you say unpredictable, do you think how much can it be more believable or non -believable?

Participant-2 : Um... I don't know. Okay.

Ujjayan : Any points on that you think it can be made more feasible or any factors that you think that it can be made more feasibility?

Participant-2 : No, not really, no.

Ujjayan : Any words based on which you chose the three words, any decision that you

made on choosing the three words on your final feedback?

Participant-2 : Yeah, I chose good potential because I think this is a really interesting way of like having these conversations and it did feel pretty natural, so I do think there's a there's a like a big potential in these kinds of technologies. Okay, okay. When you say natural it's more about believability. Yeah, that was pretty believable as well. Yeah, it still feels a little robotic mainly I think because of that initial delay when you stop talking and before they start talking. Okay, but in general it is pretty believable.

Ujjayan : Yeah, Thanks a lot Participant-2 and these insights would be really helpful for us.

C.1.3. Participant - 3

Ujjayan : Hello, this is Ujjayan , part of JIP, KLM project 1 .4 .1 and I will be interviewing participant for who is Participant-3 and he has gone through a proof of concept demonstration of the communication training and we will build upon the questions and take his insights in the upcoming questions.

Ujjayan : So hello Participant-3, could you share more about your role at the KLM, your department and your position?

Participant-3 : I'm an intern here at XR Lab, I work here as now for the yeah, just as a developer for the and my yeah my internship is now making also communication training so but then it's for pilot in the cockpit or in front of a cabin so it's a different kind of communication.

Ujjayan : it's okay okay glad and nice that you mentioned about communication so prior this proof of concept demonstration that you just had now all the communication training that you had can you reflect like how your past experiences impacted to give a rating of today in the quality section Participant-3 : Yeah, so the thing is, any features or aspects that affected or influenced today's quality of the training? Well, I don't know how to put it, since the features are basically that you as a person call this, in this case just the normal phone number, as sales.

Ujjayan : And in that aspect, in that way it really feels kind of that you are talking to a real person. Few things that makes it less intuitive or more intuitive.

Participant-3 : Well, it's really intuitive that it's just a call on your phone . Okay, you don't have to open application or Something else it's just type in the number or Call contact if it's in your contacts already, but it's just like it's normal you would call someone Yeah, so and that so it feels it gives already a kind of natural way of Yeah, communicating to someone like that.
Ujjayan : Are there any elements you think that if I did can make it more intuitive this

Participant-3 : You can say video calling Since I think most kind of meetings I'm not in sales So most of the things we do go via video calling in teams. So yeah, maybe video calling can be done or Yeah, I Think that's one of the things that Features that could be yeah, because when you say about even like an audio calling or video calling It you want to make the entire ecosystem more believable in that sense.

Ujjayan : So do you think that today's communication training that you went through was it believable or non believable if you compare it to real life communication situation.

Participant-3 : It was pretty believable for me at least . I could See that That things that I said and that that the bottom could really offer normal person could say the same thing in all cases. Except for one minor error that it just said the pre-recorded me wrong pre -recorded message for the past that was in there And that the bosses after the pre -recorded message was a bit long But for the most part that he what he said was kind of was believable Okay, okay,

Ujjayan: But was there any situation that you felt okay this conversation makes it so real and it's very realism was there any communication that you had with the bot assistant that when you asked the bot this thing and it made you feel oh this is very real?

Participant-3 : I was kind of going in with what he said but I asked him what he meant with his Indian heritage, what he meant by it, what KLM should provide. He gave a natural answer like he meant that he meant the food that there is the food that we eat is that available on a flight and that we get the same respect as the other customers that are if there are any customers that we could just get don't get... how do you say that in English? That we get the same level of surface than other people so that we don't we're not as... behind or that he explained that really well. So it's a very Equalitarian service in terms of different cultures. He was even though yeah, I got Even though you kind of you stayed within his what he wanted, but you asked him something about he just gave a natural answer

Ujjayan : when you say natural answer at the very start of the Entire proof of concept you are given with a prompt Like you have to do this and this is your target that you want to achieve If you were not aware of the prompt Would that have improved the believability in the product if you were just given out in the blue

Participant-3 : I Think it may be held, but since I got it I don't know it really since I got the prompt and kind of the So say scared like what he wanted So it's maybe for salesperson where to know what he wanted before you call in Sometimes maybe you know because you called him earlier and this just your second call But I think it could maybe yeah for other tests it may be to do not hit but okay, but when you say like If you would have got the prompt or not got the prompt the salesperson would have benefited if you're given with the direct To -do list kind of thing.

Ujjayan : Do you think that the entire then the word that comes to my mind is unpredictability? When you shared when you spoke about unpredictability in the ratings that you provide Can you share a specific situation that made you feel like all the conversation is really unpredictable?

Participant-3 : well since well It kind of I kind of hurt the schedule, but it's one I knew what he wanted But I didn't know how he would start it and He just took a pick up the phone to say just was hello. How are you and I just answered back with that And then he just kind of went in a different he said what he wanted in a different kind of way than

I expected. Like he didn't say from, I just, he kind of told a backstory before he said, I wanted this, but what was in this gift? I wanted ticket for that price. And that's what he kind of, what I heard. Okay. So that changed the direction. Okay. Okay. It wasn't directly far. He went to the ticket price, but he said the story why he wanted it and what are you more, the reasons behind it. And I, Okay. Not the full reason. Okay.

Ujjayan : In the case of this proof of concept, what do you think is the feasibility when it comes down to the entire department that you work in? Do you think any factors that should be considered to improve the feasibility?

Participant-3 : I think it could be used like if there is, Even the communication training I am in, or I'm working on, it could be used, but in the case that we just now talk to the cabin, I get a response automatically back from the trainer that not only the trainer responds, but the bot kind of responds in the way you talk to them. But I don't know if that's feasible since the delay is now there. It could be feasible with the further development of this program. Okay. We are in the last section of our interview and the last question that I want to ask you is, you described the training program in three words.

Ujjayan : Could you explain why you chose the words and your emotions and motivations behind it? The what? The words and your comments, your final comments. It was really believable, but slow, right? Okay. The first three words, do you mean, right?

Participant-3 : Yeah, it felt for me like I was, it was believable. Like I was kind of trying to get and deal with the person. Okay. But the response, timed by some answers, when you just took for my just net an awkward amount of long. It was not that it was like he was thinking only think about it. And because it was completely silent, maybe like, I don't know, but it's just completely silent and then for a pretty long time.

Ujjayan : So you think that latency plays a very important role?

Participant-3 : Maybe, yeah. Okay. Maybe some breathing noises, like, I don't know, but I don't know how good it picks up nowadays with breathing noises in phones. But I kind of, when I last called, it was okay. That was not too by accident yesterday. You should always hear someone breathing. Okay. If they're thinking that I don't know. To make it more realistic. Yeah. Okay. Like the waiting time doesn't feel that long, maybe because it's now completely silent.

Ujjayan : Okay. Okay. Okay. Thanks a lot for the insights. And here we come to the end of our interview.

C.1.4. Participant - 4

Ujjayan: Hello guys this is Ujjayan and we are part of JIP. KLM project is 1 .4 .1 and with me I have participant 5 and his name is Participant-4. And he has gone through just now with a demonstration of the proof of consent. Okay so welcome Participant-4. Can you tell me more about your role at KLM, your department and your position?

Participant-4 : I am a VR developer for the innovation track team at KLM. I am working as an external resource from Sogeti. So I am not officially part of KLM but I am a consultant.

Ujjayan : Okay. So today when you went through the proof of concept, the demo and based on your past experiences working with the employee communication training, can you tell how your past experiences with this technology has impacted your current perception in the domain of quality?

Participant-4 : I have never done any previous trainings at KLM. So I have no clue what they are currently like but I really like this kind of training.

Ujjayan : Okay. But on the basis of quality today, what do you think any features or aspects or any experiences that you felt in today's demo that you felt the quality is good or either bad?

Participant-4 : I do believe that the experience today felt more like a real life experience than for example watching videos or doing questionnaires in multiple choice or clicking on your screen and gaining information that way. So I think that this is a more hands -on realistic experience. I do like that as a feature. Okay.

Ujjayan : When you say realistic, you mean real life? experience, can you say like any elements that makes the communication training more intuitive or less intuitive?

Participant-4 : I can say that I was giving the goals of Rajeeb. I was giving his goals beforehand so I knew a little bit about what he was going to tell me so that felt a little bit too intuitive for me because I knew what was going to happen. I think if you would have set up this training in a way where I would have been given goals that I would have to convince Rajeeb of then that would have been more difficult and more of a training, more of a realistic training.

Ujjayan : But when you say this not giving you the prompt or not giving you the to - do list that you want to interact with Rajeev for example What brings to my like if if you were not aware of the prompt or if you're not aware of what to do Would that have improved the more believability in the product?

Participant-4 : Yeah, absolutely Okay, okay because then it would have felt like someone like I was calling someone for an actual sales call and I would have had more of a Goal myself that I was trying to convince him off so that would have been more close to reality Right now I felt like he was Listing his options and his questions and his answers from a script.

Ujjayan : Okay. Okay, so okay and When I say like Unpredictability in the training do you share any situations that you were having the conversations with that you felt like the conversation went a little bit of unpredictable That you couldn't predict what the next

Participant-4 : well, I was driving him of course and that seemed to work pretty well, so That was interesting But that also makes it more realistic I can imagine if you are a real salesperson calling someone up you can literally have a conversation about anything just to form a relationship Form a bond, so I think that's a good thing But it did crash at some point when I went too far off topic, but I don't know if that was really the cause it just happened to crash Yeah, to sell the ticket. No, he was trying to buy a ticket. Okay. I was

selling a ticket. Yeah, yeah.

Ujjayan : Okay. So in that aspect, do you think how much is the response time and the response that is coming back from Rajiv?

Participant-4 : How much is the delay, the depth of conversation where you felt these factors are pretty notable? Well, you could definitely hear that he was answering with filler words while waiting for the real answer to give. So there was a delay between him saying things like well or oof or okay, then there was a delay, and then the real answer would be given. So those two would have to intertwine more and merge more for it to be super realistic.

Ujjayan : Okay. So you feel like the tone of the filler words and the real conversation didn't match up or the latency between those?

Participant-4 : The latency. specifically. The tone was... well maybe the tone will become more of a hindrance when the latency is delayed or is removed. So now the tone difference is not as noticeable because there's a delay between the two. Okay. So maybe the tone would become noticeable but I can't say anything about that. Okay. Okay.

Ujjayan : And one of the things when you say when you experience through this communication chat bot today, can you say like what could be the feasibility factor that can be improved? What factors can you say like if incorporated into this communication training program or what can make it more feasible? Do you mean technological? Yeah. Technological or any features that you felt like?

Participant-4 : Well conversation with Rajiv for example. He seemed to stick to a pretty direct goal so he was very goal driven in my eyes which was felt a little bit inhuman. Okay. I think he should have been more open to new suggestions and then try to counter those suggestions with things like in real life it's more of a compromise thing than a specific criteria and goal that someone has so you will always move towards each other. I didn't feel like that was really the case. So there was no middle ground? There was no middle ground. Yeah. Okay. Okay. That's putting it very well.

Ujjayan : Okay. And after end of every survey you will ask like any comments or any words that you feel like three words about the entire communication training.

Participant-4 : Do you want to elaborate anything on those? No. I think I said very goal driven which is I think we handled that already.

Ujjayan : Okay. Okay. Okay. Those insights are really nice and we can incorporate them to take the communication platform ahead.

Participant-4 : Thanks a lot. Good luck guys.

C.1.5. Participant - 5

Ujjayan : Alright, welcome Participant-5. You just had a small preview of the product that we are planning to showcase. But before we move on, could you share a little bit about your role at KLM?

Participant-5 : Yeah, I'm an external developer at the XR Center of Excellence here at KLM. So I work for the company Sogeti, which is a secondary company I think it's called. So I'm hired here at KLM to work on different XR projects. I'm here now for two years as XR developer. That's very nice to hear.

Ujjayan : So have you had past experiences using AI and has those experiences you know affected the way you've rated this product today?

Participant-5 : Probably yes. I worked with some AI, not too much, mostly image generation and very small experience with ChatGPT. I did work with Inworld AI, which is a Unity package to incorporate kind of a voice AI also into your project.

Ujjayan : Yeah, but how did that affect your perception of this product?

Participant-5 : I think it affected my perception in a way that I know I'm talking to AI. So I might treat it differently and ask it unrelated questions, which I would not ask a human there. Yeah, but before you tested the product, what was the prompt that you were given You were just given information about the person you were talking to. The only thing that was given to me was that I'm a salesperson and I'm trying to sell someone a ticket. Yeah, that's it. Exactly. Okay.

Ujjayan : Were there any features aspects that influenced the overall training quality?

Participant-5 : Yeah, for me the delay time was very long. I think it was also put a little bit longer before me. And something else that caught my eye was that in the first stage you accepted a certain price, which in a later, at the end of the interview, he refused. Okay, okay. So he slightly inconsistent in how he behaved.

Ujjayan : Okay. So what elements of this AI employee communication training contribute to the level of intuitiveness and what do you think should be improved?

Participant-5 : What I think is very nice is answering in time, but he has behaves pretty human to say it like that, which I think was very nice. I also tried to Pursuade him Into strange directions Which he knew were not true. So he asked me about waste management If KLM is up to stand with waste management And I told him that KLM is the greenest company and we dump all our waste into the ocean and he disagreed with that Which I did not expect actually.

Ujjayan : So do you think that plays a little bit into the unpredictability of the bot?

Participant-5: Yeah, for sure. Okay Also in terms of believability What aspects of this AI Bot makes it very believable or not believable.

Ujjayan : I see you've mentioned about the time lag Which is being taken into account,

but apart from the time lag is that anything that? Would make it more believable than it already is I think the voice is already pretty good.

Participant-5 : Yeah, very believable How to make it more believable I think if I was not given the instruction that it was a AI bot Yeah, it might take me. I might think it's just a very slow human Or just a bad connection or a bad connection I noticed that he uses some kind of Prequels before he started sentence which had a very different intonation So if that would be maybe in the same intonation as what he would start with That would already make it very much more believable Okay, and

Ujjayan : Now this just moves into a little bit of the feasibility of this sort of training based on your experiences Can you explain in terms of feasibility the factors that you considered?

Participant-5 : Yes, I would kind of well, it's not there yet, but I would very much consider this as a training Yeah, I think it's it offers a very wide range I know how much you can influence how a character behaves, but maybe you have some kind of sliders that would make it very useful. So if you have, as a trainer, a control panel in which I could control if the customer is angry, happy, I think that would greatly increase the value of the product. Okay.

Ujjayan : And finally, you filled a survey after the product and you were asked to describe this product in three words. Could you tell me why you chose the words that you chose?

Participant-5 : Yes, I chose feels pretty human. I think it was more human than I expected it to be, even with the long delays. Yeah, which I very much agree with. I mean, it's almost their kind of, actually.

Ujjayan : Well, it comes to the end. Thank you so much for your time. Your input was very valuable. No problem.

C.1.6. Participant - 6

Deepak : Alright, I'm here now with Participant-6 and you've just been back after testing the product. But before we start with our interview, I'd like you to talk about your role at KLM.

Participant-6 : Alright, so my role at KLM, I recently started in September and I am now a XR technical specialist, which isn't a developer role, but it is a role where I am doing engineering and support. Make sure that hardware works as it's supposed to and also make sure that the VR experience that we create gets to the people that need it.

Deepak : Okay, so I'm assuming you've already had a little bit of experience with AI.

Participant-6 : I have a little bit of experience with AI.

Deepak : And how did that experience with AI help you rate our current product? How

did it influence your perception of our product? You mean the AI that's right now in the product that you showed me or in general? No, in the product. But how did your previous experience with AI influence your rating on our product?

Participant-6 : My previous experience with AI is mostly based on visual text that appears, so the normal GPT use. So to hear it first time from an actual voice is a new experience for me, which is quite exciting because I like the sound better than the whole wall of text that might appear.

Deepak : But apart from voice, were there any other features or aspects that you considered while giving the overall rating? I think the interaction, the interaction's possibility is one of them. I felt like the talk I had with the AI brought was really exciting. was quite linear, so it was really searching for specific sentences from me. So, yeah, that way I did feel like I was pushed in one direction.

Deepak : Yeah, but did you still find that human?

Participant-6 : In a way, yes. In others, it felt like I can't go to another topic at all. I don't know if that is what the purpose is, of course, but I did feel like because of this, it did keep me on the road towards saying the right thing and at some point moving forward.

Deepak : Yeah, but did you at any point feel that this conversation was too predictable in terms of unpredictability in a conversation?

Participant-6 : No, I think it felt unpredictable, because you never know what the AI was going to say.

Deepak : Okay, a little bit into intuitiveness of the training. Did it feel straightforward? Did you feel that the communication feel very intuitive?

Participant-6 : I think the communication did feel intuitive, although I didn't really know because I'm not a salesperson I didn't really know how to respond or what to say. You're a bit thrown into the deep there You can say anything and the Bot will still nudge you towards something there But yeah, because I didn't really have a goal let's say I didn't really know what to ask the AI bot Yeah, but with the best of your knowledge

Deepak : Would you say that our training was believable or yes, but other aspects that we could improve to improve the believability?

Participant-6 : I think I think it was believable. I really liked the The sounds in between as well like that made them a little bit more human as well like he was thinking But sometimes there he would say something strange and Then continue with the sentence and then it felt like oh wait, is he having a stroke or something? Well, what's happening here and that There I fell out of the immersion. Let's say What what is your take on the response time that the bot right now? It did feel a little long which made it feel a little unnatural.

Deepak : Yeah yeah but as a product as a whole How feasible do you think products like this will be ?

Participant-6 : I think if this is sharpened a little bit better, you know and I think this could be a very good way to practice still be in a safe zone and doing it while doing it and not doing the Participant-1retical part, you know? It's because there is a difference there and I think it will work and be faster to learn it this way than from just a piece of paper.

Deepak : Also, if you did not know that it was a bot you were talking to, would you have doubted that you were talking to a bot?

Participant-6 : I think with my knowledge I would be thinking that there was a bot talking to me but I don't know how someone who isn't in this environment would react to this. It's completely believable I think.

Deepak : Finally, after your use of the product you had to fill a survey and you had to describe the product in three words. I want you to tell me why you chose the words you chose.

Participant-6: Let me rethink real quickly what I wrote down. I believe I said... uh... uh... searching uh... friendly and linear those were the three words that i described it with uh... linear because it as i said earlier it felt like i was really nudged towards one thing and i couldn't uh... get off the track there before going back to his question uh... friendly because uh... it i didn't feel offended in any way uh... it felt like he was just having a nice conversation with me uh... and the last one was uh... uh... searching friendly and linear linear yeah linear so i i i think it was the linear is the same thing i was put into the place you know uh... that's my uh... two soft words thank you for your input

Deepak : we've come to the end of our interview thank you very much

C.1.7. Participant - 7

Deepak : Alright, I'm with Participant-7 and you've just come back from testing the product. But before we start our interview, I'd like you to talk about your role at KLM.

Participant-7 : Okay, I'm a visual designer. I've worked at KLM for six years and also done a VR for six years. So I know a lot about the product inside of KLM. And I just do the design so I visually decide how it looks now it feels.

Deepak : With your past experiences in VR and have you had experience in AI before?

Participant-7 : No, not work related. I'm interested in it myself and I use ChatGPT for my work. So I do know a lot about the knowledge and I try to stay into an innovative mindset. So I do Google a lot or search a lot online. I find a lot. And I'm around co -workers who work with it constantly but I've never used it work related. But with your knowledge that you have so far,

Deepak : How did that impact your perception of the quality of training that you just witnessed?

Participant-7 : That's an interesting question. I was kind of seeing how far the conversation was able to go. But also hit the mindset that it's a project. So maybe that was a little bit of a bias. But for me it was very interesting how far the AI voice could understand what

I was trying to tell them. I think there was one part in the beginning where he ignored my question. And then I tried to get back to it again, like asking like did you hear what I said? And then he tried to ignore it again and then we reset it and then we did it again. And then at the beginning it was a little bit iffy again. And then it started into a full on discussion where I kind of forgot that I... that I wasn't having a discussion with a person. So that one was really successful in my opinion. So I was searching for that kind of realism and I got it. Okay, nice. Yeah.

Deepak : So apart from the realism factor, whether other experiences or aspects that you witnessed that influenced your overall rating of this training?

Participant-7 : Um... Well, I normally only have conversation with AI via text and never via speech. So for me this was really interesting in how it sounded and if it was an actual person. I think it was... I always with chat GPT, you see it adds too many words sometimes. Like this is a very ohh and ummm, like special... It adds a lot of words. I think this... You could still see that in the way the AI talked to me that it was... Adding like oh this is such a prominent No person would really maybe say that. Maybe in this business world they would I don't have enough experience into that but

Deepak : But if you if you knew I mean or if you did not know that you were going to talk to a bot Would you have realized?

Participant-7 : Yes It was because it it took pretty long before it answered back Um And he didn't sound nervous, but sometimes like if he ignores your your stuff you're like, oh, maybe he's nervous But yeah, he sounded really well confident So I think I would have noticed yes for sure in the in the state of this right now with how fast it can answer back For me it took a while. Oh, and there were a lot of um Because it does the ums and the actually those didn't really sometimes he answered afterwards like oh, yes That's really nice. And then he actually did like hmm. It was like Did he feel that my answer was nice or did he feel like it wasn't yeah And the little umson always threw me off and then he actually gave another answer

Deepak : Okay, but how did all of this come into uh the intuitiveness of this training program? Did you find it very intuitive?

Participant-7 : Yes, very intuitive because I could just Even if it didn't went the way I wanted it or that I expected you could always sidetrack and be like oh, okay Let's ask something else. Okay But I'm not sure if That would work if the person wouldn't know if it was an ai. Yeah Would they know like oh, okay, let's he doesn't really get what I'm saying. Let me try something else You wouldn't do that with the real person. I think Yeah, um, so in the intuit intuitiveness it depends on if the user does or doesn't know I knew So for me it was different. I was biased in that way In terms of feasibility, do you think this product has potential in replacing existing training? Very much. I even think it's necessary, in my opinion. Yeah, very much. I think this is what AI has created for one of the biggest enhancements of AI. This type of training, this type of conversation, I think it's really realistic.

Deepak : Do you think it's because of how unpredictable the conversation was?

Participant-7 : Yeah, there's just no script. So it becomes fully unpredicatble and that's a

motivating factor to look upto in a conversation.

Deepak : Okay. Just one last question. You were asked to fill a survey after you used the product. Yeah. And you were asked to describe the product in three words. I want you to tell me why you chose the words that you chose.

Participant-7 : At the second part, right? I think I answered something like, felt like real. Yeah. Because that was just my experience. At the end, I totally, I kind of had that, I work with VR. And in VR sometimes it can really make you feel like you are actually there. And sometimes, even when you know I'm standing in some office room, your mind tells you, no, I'm in this environment. I had that at the end of the call. Okay, that's very interesting. Oh, I'm not talking. I needed to tell myself, oh, I'm talking to something these guys created, you know? Instead of having an actual really bossy Indian man that we're saying, like, no, we're gonna do it my way. So I thought it was, there was like an interesting perspective for me, that when it works really well, it works really well. I'm glad you feel that way.

Deepak :Yeah, with this we've come to the end of our interview and your answers have been very valuable.

Participant-7 : Thank you. Thank you.

C.1.8. Participant - 8

Deepak : Alright, welcome back Participant-8. You just had a small demo of the product but before we start our interview I'd like to know about your role at KLM.

Participant-8 : Yeah, what's my role? My role at the KLM, I'm the senior researcher for new technologies. So AI technologies, XR technologies and other technologies such as blockchain and that sort of thing. My focus at the moment is basically AI and XR technologies. I make sure that everything we do has academic grounding that it's based on something.

Deepak : You mentioned that you have experience in AI. How has your previous AI experience impacted your perception of this training program?

Participant-8 : Positively. I think if you're able to use AI especially for the unpredictable aspects of AI I think it could be very good and very worthwhile to use these in trainings.

Deepak : Apart from unpredictability, were there any other features or aspects that you considered while giving your overall rating of the product?

Participant-8 : Usability. Usability. And difficulty. Because the questions we got, there was one question in between which says how difficult was this to do. It sounds like a close question but actually it's really a wide open question because everybody's perception of difficulty means something else. Difficult can be, I find it very difficult to communicate. That's the same. Difficulty could be I couldn't get the phone to work. So this is a very open question. My aspect of difficulty is the communication, understanding and trying to get a feeling of what somebody is saying.

Deepak : But in terms of usability, how intuitive did you find this method of training?

Participant-8 : very very Very very usable for me because it's a phone number you phone It's easy. It's a click of a button. I work in a technology Environment where everything is quite complicated But this I press a button and it works So that's very important because you're making products for the general public.

Deepak : Yeah, and you also told that you found the product to have a bit of unpredictability But if we move into the aspect of believability Did you find the conversation believable?

Participant-8 : With after the first two minutes. Yes, okay the first two minutes is adjusting So I know this is an AI bot The next time you do this test don't tell anybody it's an AI bot to say to somebody on the line But after a while you still go into the immersion and you still go into the conversation, your natural talent takes over.

Deepak : Yeah, but do you think your opinion would be different if you were not informed of the fact that you were talking to a bot from the very beginning?

Participant-8 : Yep, definitely. Do you think you would not have even questioned the fact that it was a bot or it would have taken a lot longer for you to realise? At the moment it's not technology strong enough to be considered. But it's getting close.

Deepak : So could you give me a few reasons as to why you think currently you don't find it extremely believable?

Participant-8 : I find it believable only that there are a few things which are missing still for me. For me the boss, the umm oohhh, they're not correct in some places. And last but not least, I'm not able to interrupt. Interruptions is very important.

Deepak : But in terms of the feasibility of this product, do you think there is future where this product will really kick in?

Participant-8 : 110% I believe it. I believe really in it. Alright, with this we come to our last question.

Deepak : After you used the product you were asked to fill a survey and you were asked to describe the product in three words. I want you to tell me why you chose the words that you chose.

Participant-8 : What did I choose? I chose a fresh, fun, I don't know what third word was, I don't remember anymore. But any fresh because it's new. It gets rid of the traditional way of training. It's finished, it's gone, we don't need any actors anymore. We got this now. So for me that's fresh. you gives new insights and gives me a new way to learn skills. Fun, yeah, it's fun because you can ask the bot anything you want and you can again answer back. And you have also control, which is, and again, going back to the unpredictability, you can anticipate any way you want, then you're not restricted by a script. Yeah, yeah. I can't remember what third word was. Thing, yeah. Is it fair, fun and fresh fun? I can't remember. Would you like to make up a word? No, yeah. Oh yeah,

Deepak : if I can make up a word.

Participant-8 : I would say something like innovative, very, very, very innovative. Yeah, very innovative.

Deepak: All right. With this, we come to the end of our questioner. Your inputs are very important as they will help shape the future of this product. So thank you for your time.

C.1.9. Participant - 9

Deepak: Alright Participant-9, you are participant number 9. You just had a small preview of the product and but before we start I want you to tell me your role in this KLM project.

Participant-9 : I'm the programmer I would say. I'm responsible for making the app work. Like make sure that the telephone app is working.

Deepak : With your prior experience into creating this program, how did this experience impact the way that you perceived the product?

Participant-9 : With my prior experience?

Deepak : Yeah, your prior experience into the field of AI. How did those experiences help in creating a perception about this product in terms of giving it a rating?

Participant-9 : Are you asking me as a developer or just as a user of this training program?

Deepak : What I'm saying is if your prior AI experiences helped or impacted your perception of this AI.

Participant-9 : You're asking if I'm biased? Yeah, of course I'm biased. I already knew the 11Labs voice. It's a standard voice that I already knew. Well, I built the whole thing so yes, of course I'm completely biased.

Deepak : Okay, hypothetically let's say you did not know anything about this product. Would you have suspected this product?

Participant-9 : I expected it to be something other than human. And if so, why? I think so, yes. And the reason is because it's not fast enough yet.

Deepak : So you think the response time is the only thing that, you know, goes into the believability of the training or you think there are other aspects as well?

Participant-9 : Yeah, the voice is not completely, completely human, but it comes super close. So I think you could fool me if I did, had not known about it and I was just a farmer and never knew about technology. And what else there? Yeah, there's the ChatGPT response. I think this is very, very good. Sometimes when it lists things, it says one, two, three. This is very non -human. Non -human, okay.

Deepak : But apart from this, in terms of user experience, did you find, the bot to be unpredictable in terms of speech.

Participant-9 : Yeah, it was very unpredictable. I had no idea what it was going to say at any point in time.

Deepak: Okay, but did you find the user experience intuitive? Was it easy to use?

Participant-9 : Intuitive, yes, yes. I mean, you talk, you respond. Yeah, fairly straightforward. Normal, for example.

Deepak : Yeah, but in terms of product feasibility, do you see a future where this product will come into fruition and if so, what?

Participant-9 : I think so, yes. I think we're very close to having it very believable. And then I think the opportunities are endless. Yeah.

Deepak : Final question, you were asked to fill a survey after you reviewed the product and you were asked to describe the product in three words. I want you to tell me why you wrote what you wrote.

Participant-9 : Okay, so I wrote, I believe, fun, I believe, future, I believe, one more word. Okay, that's fine.

Deepak : You can make up a word...

Participant-9 : Innovative. Innovative.

Deepak : And why do you think it is all these?

Participant-9 : Okay, so fun, well, it's fun because it's almost human. It makes it fun that you managed to build something like that. Yeah. I'm not sure if I would have believed at the start of this project. Innovative. Yeah, it's new. Well, we found one GitHub where they did it, but most of the building process what I did, I just tried... I didn't have an example. So that's why I think it's very innovative. And then future, I think, because I would see this being used widespread in the future.

Deepak: Yeah. Well, that comes to the end of our interview. Your results, I mean, your answers have been very helpful in describing the nature of our product. More. So thank you. Thank you.

C.1.10. Participant - 10

Kshitij : Okay, so let's do the qualitative question. The first domain is about quality of training and the question is like, so how do you do the ratings for the AI communication training based on their experiences?

Participant-10 : So how do you talk about the previous training experience that you felt that the KLM offered before implementing the product?

Participant-10 : Yeah, did I think they failed? No, what did you expect?

Kshitij : Like, was your experience talking to them before they started testing this product? How was the possibility of the training experiences that they faced?

Participant-10: Yeah, I think it was very predictable, very intuitive actually. So people could learn a tParticipant-3 instead of learning a skill. And yeah, that's what I believe was going on.

Kshitij : So what was their experience after they got introduced to this product? And what was their idea about this new product that we...

Participant-10 : Well, it surprised me, they were very convinced of the product. They felt it was realistic even though they knew that they were talking to an AI bot. It felt realistic for them and they also mentioned that it would be a good addition to their training programs where it could definitely make an impact.

Kshitij : Okay, so the next question is about intuitiveness of the training. So what difference did you find before and after introducing the product?

Participant-10 : So before I noticed that people didn't think that the training programs were intuitive and they were very easy and predictable. And after I noticed that... Yeah, that they found it very unpredictable and very intuitive and they could just play along and they found their questions being asked on a very complex level and intellectual level as well.

Kshitij : And the third one is about the unpredictability of training. So what experience did you, what feedback did you get from the employees who were then predictability of the training?

Participant-10 : Yeah, they found it very unpredictable and they often found it very hard to communicate with the bots and not in a bad way. They found the questions very hard and what I noticed is I said like, oh shit, I don't have the information right now but I have to check up on this and then I'll send it later to you.

Kshitij : So yeah. Nice. And the fourth one is about the unpredictability of training. So, was there feedback after being introduced to the product? How was their ability?

Participant-10 : Some of them found it believable, some of them did not, but it depended on how they were biased on the AI used. So some of them believed that it didn't go in depth enough. I made some small adjustments in the prompt and then afterwards the believability went a lot higher actually. And everybody could have a more believable conversation. And if you play with latency as well, the less latency it has, the more believable it was for them. No, okay. And the next one is about the feasibility of the training.

Kshitij : So what was your experience being in contact with the KLMM plus about the feasibility of training?

Participant-10 : Like, what do you mean with the feasibility?

Kshitij : Like, how, did they find the product more user -friendly or kind of experience, made it comfortable or something? Yeah, definitely. More interaction with the product.

Participant-10 : Yeah, they could because they could really answer true and they could ask more. And it was more feasible because they didn't have to feel that there was an actor. They had to feel eventually that they were talking to something. It was not human. They didn't have to feel the feeling that they were talking to a human. They were talking to something that was intellectual and they could have a conversation with them on a deep level. So another aspect of the feasibility is that it would be very easy to conduct because you don't need to hire the actor. It could be done from home. It could be done from the office. It could be done anywhere. So that makes it feasible as well.

Kshitij: Okay. Thank you, Participant-10. And we come to the last question in this session. And can you describe what was your experience being in contact with them and then how were their expressions and emotions about the product?

Participant-10 : They were very surprised and they were nodding with an impressive nod. They were saying this is beyond what we imagined and it was a good product. There were some minor adjustments and improvements such as the cracks in between the audio and the latency even. For example, the first lady who came in was very skeptical but another guy who was a big fan of it as well. We had mixed feelings about it but usually I think everybody had a really good idea about the product. To be sure that they were not aware of the product, they wanted to communicate with the product. It was just a random thing. Yeah, it was a random thing and they didn't know it and they were very surprised. Some even found it hard. to make a sale. So you could see who was a salesperson who had it in them and who didn't.

Kshitij : Thank you for your feedback. So we conclude with our question for today. Yeah, okay.

Participant-10 : Thank you.

C.2. Phone Instructions

Number: +31970102058 78

Introduction: You'll be interacting with Rajeev Malhotra, the CEO of Malhotra Tech Innovations, over the phone. However, please be aware that Rajeev is an AI-powered chatbot designed for this call.

C.2.1. Objective

Your goal is to engage in a realistic sales negotiation with Rajeev, with an emphasis on meeting his specific objectives.

C.2.2. Who is Rajeev Malhotra?

Rajeev is known for his commitment to business ethics, cultural values, and sustainability. He has three primary objectives for this conversation:

1. Negotiate a ticket price below 500 euros.
2. Ensure the partnership respects cultural values and norms.
3. Define and agree upon sustainability targets.

C.2.3. Instructions

1. **Initiating the Call**: Wait for the audio to be played. After the introductory message, you may begin speaking.
2. **Speaking**
   (a) Ensure you speak clearly and at a steady pace
   (b) Once you begin a sentence, make sure to complete it without lengthy pauses. If there's a prolonged silence mid-sentence, only half of your sentence may be captured and sent.
   (c) If you need to think or pause for a moment, use the mute function. Once you're ready to speak again, unmute and continue.
3. **Focus Area**
   (a) Understand Rajeev's three main objectives and ensure that your responses and offerings align with them.
   (b) Remember that while price is essential, cultural values and sustainability are equally important to Rajeev.
   (c) While Rajeev is receptive to negotiation, he has strong principles, so ensure mutual respect and understanding in your communication.
4. **Technical Tips**
   (a) Stay in an area with good reception to prevent call dropouts.
   (b) Keep your phone charged or have a charger handy to ensure the call isn't interrupted.
   (c) Avoid background noise which might interfere with the call clarity.
5. **Concluding the call**
   (a) Once you feel that the objectives have been met or the discussion has reached a natural conclusion, you may end the call.