

Politecnico di Milano

SCUOLA DI INGEGNERIA INDUSTRIALE E DELL'INFORMAZIONE
Master Degree – Computer Science and Engineering



**VIRTUAL AND AUGMENTED REALITY
+ CONVERSATIONAL AGENTS FOR
PERSONS WITH AUTISM: A CASE
STUDY**

Relatore

Prof.ssa Franca GARZOTTO

Co-Relatori

Dott. Alberto PATTI

Dott. Francesco VONA

Candidati

Giacomo Bacchini – 10688201

Anno Accademico 2020 – 2021

Ringraziamenti

Il mio percorso per arrivare fin qui è stato lungo e non privo di intoppi. Per quanto non poter partire per l'esperienza di exchange per via della pandemia sia stato inizialmente un duro colpo, questo mi ha dato la possibilità di partecipare al corso di AUI e di conoscere la Prof.ssa Garzotto e i ragazzi dell'I3Lab. Ringrazio Monica, la mia compagna di corso: insieme a lei ho preso la decisione di iscrivermi al corso che ha poi portato allo sviluppo di questa tesi. Un grazie anche a tutti i miei amici del corso: Kelly, Andrea, Angelo, Irene, Mariachiara, Emilio, Monica, Ninni. Non dimenticherò mai il periodo trascorso con voi a Milano. Un ringraziamento speciale a Gianmarco, che conosco da sei anni e con cui ho completato il percorso universitario triennale a Bologna e con cui ho preso la decisione di affrontare questa magistrale al Politecnico. Ringrazio i miei colleghi con cui ho realizzato questo progetto, Davide e Andrea, due persone che non conoscevo ma con cui sono riuscito a costruire un'amicizia e con cui nell'ultimo periodo ho trascorso tantissimo tempo, pur essendo in città diverse. Rimanendo sempre alle persone che ho conosciuto a Milano, un pensiero anche alla mia maestra Chihiro: in questi anni grazie a te ho potuto scoprire la bellezza di studiare una lingua completamente diversa dalla nostra.

Uscendo dall'ambito universitario, i ringraziamenti più importanti vanno alla mia famiglia per il supporto che mi hanno dato in questi anni. A mio fratello Matteo che sta iniziando ora il suo percorso universitario, a mia madre Cristina che mi ha ascoltato e consigliato nei periodi più difficili, a mio padre Claudio che mi ha sostenuto nelle mie scelte.

Ringrazio infine tutti i miei amici di sempre delle mie zone che conosco da tempi immemori e che anche con una semplice uscita possono svoltarti la giornata.

Ai miei due amici e coinquilini Maina e Diste con cui ho creato tanti bellissimi ricordi e modi di dire che ci portiamo appresso ancora oggi. Ad Ambra, sempre propositiva e disponibili per ascoltarti. Ad Alessia, con cui ho creato indelebili ricordi negli anni dell'adolescenza. A tutto il gruppo di giocatori di ruolo che mi hanno alleggerito le serate in questo periodo.

Ringrazio la prof.ssa Garzotto per avermi dato la possibilità di partecipare alla realizzazione di questo progetto. Un sentito ringraziamento ai correlatori Francesco e Alberto per avere seguito me e gli altri ragazzi durante tutto il nostro percorso di tesi. Un altro ringraziamento va a MariaGiovanna per il supporto ricevuto durante la fase di design dell'applicazione e a Jessica per l'aiuto offerto nella fase di studio sperimentale. Infine ringrazio tutti i collaboratori dell'I3Lab per avere fornito il materiale e la strumentazione necessaria allo svolgimento della tesi.

Giacomo

Sommario

Il Disturbo dello Spettro Autistico (ASD) è un gruppo di disturbi dello Sviluppo Neurologico (NDD) causato da una condizione correlata allo sviluppo del cervello. Ha un impatto sulla comunicazione verbale e non verbale e sulle interazioni sociali, causando mancanza di abilità nello sviluppo, nel mantenimento e nella comprensione delle relazioni, interessi ristretti e comportamenti ripetitivi. Tutte queste difficoltà rendono le attività quotidiane, come andare a scuola o uscire a fare la spesa, molto difficili da completare e talvolta impossibili da compiere da soli. 5A è un progetto che mira a migliorare l'autonomia delle persone affette da autismo nei contesti di vita quotidiana attraverso l'uso di tecnologie immersive, ovvero: realtà virtuale e aumentata e agenti conversazionali. Il caso d'uso scelto per questa tesi è il museo, appartenente al contesto di accessibilità e fruizione di servizi commerciali, culturali e ricreativi. L'esperienza inizia con una formazione preliminare in Realtà Virtuale in cui gli utenti sono immersi in un ambiente museale artificiale e sono guidati attraverso una serie di compiti grazie all'uso di un Assistente Virtuale e altri elementi virtuali. Al fine di facilitare la comprensione e l'apprendimento dell'esperienza svolta in VR, questi elementi sono condivisi come elementi transmediali nell'app in realtà aumentata, che può fungere sia da formazione secondaria in loco, sia da guida completa per l'utente durante la visita. Entrambe le applicazioni sono altamente configurabili in base alle esigenze dell'utente per massimizzarne i benefici apportati. Il seguente documento descrive il processo di progettazione e implementazione delle due applicazioni. Infine, è stato condotto uno studio esplorativo per testare l'usabilità e l'accettabilità del sistema.

Abstract

Autism Spectrum Disorder (ASD) is a group of Neurodevelopmental Disorders (NDD) caused by a condition related to brain development. It impacts verbal and non-verbal communication and social interactions, causing lack of skills in developing, maintaining and understanding relationships, restricted interests and repetitive behaviors. All these difficulties make daily tasks, like going to school or going out for grocery shopping, very difficult to complete and sometimes impossible to accomplish alone. 5A is a project that aims to improve the autonomy of people with autism in everyday life contexts through the use of immersive technologies, namely: virtual and augmented reality and conversational agents. The use case chosen for this thesis is the museum, belonging to the context of access and use of commercial, cultural and recreational services. The experience starts with a preliminary training in Virtual Reality where users are immersed in an artificial museum environment and are guided through a set of tasks thanks to the use of a Virtual Assistant and other virtual elements. In order to facilitate the understanding and learning of the experience carried out in VR these elements are shared as transmedial elements in the Augmented Reality app that can act both as a secondary training in loco as well as a complete guide for the user during the visit. Both the applications are highly configurable according to the user's needs to maximize the benefits. The following document describes the design and implementation process of the two applications. Moreover, an exploratory study has been carried out to test the usability and acceptability of the system.

Contents

Ringraziamenti	ii
Sommario	iii
Abstract	iv
Contents	vii
List of Figures	x
List of Tables	xi
1 Introduction	1
1.1 Contribution	3
1.2 Thesis structure	4
2 State of the Art	6
2.1 Autism of Spectrum Disorder (ASD)	6
2.2 XR technologies for ASD	7
2.2.1 Virtual Reality for ASD	7
2.2.2 Augmented Reality for ASD	11
2.3 Conversational Agents	13
2.4 Hardwares for Virtual Reality	17
3 Needs, Goals and Requirements	18
3.1 Needs	18
3.2 Goals	19
3.3 Requirements	19
3.3.1 VR Requirements	20
3.3.2 AR Requirements	21
4 Design	22
4.1 Tasks	24
4.2 Transmedial elements	28
4.2.1 Virtual Assistant	29
4.2.2 Virtual Guide	33
4.2.3 Minigames	36
4.2.4 Pathfinding	37
4.2.5 UI design	37

4.3	Virtual Reality (VR) design	41
4.3.1	UX flow diagram	41
4.3.2	Environment	46
4.3.3	Interactions	51
4.3.4	User Interfaces	55
4.3.5	Customization of the Experience	60
4.4	AR design	62
4.4.1	UX Flow diagram	62
4.4.2	Interactions	76
4.4.3	Levels of difficulty	78
4.4.4	User Interfaces	78
4.4.5	User Tracking	84
5	Implementation	87
5.1	Unity	87
5.1.1	Libraries and Plugins	88
5.1.2	Packages and Assets imported from external sources	89
5.2	Blender	90
5.3	Speech-to-text and Text-to-speech	92
5.3.1	Microsoft Azure STT and TTS	92
5.3.2	Android and iOS native STT and TTS	92
5.4	5A-Museums - VR architecture	93
5.4.1	The modular structure	93
5.4.2	Main components	94
5.4.3	Gamification process: Treasure Hunt	96
5.4.4	Conversational Agent and dialogues	98
5.5	5A-Museums - AR architecture	100
5.5.1	AR Session	100
5.5.2	Modularity through QR codes	100
5.5.3	Pathfinding	101
5.5.4	GPS Positioning	102
5.5.5	Image Recognition	103
5.5.6	Main Components	104
6	Exploratory Study	108
6.1	Goal of the Study	108
6.2	Variables	108
6.3	Hypotheses	109
6.4	Subvariables and their Measurement	109
6.5	VR	110
6.6	AR	110
6.7	Research Design	111
6.8	Participants	112
6.9	Execution	112
6.9.1	VR session	112
6.9.2	AR session	112
6.10	Questionnaire	114

7	Conclusions	115
7.1	Contribution	115
7.2	Limitations	115
7.3	Future works	116
A	First Appendix	118
A.1	ATI Scale Questions	118
A.2	Usability	118
A.3	Interface	119
A.4	Inputs commands	119
A.5	Display and Audio	119
A.6	Engagement	119
A.7	reliability	119
A.8	Virtual Assistant	119
A.9	Virtual Guide	120
A.10	Facilitator Graphics	120
A.11	Text-to-Speech	120
A.12	Speech-to-text	120
A.13	Facilitators	120
A.14	Exclusive questions on VR questionnaire	120
	A.14.1 Facilitators	120
	A.14.2 Distractors	121
	A.14.3 Treasure Hunt	121
A.15	Exclusive questions on AR questionnaire	121
	A.15.1 Image recognition	121
	A.15.2 QR codes	121
	Acronimi	122
	Bibliography	126

List of Figures

Figura 1.1	Transmedial elements	3
Figura 2.1	VR-ATT	8
Figura 2.2	Project from Emil Rosenlund Høeg at al.	9
Figura 2.3	Knowledgemon Hunter	11
Figura 2.4	Minigame inside ARTis	12
Figura 2.5	Mosoco helping children with autism	13
Figura 2.6	Child interacting with Echoes	14
Figura 2.7	FaceSay experience	15
Figura 2.8	Bob Fishing Shop conversational agent	15
Figura 2.9	INDOAR assistant	16
Figura 2.10	Oculus Quest 2	17
Figura 4.1	Key areas of a generic museum	24
Figura 4.2	Pictures of the key areas	25
Figura 4.3	Key areas inside the <i>Museo di Storia Naturale</i>	25
Figura 4.4	UX flow diagram	28
Figura 4.5	Virtual Assistant 3D model	30
Figura 4.6	Analysis of the design of popular characters	31
Figura 4.7	Sketches	31
Figura 4.8	From sketch to 3D	32
Figura 4.9	Expressiveness of the VA. Some of this were unused in the final version.	32
Figura 4.10	Reconstruction of the Scipyonix specimen nicknamed "Ciro"	33
Figura 4.11	Analysis of the design of popular characters	34
Figura 4.12	Character design	34
Figura 4.13	Virtual Guide 3D model	35
Figura 4.14	Jump animation	35
Figura 4.15	Treasure Hunt	36
Figura 4.16	Encyclopedia menu	37
Figura 4.17	Arrow pointing to the next task in AR	38
Figura 4.18	Typical conversations with the Virtual Assistant (VA) in the two applications	39
Figura 4.19	Comparison between PCS and the graphics used in 5A	40
Figura 4.20	Oculus controllers	41
Figura 4.21	UX flow chart of the VR experience	42
Figura 4.22	Outdoor scenario extended UX Diagram	43
Figura 4.23	Hall scenario extended UX Diagram	44
Figura 4.24	Exhibit room scenario extended UX Diagram	45

Figura 4.25	Treasure hunt extended UX Diagram	45
Figura 4.26	The city in the outdoor scenario	46
Figura 4.27	Comparison between the real and virtual museum	47
Figura 4.28	Moving objects and people on the city streets	47
Figura 4.29	Houses and skyscraper	47
Figura 4.30	Comparison between the real and virtual hall	48
Figura 4.31	Door communicating with the first exhibit room	48
Figura 4.32	The first exhibit room	48
Figura 4.33	Different methods of item display inside the museum	49
Figura 4.34	The second exhibit room	50
Figura 4.35	Allosaurus reconstruction	50
Figura 4.36	Showcase of the evolution of dinosaurs	50
Figura 4.37	Virtual Assistant extended UX Diagram	52
Figura 4.38	Interactions with the Virtual Assistant	53
Figura 4.39	Interactions with the Virtual Guide	53
Figura 4.40	The process of clicking on a Treasure Hunt item	54
Figura 4.41	Ticket clerk asking the user if they want to buy a ticket	54
Figura 4.42	User bring the ticket to the Guide to check it	54
Figura 4.43	Activating the microphone	55
Figura 4.44	Main menu with language panel	56
Figura 4.45	Main menu with starting zone panel	56
Figura 4.46	Main menu with difficulty panel	56
Figura 4.47	Virtual Assistant's default dialogue box	57
Figura 4.48	Virtual Assistant's help dialogue box	57
Figura 4.49	List of Facilitator graphics.	58
Figura 4.50	Encyclopedia UI	58
Figura 4.51	Encyclopedia UI	59
Figura 4.52	Facilitators	60
Figura 4.53	Attractors	61
Figura 4.54	Temporary	63
Figura 4.55	Examples of menu and tutorial screens	64
Figura 4.56	Example of a screen displayed while outdoor	65
Figura 4.57	Outdoor AR UX Diagram	66
Figura 4.58	Example of a screen displayed while indoor	67
Figura 4.59	Hall AR UX Diagram	68
Figura 4.60	Virtual Assistant help AR UX Diagram	69
Figura 4.61	Example of help screen inside an exhibit room	70
Figura 4.62	VA help AR UX Diagram	71
Figura 4.63	Encyclopedia AR UX Diagram	72
Figura 4.64	Screens showing how the Encyclopedia works	72
Figura 4.65	The Treasure Hunt minigame	74
Figura 4.66	Virtual Guide AR UX Diagram	75
Figura 4.67	Holding the Vocal input button	77
Figura 4.68	Used and proposed solutions to give suggestions to the user	77
Figura 4.69	Menu screens	79
Figura 4.70	The three animations of the VA sprite, respectively: Idling, Helping and Vocal input	80

Figura 4.71	The two dimensions for the VA dialogue boxes and menus . . .	80
Figura 4.72	Color palette used for the VA dialogue boxes	81
Figura 4.73	The two dimensions for the VG dialogue boxes and menus . . .	81
Figura 4.74	Color palette used for the VG dialogue boxes	81
Figura 4.75	Facilitator graphics used in the AR app	82
Figura 4.76	Instructions to play the Treasure Hunt	82
Figura 4.77	Tutorial screens	83
Figura 4.78	Beacon technology	84
Figura 4.79	5A-Museums QR codes	85
Figura 5.1	Unity IDE	88
Figura 5.2	Guide's body is sculpted	90
Figura 5.3	Assistant's body is poly modeled	91
Figura 5.4	Poly modeling vs. sculpting in shaping a head.	91
Figura 5.5	Animations were made by rigging characters and then exporting to Unity	92
Figura 5.6	Museum items	96
Figura 5.7	Example of a loader method in the DialogueManager	99
Figura 5.8	Example of virtual maps generated by the AR application . . .	102
Figura 5.9	Serializable classes instantiated through QR code scanning . .	106
Figura 6.1	Map of the Museo Civico di Storia Naturale di Milano	111

List of Tables

Tabella 2.1	Comparison of the features of several conversational agents .	16
Tabella 3.1	VR Goals and Requirements matrix	20
Tabella 3.2	AR Goals and Requirements matrix	21

Chapter 1

Introduction

This thesis is a collaborative work between me and two IT engineers, Andrea Panizzi and Davide Perugini. It is difficult to separate the contribution, as it was a strict collaboration in which everyone engaged. That said, my contribution has been stronger in the research, user experience, dialogues and transmedial elements parts, while most of the parts strictly related to coding was done by my colleagues. Regarding Virtual Reality, most of my work concerned the creation and the import of 3D contents (3D environments, 3D assets and animations) from Blender to Unity. Inside the Unity project instead, the work focused on handling animations, texturing, lighting and sounds. In relation to Augmented Reality, I tested the range and the possibility of different applications for image-based recognition and user tracking system technologies. Further, I worked on measurements and creation of *DynamicMap* virtual maps and on all user interfaces and sprite animations. To document the work in its completeness, all the parts are fully reported in both theses, with a more pronounced focus on the respective contributions.

According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), ASD is a group of Neuro-Developmental Disorders characterized by deficits in social communication and social interaction across multiple contexts, including deficits in social reciprocity, nonverbal communicative behaviors used for social interaction, and an impairment of growth, maturation, or development of the central nervous system [9] [15]. Autism spectrum disorder begins in early childhood and eventually causes problems functioning in social, personal, and academic areas. The ASD population has a difficult time generalizing knowledge from one context to another, making it difficult for them to apply what they learn in a general context to specific situations [19]. Furthermore, people with ASD are frequently self-absorbed and appear to live in their own world, with limited ability to speak and engage with others [21]. Autism is referred to be a "spectrum" disorder since the kind and intensity of symptoms that people experience vary greatly. Individual clinical characteristics are noted through the use of specifiers such as "with or without accompanying intellectual impairment" or "with or without accompanying structural language impairment" during the diagnosis of autism spectrum disorder, whereas historically ASD was subdivided into several different disorders (e.g. Asperger's Syndrome). These specifiers allow clinicians to personalize the diagnosis and communicate a more detailed clinical description of the

affected individuals.

Regarding the levels of severity of the symptoms, the DSM-5 characterizes ASD by dividing it into three levels of severity. The third level is the most severe, requiring very substantial support in Daily Living Skills (DLS) due to serious deficits in verbal and nonverbal communications skills [DSM-5]. The higher functioning individuals placed in the first and least severe level of the DSM-5 are capable of verbal communication but still require support due to difficulties processing sensory input and understanding social cues. According to the latest data in recent publications in 2017 the prevalence of autistic spectrum disorders globally was roughly 1% of the total population, or around 62 million people [48], while in Italy, around 442 thousand people, or 0.62% of the population, were diagnosed with ASD in the same year, though the data are scarce and uncertain [39]. Also, males are approximately 4-5 times as likely to have autistic spectrum disorder versus females [48]. People affected by ASD show severe difficulties in several aspects of daily life, and find themselves overwhelmed and anxious in unfamiliar or crowded environments such as a public spaces. This aspects will be further analyzed in the first paragraph of the State of the art (see 2.1).

Virtual Reality is the simulation of a world through a computer modeled space that lets the user interact with an artificial three-dimensional visual or other sensory environment. Today, the expression mainly refers to the immersive simulated experiences provided by modern technological devices such as VR headsets and multi-projected environments. Thus, the term VR refers to technologies that can interact with the sensorial perceptions of users and present them with an artificial world perceived as real. Therapists and researchers in the field of ASD started to use VR in the mid-1990s, with the first examples of its use as a mode of exposure therapy for specific phobias [4]. They often deployed the technology to create virtual environments to help autistic people prepare for encounters or situations that could be stressful [49]. Nowadays, the technology has improved greatly and it is possible to provide a different level of support. Advances in voice recognition, natural language processing, and artificial intelligence, as well as technological devices becoming more common and affordable, have led to the increased use in therapy. One of the most notable benefits is that virtual reality affords incomparable control over the environment, allowing researchers and clinicians to arrange environments to best promote learning and generalization [4]. Furthermore, the use of virtual conversational agents with unconstrained natural language input capabilities in healthcare is an emerging field of research that may have the potential to benefit health across a broad range of application domains [28].

Augmented Reality is defined as an interactive experience of a real-world environment where the objects that reside in the real world are enhanced by computer-generated perceptual information, sometimes across multiple sensory modalities, mainly visual and auditory [23]. The three basic features are a combination of real and virtual worlds, real-time interaction, and accurate 3D registration of virtual and real objects, seamlessly interwoven with the physical world such that it is perceived as an immersive aspect of the real environment [11]. Augmented reality is used to enhance natural environments and offer perceptually enriched experiences. With the help of advanced AR technologies (e.g. adding computer vision, incorporating AR cameras into smartphone

applications and object recognition) the information about the surrounding real world of the user becomes interactive and digitally manipulated. In recent years, augmented reality has been explored for many applications for smartphones, from gaming and entertainment to education and business [22]. Several studies have focused on using augmented reality (AR) to help children with autism in many areas of development, in particular by extending children’s interests, “augmenting” objects available in the surrounding environment and simultaneously increasing the intrinsic motivation [3].

1.1 Contribution

On this basis, the project 5A, or Autonomie per l’Autismo Attraverso realtà virtuale, realtà Aumentata e Agenti conversazionali, was created. 5A is an open-source project that has been conceived and developed in collaboration with *I3Lab*. It is financially supported by *Fondazione Tim*. The project aims to improve the quality of life and autonomy of people affected by autism with the use of Virtual Reality VR and Augmented Reality Augmented Reality (AR) used in tandem. The main idea is that the user is at first introduced to a real-life scenario in a simulated virtual environment where they can experience and do a preliminary training in a safe and controlled environment a situation similar to that they would encounter in real life. Subsequently, the user is introduced in the actual real-life environment with the support of an AR app easily accessible on modern smartphones that superimposes interactive multimedia elements on the vision of the surrounding physical world. By having multiple selectable levels of help features the app can function both as further training and as an actual helping guide that acts as a support for the user while moving through the environment.

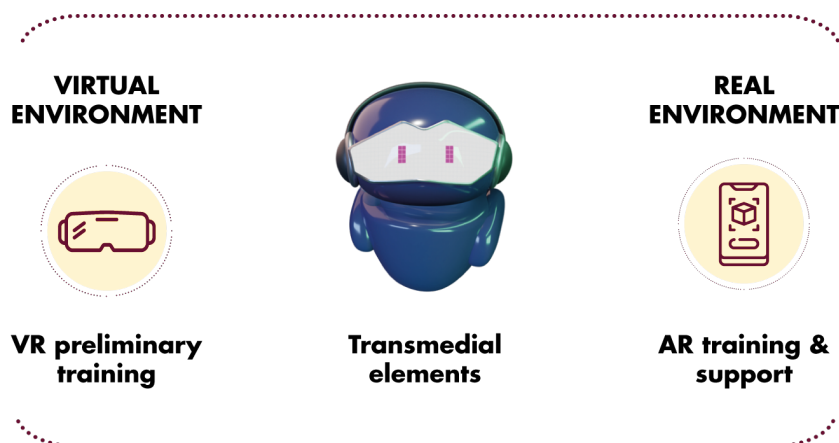


Figure 1.1. Transmedial elements

This two-step approach helps them generalize the skills acquired while in the VR environment, by finding the same elements in AR: in both Virtual and Augmented Reality apps the user will have a similar experience, feedback and help while performing the various tasks. To convey the similarity and increase the generalization of the experience inside the user’s mind, similar virtual elements, called Transmedial elements, are used. One of the key parts of Trans-medial elements are the conversational agents

that provide suggestions, vocal and textual feedback. They function as guides, assisting people with ASD in identifying specific elements and performing fundamental tasks with the goal of completing in autonomy the scenario.

Various daily-life scenarios like taking the subway or shop in a supermarket have been created inside the family of 5A apps. The project's main strength, that is its extended reality nature, which includes a two-phase training and support, is unique and exploratory of the field. Within the 5A project, this thesis attempts to design and implement a VR training experience and AR support to serve people with ASD in the context of a visit to a museum. The visit at a museum is a type of experience where pragmatic tasks like buying a ticket or finding an entrance are equally as important as having a enjoyable and engaging experience inside the exhibit. With the help of a context-aware conversational agent, the deployed solution helps autistic people become more independent through a museum visit. The applications' design and implementation basis were the foundations laid down by the the previous 5A projects, with the goal of having the best solutions for both the Virtual and Augmented environments. The main steps of the process were:

- analysis of the context, its strength, limitations and problems;
- definition of the main kind of experience the user has to go through in a both spatial and social way;
- specification of the contents, features and their style of presentation;
- development and testing of the various functions;
- study and evaluation of the actual benefits for the user based on a study conducted with volunteers from a therapeutic center.

1.2 Thesis structure

This thesis is composed by the following chapters:

- **Chapter 1 - Introduction:** general introduction to the Autism Spectrum Disorder, Virtual Reality, Augmented Reality and the nature of the 5A project.
- **Chapter 2 - State of the Art:** presents an in-depth description of Autism Spectrum Disorder, then it gives an overview of the existing projects that make use of Extended Reality technologies in order to give support to people with Neurodevelopmental Disorders Neuro-Developmental Disorders (NDD). Finally, some projects that use conversational agents are described.
- **Chapter 3 - Needs, Goals and Requirements:** outlines the needs, goals, and requirements that revolve around the users of 5A. Every app of the 5A family shares these.
- **Chapter 4 - Design:** it reports the complete design process of 5A-Museums, first describing the common elements between VR and AR, and then focusing on the respective parts and their peculiarities and differences.

- **Chapter 5 - Implementation:** gives an in-depth analysis of the technical details of the two developed applications, focusing on the technologies and packages used, on the architecture and on the main features and script components.
- **Chapter 6 - Exploratory Study:** reports on the stages of the study conducted to test the effectiveness of the application.
- **Chapter 7 - Conclusions:** summarizes the thesis overall results and presents some critical remarks on the work performed.

Chapter 2

State of the Art

2.1 Autism of Spectrum Disorder (ASD)

As briefly mentioned in the Introduction (see 1), only a minority of individuals with autism spectrum disorder live and work independently in adulthood.

Hus Bal et al. [14] investigated the independent variables for Daily Living Skills DLS attainment in a sample group of ASD patients ranging in age from 2 to 21 years. The study found a decline in the ability to develop DLS during adolescence, particularly for domestic skills such as food cooking, shopping, and household safety, based on interviews and questionnaires administered to both parents and diagnosed individuals. The study further highlights the connection between the fear of the new, unknown and uncertain with an immobility of the subjects both in mental state and physical state. The study concludes that DLS should be a focal point of educational plans for young people with ASD. Consequently, a vast number of adults diagnosed with ASD rely on support from their parents or social services [19].

All of these implications highlight the importance of well-thought training activities in order to successfully teach these subjects. Interestingly, many children with autism are very interested and motivated to use smart devices such as computers, tablets, smartphones [15]. These types of assistive technology devices allow people with autism to interact with their surroundings, make choices, express needs, and interests, and communicate with their friends and parents. Using these devices also increases their sense of self-efficacy, allowing them to gain more and more skills and knowledge while improving their abilities [25]. For these reasons, using technological devices may represent an excellent tool for DLS training of individuals diagnosed with ASD. Devices such as visors, smartphones and smart objects are quite promising at present and progress has already been made in ASD individuals as shown by the statistics. One key feature of this technological devices is that they allow for an experience that is keen to that found in real-life board games and videogames. This effect, obtainable for example by the use of rewards, badges, checkpoints, feedbacks, is more widely defined as "gamification". In fact, it's been observed that people with ASD are more likely to learn or to improve certain behaviors if they have stimuli that clearly gives them a feel of "reward". The response is even stronger if the stimuli are given by their favorite cartoon, film or videogames characters [37].

A study by Goldsmith and LeBlanc highlighted how modulated simulations in which

practice and learning activities can be achieved in a safe environment could have a positive influence on the communication, navigation, and focusing skills of children with autism [4]. The field is open and prepared for further development and better applications for the ASDs, especially if these are implemented by all parties, such as family members, educators and therapists who are involved in the rehabilitation of the autistic people. [25].

2.2 XR technologies for ASD

2.2.1 Virtual Reality for ASD

As stated in the Introduction (see 1), VR is the use of computer modeling and simulation to allow a person to engage with a three-dimensional visual or other sensory environment that is artificially created. The main VR applications' characteristic is their immersive experiences. These experiences generally involve execution of more complex tasks (e.g., riding a motorbike) and more straightforward ones (e.g., walking around a city). To complete these tasks, users can interact with the virtual environment in a more immersive and engaging way. Immersion is, in fact, one of the reasons VR is increasingly used technology for therapeutic purposes [[18]]. Although the term immersion is often used in the restricted meaning of spatial immersion [10], the expression can generally be extended to the whole sensorial experience. Indeed, immersion refers to the capacity of VR experiences of submerging the perceptual system of the user in such a way that they can effectively feel "inside" the digital environment. This plays well with the wide and complex inner world of people affected by ASD and allows the technology to communicate with them in more effective ways [5].

According to a study conducted to investigate the feasibility and effectiveness of VR applications applied to DLS training of individuals diagnosed with ASD:

- multi-modal interfaces create better learning conditions through cognitive integration and limited cognitive load compared to single modality learning interfaces [2];
- technology-based learning has been suggested to be an especially effective learning tool for people affected by ASD [5];
- VR simulation can immerse the learner in a situation where tasks can be performed and repeated in safe surroundings.

Shahira Sharafeldin and HadeelEl Naggar remarked at the Eurasia International Conference in 2018 that virtual environments enhance interactive learning and provide a variety of chances for the student to have influence over the learning process [27]. They also highlight four major benefits of employing virtual technologies:

- VR creates an immersive setting in which people with ASD feel like protagonists, which increases motivation and engagement.
- create a constructive learning approach: Virtual reality experiences allow persons with ASD to study, explore, receive feedback, and engage with a virtual

environment with a modular degree of freedom that is adapted to their specific needs, allowing them to learn more effectively.

- give low-cost yet effective solutions: sophisticated and costly devices are no longer required. Smartphones, tablets, and video-game devices now have simple access to VR/AR thanks to recent technology advancements.
- experiment environments: The level of immersion gained enables for more learning and access to genuine objects.

In order to improve the autonomy of people with ASD in orient themselves inside an airport and ultimately board a flight, in 2020 the Interactive Media Institute of San Diego built a virtual reality experience. The application, called VR-ATT, was developed for *Google Cardboard* and simulates the procedures that passengers must follow in order to book a flight. It's a five-minute virtual reality simulation film that takes players through the check-in process, security procedures, waiting at the departure gate, and boarding. The project was completed using a Google Cardboard headset and the iPhone app. VR-ATT aids the spatial navigation of the users by highlighting important information such as "Look, there's the ticket counter." Let's wait and see what the attendant says." Social stories, which are widely used to teach communication and social interaction strategies to people with autism, were used as the basis for the narrative script [34].

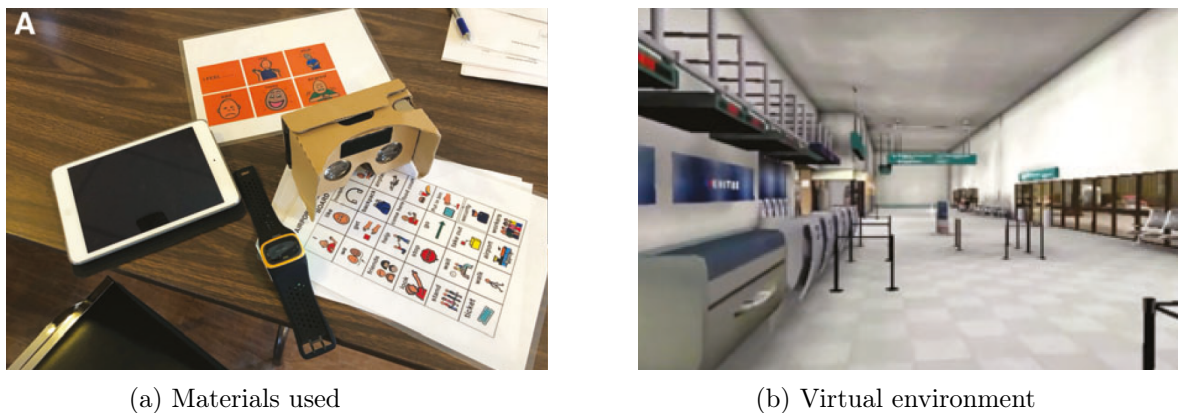


Figure 2.1. VR-ATT

Emil Rosenlund Høeg et al. in Copenhagen presented an experiment aimed to train people with ASD at shopping at a supermarket with the goal of making them more autonomous. The virtual environment was designed and implemented to look like a local supermarket next to the school in which the baseline- and post assessments were conducted. The user could move around in a 1:1 virtual version of the supermarket via walking and VR teleportation using the HTC Vive controllers. The user could look at a shopping list, grab items from the shelves and place them in the virtual shopping basket [19]. The feasibility and effectiveness of the system was investigated in collaboration with a teacher at a school for children and adolescents with mental disorders; a head-mounted display based VR simulation of a supermarket was built and

evaluated with the purpose of developing the shopping skills of students diagnosed with ASD. A preliminary test inside a supermarket without VR training was conducted, then the subjects were asked to do the same tasks inside the Virtual Environment. The aim of the repeated usage of the system was to allow the transfer of adaptive shopping skills from VR to real shopping experience. This was achieved through the repetitions of product search processes in the safety of the simulated environment, and stripping the environment from additional obstructive and intimidating elements, allowing the participants to focus on the improvement of their skills. Days later, the subjects were brought again in the real supermarket to test the effectiveness of the training. The results were mixed, because during the final test the supermarket was much more crowded than the first time. In particular, the authors make the hypothesis that the treatment group's negative development of self reported sense of confidence during the post-treatment assessment could be due to the lack of constant training multiple times on the VR simulation and the environmental discrepancies between the real supermarket and its VR version, as the VR supermarket simulation did not include any other human shoppers. This underlines the importance of having a high fidelity model and also humans and crowds, which can be considered “distractors” that are present in real life experiences. Moreover, there were also positive results, as many participants found the product they were searching for by themselves. Also, one of the participants who usually was uncomfortable with shopping scenarios and who refused to enter the supermarket alone during the pre-assessment accomplished the post-treatment shopping assessment with perfect results.



(a) Inside the virtual environment



(b) User interacting with the environment

Figure 2.2. Project from Emil Rosenlund Høeg et al.

A study conducted from the University of the West of England of Bristol by Nigel Newbutt has experimented a VR experience of a museum on 11 autistic subjects between 10 and 14 years [47]. The aim of the testing was to make the users feel comfortable in the museum environment, training them for a real museum visit. The process was divided in 5 steps:

- Answering questions related to their experiences of visiting museums/cultural spaces.
- They had the chance to experience a museum tour using a VR 360-degree app (using *Google Cardboard*).

- Completed a questionnaire about the VR experience.
- 4 days later visited the same museum in real life.
- The next day completed a post museum visit questionnaire related to their experiences.

The key elements that were discovered through this study were:

- Using the VR app helped to calm and relax the children.
- Using the VR app seemed to help the children know what to expect before they visited the real world museum.
- The VR app might have helped to alleviate issues of confusion when visiting in real life.
- The VR app was reported as helping the children to visualize the space before visiting in real life.
- Reports of museums being distracting were similar before and after the visit.

Other two projects that we took in account during the study and the development of 5A Museums were the already implemented 5A projects: 5A-Subway and 5A-Supermarket. In 5A-Subway users are guided through the virtual reconstruction of the Milan subway. They have to find the subway road signs, enter the underground, buy a ticket and take the subway. During the experience, users are followed by a robot assistant that helps them in the tasks execution [35].

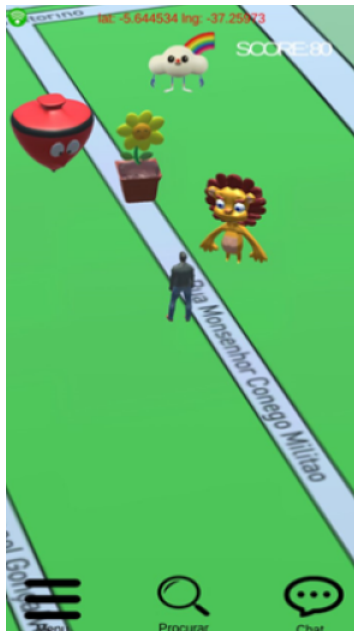
5A-Supermarket is a VR experience, developed for mobile smartphones and Oculus devices, where users have to go grocery shopping in a simulated supermarket. The application guide the users to all the steps that have to be followed when going to the supermarket, from taking the shopping cart to the payment to the cashier. Even in this case, the user is followed and helped by the same robot assistant [36]

Unfortunately 5A-Supermarket, due to Covid-19 related problem, has not been tested on final users. Concerning the testing of 5A-Subway these are the main results that emerged:

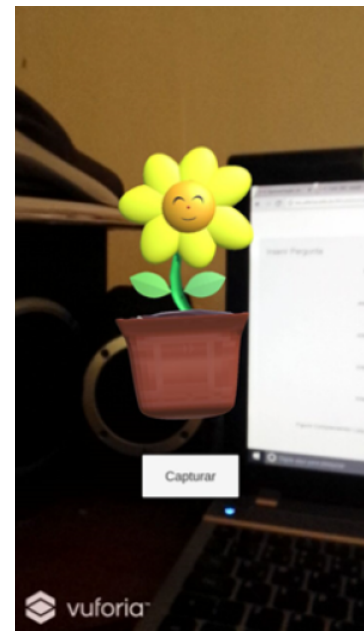
- users extremely appreciate the experience and express their will of having more activities similar to the one proposed;
- 5A-Subway has been considered useful to its purpose;
- the conversational agent is helpful and contributes to support users end the experience.

2.2.2 Augmented Reality for ASD

AR, on the other hand, is also emerging as a promising technology to help those with ASD understand the world more fully by bridging the physical and digital worlds [33]. About the impact of AR on ASD, research shows its effects to be positive in a variety of situations. AR can encourage play and increase motivation, attention, and the learning of new tasks. AR-based apps are used to enhance engagement, motivation, and learning for people with ASD [33]. AR overlays, like 3D videos, figures, and information, can be added to any real object and multiple studies have shown that these AR experiences result in increased engagement, enjoyment, motivation, and attention. One particular study designed to teach object discrimination revealed a 62% increase in on-task participation and happier, more determined students [13]. Some studies even suggest that some child can prefer AR-based apps to VR, as they are less overwhelming and have a softer approach by mixing the virtual world with the real one [29]. Another study was executed for children where a five-step picture schedule of a toothbrushing task was augmented with a video clip demonstrating all the steps. After treatment, the AR intervention showed an immediate increase in independent performance by the participants, and after two months, completion of the task remained high [17]. The kids enjoyed the AR content, and the teachers reported the app to be an effective teaching tool. This systematic literature review shows that the development and evaluation of systems and applications for users with ASD is very promising. The use of technological advancements such as virtual agents, artificial intelligence, virtual reality, and augmented reality undoubtedly provides a comfortable environment that promotes constant learning for people with ASD [31].



(a) User navigating through the environment



(b) Display of the AR object

Figure 2.3. Knowledgemon Hunter

KNOWLEDGEMON HUNTER, developed by a team of students at the Universidade Federal Rural do Semi-Árido in Mossoró, Brasil, is a serious augmented reality game developed for smartphones inspired by the popular app Pokemon Go, designed to support learning and minimize the autistic children's social isolation. It encourages children to explore the geographic space in their proximity, demanding the user to search fictional characters [24]. The main interface features a 3D avatar located in a map mirroring the real one corresponding to the user's position. The device captures the user's movement and coordinates through the GPS system. To process the coordinates, MapBox API was used. When the user moves in the real environment and approaches a virtual character geographical position, it appears in the main interface as an AR object and can be captured. The characters' shape, color, movement, and sound are based on animals present in assisted therapies, as they generally generate a particular empathy in children. When a child captures a character, it can access the Learning Object related to that character and, through interactions with the game menus, view the information about it. The application saves the Learning Object so that the user can access the information at any time. After seeing the content, the user can answer questions related to it in a sort of minigame that tests the user's cognitive abilities. When the answer is correct, a congratulations image will display, and the algorithm adds points to his score. Otherwise, an incentive message is displayed to it to try again.



Figure 2.4. Minigame inside ARTis

The archaeological park context is explored in ARTis, an APP by Garage94 and Naepolisanit, that allows seeing virtual content overlapped with the surrounding reality through a smart device's camera, creating a path inside a museum or park to help the user orientate and experience an appealing and interactive visit [38]. Interestingly, the APP is created with the fact that the user, a child/teen in this case, is accompanied by a tutor during the whole visit. The tutor can select the difficulty level for the user to tweak the help provided and the level of complexity of the several activities offered. ARTis allows users to interact with this kind of content thanks to the presence of a virtual guide (a friendly avatar) who will accompany and help the user throughout the visit to the museum and will allow a better and safer exploration by providing information and support in case of need. Particularly interesting here are the activities proposed in the form of minigames, such as virtually breaking down

the artworks as if they were pieces of a puzzle or adding graphic elements to a statue to allow interaction with it or simulating the restoration of a painting, thus creating interest in the user about the artwork.



Figure 2.5. Mosoco helping children with autism

Lizbeth Escobedo et al. developed the Mobile Social Compass titled MOSOCO, a mobile assisting application that uses augmented reality and the visual supports of a validated curriculum, the Social Compass, to help children with autism practise social skills in real-life situations. They presented the results of a seven-week deployment study of MOSOCO in a public school in Southern California with both autistic students and neurotypical (NT) students [8]. MOSOCO provided autistic children with interactive features to encourage them to make eye contact, maintain appropriate spatial boundaries, reply on conversational initiators, share interests with partners, disengage appropriately at the end of an interaction and identify potential communication partners. MOSOCO runs on an Android smartphone which is wirelessly connected to a server. It uses the smartphone's camera to augment a real-life social situation with visual support mimicking the Social Compass curriculum which augments social stories using paper-based visual supports. The results of the study demonstrated that MOSOCO facilitated practicing and learning social skills, increasing both quantity and quality of social interactions, reducing social and behavioural mistakes, and enabling the integration of children with autism in social groups of neurotypical children.

2.3 Conversational Agents

Conversational agents are being employed as context-aware guides, a point of reference and a supporter. This concept has been used in extended reality apps to help end-users in managing anxiety and frustration while doing daily activities. The majority of Conversational Agents for children with disabilities are focused on improving language skills and the identification on body language, so their end goals differ from 5A's. Nevertheless, a comparison was useful to understand better which features would be implemented in the app and how it compares. The analysis cover examples of three Agents for people with disabilities and one for non-impaired people. In the end, a comparison table is presented.

ECHOES is a serious game developed by The Department of Informatics of King's College of London to help young children with Autism Spectrum to practice and

acquire social communication skills. In ECHOES, children interact with an intelligent virtual character in the context of social situations through a multitouch LCD display (not VR) with eye-gaze tracking [12]. All ECHOES activities take place in a sensory garden with an intelligent virtual character and interactive items that react to the user touch gestures. For example, touching the petals of a flower transforms it in a bubble. The relationship of action-reaction implemented through touch promotes children's understanding of cause and effect. Since it is not obvious to the children from the outset how the different objects will behave and react to their gestures, the sensory garden fosters children's imaginative play and exploratory behavior, which are rarely observed in individuals with ASD. The conversational agent is a human companion who acts as a friend and instructor. It is meant to help end-users to expand and enhance their expressivity and communication system, as well as helping them understand not only verbal language, but also non-verbal behaviors such as body language and expressions.



Figure 2.6. Child interacting with Echoes

Another good example, also using avatars on standard computer screens, is FaceSay: a computer program that contains three different games with realistic avatars designed to teach children with ASD specific social skills [7]. The avatars were animated photos of real persons that could interact with the children by drawing on a pre-programmed knowledge base. The overall goal of the games was to promote awareness of the movements and features of the face, particularly the area around the eyes. Interactive features of the software provided opportunities for children to respond to social situations. Targeted social skills included teaching specific social skills for responding to joint attention, face puzzle, and matching and manipulating facial expressions. Results of studies conducted demonstrated that when students with ASD were provided with an opportunity to learn and practice specific social skills in FaceSay's controlled environment that simulated a natural setting, their social interactions improved also in natural environments.

The next example of conversational agent is demonstrated by the app Bob Fishing Shop, built by a group of researchers of Chapman University [30]. Bob Fishing Shop is a virtual reality simulation aimed to assist persons with ASD in practicing social



Figure 2.7. FaceSay experience

interactions and language. The purpose of Bob Fishing Shop, which was created for the Oculus Rift VR headset, is to improve social and conversational etiquette while keeping end-users interested in a secure and friendly setting. The environment was specifically designed to be similar to cartoon animations to have characters and gestures familiar to the users. A baseline script that maps an example dialogue between the user and Bob is used to retrieve the interaction with Bob. Repetitions were avoided by having a few versions of the same sentences. To give the human conversational agent additional personality, a professional voice artist recorded the script phrases. A brief research with two users with ASD found that the VR headgear provided a favorable experience for the users and that minimal training was necessary to utilize it.



Figure 2.8. Bob Fishing Shop conversational agent

As a last example, INDOAR is an app for everyone that allows user to navigate through either short or complex routes displaying additional information or offers on the way to their destination. It can be used for small locations as well as complex and large buildings without having any hardware installations on site. Routes are defined either through a browser-based admin interface or created directly within the app. A

variety of tracking technologies is available to suit any location scenario. The system offers a standard arrow navigation or an animated digital guide (GuideBOT) which can be replaced with a customized character. The template can be extended with AI chatbot features and supports both text and voice input. It can record speech and convert it to text by using Google Dialog Flow or SAP Conversational AI to interpret voice/speech input and answering accordingly. INDOAR is interesting also for its use of indoor navigation technologies. The app on request of the client buying it can work on different tracking technologies. A further analysis on tracking technologies is reported in Chapter 4.

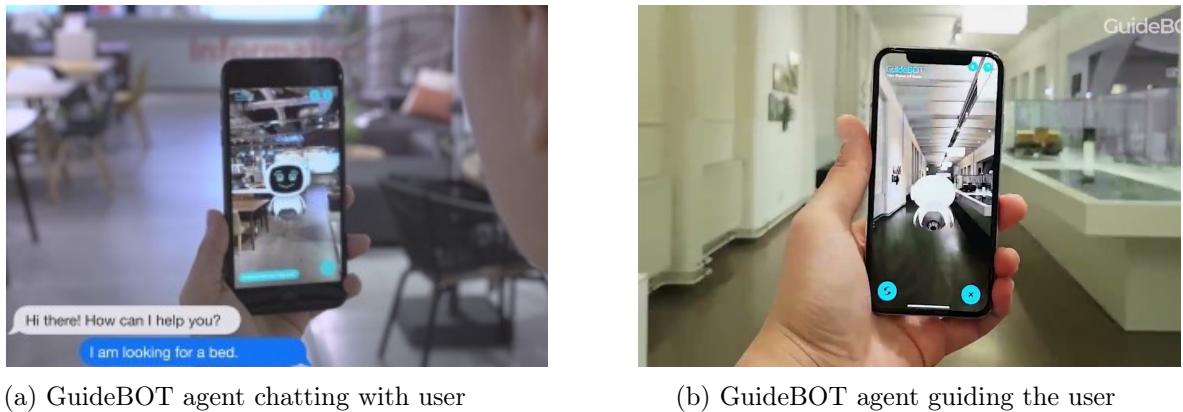


Figure 2.9. INDOAR assistant

Project	Conversational Agents				
	5A	ECHOES	FaceSay	Bob Fishing Shop	GuideBot
Character	Robot	Human (cartoon)	Human (realistic)	Human (cartoon)	Robot
Is a real-life prop	No	No	No	No	No
Context-aware	Yes	No	No	No	Yes
Helps users in orientation	Yes	No	No	No	Yes
Helps users in understanding expressions and body language	No	Yes	Yes	Yes	No
Guides users in daily tasks	Yes	No	No	No	Yes

Table 2.1. Comparison of the features of several conversational agents

2.4 Hardwares for Virtual Reality

Oculus devices are a relatively new technologies that are currently ever-growing. Although firsts Virtual Reality hardware devices were developed in the early 1990s, these technology began to spread only by the beginning of the 2010s. One of the first prototypes to be released was the Sony Glasstron, released in 1996. The products included two LCD screens and two earphones for video and audio respectively.

Concerning more modern devices, the first real VR headset was the Oculus Rift released in 2016 and developed by the company Oculus VR. The headset was composed of the visor and two hand controller with several movement sensors. It had a Pentile AMOLED screen with a resolution of 2160×1200 , with a 110° field of view. Since it was not equipped with internal CPU and GPU, it had to be constantly linked to a computer with the minimum requirement of Windows 7 OS.

In 2016 another VR headset have been released: the HTC Vive. It had similar specifications and became a competitor to Oculus. Another competitor device that came out in the same year was the Playstation VR, that had cheaper price and performances. The first standalone VR device was the Oculus Go released in 2018. It runs on Qualcomm Snapdragon 821 with an Adreno 530 GPU. The novelty of this product was the possibility to use the device both as a all-in-one device or connecting it to a Windows device to exploit a higher quality hardware.

In 2019 the successor of Oculus Go was released: the first Oculus Quest. Compared to its predecessor it has a screen with better resolution and an higher frequency update. Moreover it offers the possibility to adjust lenses and supports positional tracking. Finally, it has better processor components, with a Qualcomm Snapdragon 835 CPU, 4GB of RAM and an Adreno 540 GPU.

Recently, the successor, the Oculus Quest 2, made its debut on the market with slightly better system specifications.



Figure 2.10. Oculus Quest 2

Chapter 3

Needs, Goals and Requirements

The analysis of the needs and subsequent goals and requirements here reported has been identified through research on previous 5A projects, like 5A-Subway (1st and 2nd version) and 5A-Supermarket. Clinical and therapeutical knowledge was kindly offered to us by the advice of neuro-psychiatrists, psychologist and therapists from two institute that collaborated with I3Lab, the research and rehabilitation center for developmental age *IRCCS Medea* and the Milan autism center *Fondazione Sacra Famiglia*. This chapter looks into the needs, goals, and requirements that were used in order to successfully designed 5A-Museums. The main target group were people affected by Autism Spectrum Disorder (ASD) who are also the main stakeholders of the project. There were also two other minor stakeholders:

- tutors and therapists that accompany this people in their everyday life;
- museum that are willing to offer the service of the application inside their exhibits

3.1 Needs

- **Fear of open and public spaces:** they need to be helped in dealing with public spaces that are often plenty of distractions, noises and other external stimuli, which can cause exaggerated reactions or anxiety.
- **Orientation difficulties:** they need guidance to orientate in complex public spaces.
- **Repetitive behaviours and restricted interests:** this is one of the most common trait of people affected by autism. They need something to interest them and keep high their attention as they tend to have few interests and quickly lose attention in arguments they are not interested in.
- **Persistent deficits in social interaction:** people affected by autism need to be guided in social interaction with other people, since they often have difficulties in starting a conversation.

- **Different behaviours and capabilities:** adapt to different people that have different levels of capability, or one person that changes their level through repetition and learning.

Regarding the tutors, their main needs is to have the possibility to help their patients in the most efficient and easy way possible. Finally, concerning the museums, there are two main needs that we took in account:

- freedom to change paths and indoor disposition of items and rooms
- improve the engagement inside their exhibits

3.2 Goals

Based on the previous stakeholders' needs, these are the main goals of the two applications:

- **G1:** the initial goal is to assist end-users in overcoming their spatial orienting issues by finding the proper path and landmarks. This goal can be accomplished by demonstrating to them the proper path to take.
- **G2:** help controlling anxiety.
- **G3:** familiarize with public spaces: people affected by ASD can feel very uncomfortable in public spaces. The applications try to make users aware of spaces and procedures they will do in advance, and assist them during the real-life experience.
- **G4:** draw the user attention: as we already said, people affected by ASD have restricted interests, so the application should be immersive and engaging.
- **G5:** easy setup of the applications.
- **G6:** applications can adapt to different levels of capability.
- **G7:** applications are easy to adapt to sudden changes: museums structure are not always the same and can change very much during time. Moreover the application should be adaptable to more types of museum.
- **G8:** applications improve the user's willingness to interact with the exhibit and items.

3.3 Requirements

The requirements for the two applications had to be split because of the approaches to the different technologies.

VR Goals and Requirements Matrix								
	G1	G2	G3	G4	G5	G6	G7	G8
R1	X		X					
R2	X		X					
R3		X						
R4		X		X				
R5		X		X				
R6		X		X				
R7			X					X
R8						X		
R9							X	
R10						X	X	
R11					X			

Table 3.1. VR Goals and Requirements matrix

3.3.1 VR Requirements

- **R1:** application has to implement a true-to-life experience.
- **R2:** application has to guide and assist the user through the virtual environment to guarantee spatial orientation.
- **R3:** application has to assist the user during the execution of tasks.
- **R4:** by incorporating a context-aware conversational agent, the application must be able to handle the possibility of anxiety produced by the presence of too many stimuli or the fact that the user is unsure on what to do.
- **R5:** by incorporating a context-aware conversational agent application must be able to recognize the user's vocal communication and to responds verbally.
- **R6:** application must be able to communicate not only by text and speech, but also in a visual way with images.
- **R7:** the application must allow for a certain amount of exploration of the area.
- **R8:** the application must have customizable difficulty according to the user specific needs and abilities.
- **R9:** the whole experience must be modular so that the user can start the experience wherever they want.
- **R10:** the whole experience must be modular to improve scalability to different museums.
- **R11:** application must be easy to use and to setup.

AR Goals and Requirements Matrix								
	G1	G2	G3	G4	G5	G6	G7	G8
R1	X	X	X					
R2	X							
R3	X							
R4		X	X					
R5		X		X				
R6		X		X				
R7				X				
R8						X		
R9							X	
R10						X	X	
R11					X			

Table 3.2. AR Goals and Requirements matrix

3.3.2 AR Requirements

- **R1:** application must be context-aware of its surroundings at every moment.
- **R2:** application must provide spatial orientation in indoor environment to the user through the tracking of the device using QR codes.
- **R3:** application must provide spatial orientation in outdoor environment to the user through the tracking of the device using GPS.
- **R4:** application has to suggest the next task and then assist the user.
- **R5:** by incorporating a context-aware conversational agent, the application must be able to handle the possibility of anxiety produced by the presence of too many stimuli or the fact that the user is unsure on what to do.
- **R6:** by incorporating a context-aware conversational agent application must be able to recognize the user's vocal communication and to responds verbally.
- **R7:** the application must have a simple and clear UI.
- **R8:** the application must have customizable difficulty according to the user specific needs and abilities.
- **R9:** the whole experience must be modular so that the user can start the experience wherever they want.
- **R10:** the whole experience must be modular to improve scalability to different museums.
- **R11:** application must be easy to use and to setup.

Chapter 4

Design

The target audience was the primary focus throughout this phase, thus all decisions were made with the idea of the applications being used by persons with NDD and their therapists. First of all, the context of use is a museum, intended in a general sense as all kinds of museums. That means by definition “a building or institution that cares for and displays a collection of artifacts and other objects of artistic, cultural, historical, or scientific importance, available for public viewing through exhibits that may be permanent or temporary” [20]. Moreover, the context was always assumed to be an indoor exhibit, because open air museums like archaeological parks have too many differences with indoor exhibits and should be treated alone as different experiences through and throughout. As said, the context is assumed as in the other 5A family apps as a general context; in other words, both the applications should have the same effectiveness in different museums with different inside exhibitions. Differently from the other apps of the 5A family though, that work with more standardized context like supermarkets and metro stations, museums rarely share similar layouts, practices and types of exhibits.

- Practices: museum can have different kinds of practice during the entrance, the purchasing of tickets, the checking of tickets, the price and promotion of tickets;
- Layout: the internal layouts of museums vary greatly, they can be on one or multiple floor, have a regular or a complex pattern inside rooms, have a regular or complex itinerary between rooms;
- Types of exhibits: all kinds of museums exist, ranging from natural history museums, to historical museums, to art museums, to science museums.

All these different traits had to be simplified in a straightforward and generalized experience. Of course, developing an augmented reality application meant that finding a suitable real-life museum was a mandatory prerequisite to develop the application.

The right kind of museum was found in the *Museo di Storia Naturale di Milano* (Natural History Museum of Milan), which provided a perfect case study for the application. The reasons in favor of this choice were mainly two. First, the fact that this type of museum is more engaging in the interaction that the user can have with the exhibit, having a collection of three-dimensional mounted skeletons and reconstruction of wild animals that can be explored in 360°. As previously said, the

selected users have difficulties in focusing and maintaining attention, so a natural history museum was deemed as more keen to offer a gamified experience. Secondly, art and other kinds of museums sometimes prohibit the use of smartphone cameras to take photos or videos for privacy and/or reproduction reasons, while the *Museo di Storia Naturale di Milano* allows its use without problems. Therefore, it was easier and faster to have permissions to use cameras and document the interior exhibits.

The museum was contacted to be informed about the project, and the first exploratory survey was made along the itinerary. A general map and scheme of the interiors, taking into account the location of the key areas, was conceived. The key element of this scheme is that, while born from and applied to the specific case of the *Museo di Storia Naturale di Milano*, it had to be generalized enough to be usable with minimal changes inside every kind of museum. Two primary areas with different main needs were identified: the outside proximity and entrance hall, with more pragmatic needs, and the inside of the exhibit rooms, with more focus on the engagement needs of the user. Pragmatic needs are found throughout the whole experience, but mainly at the start, where the user must find the entrance, enter in the hall, locate the ticket booth, buy a ticket, check the ticket and find the entrance to the exhibit rooms. Once inside the exhibit rooms, the needs of this kind are how to go forward or backward through the whole itinerary, and eventually know who to ask for more information about the exhibit and/or the location of particular zones like the restroom or coffee shop. In this area, however, attention is paid also to the engagement needs of the user.

- A. ENTRANCE
- B. (TICKET BOOTH)
- C. EXHIBIT ENTRANCE
- D. EXHIBIT ROOMS
- E. EXIT

The Entrance area (A) is composed of entrance doors, usually of a medium to large size, that in some cases face a street or in other cases a park or enclosed private outdoor area. A sign or label indicating the name of the museum may or may not be present. Generally, the Hall is the place where the Ticket booth (B) is located, and presents one or more ticket clerks designated to sell the tickets, although in some cases it is possible that the booth is absent. The Ticket check (C) area can be located inside the Hall or be at the end of a path from the hall to the Exhibit rooms. The checking of tickets can be automatic via an automatic machine, or manual with a designated worker of the museum staff. The Exhibit Entrance (C) is located right after the Ticket Check area. The Exhibit rooms (E) are where the main exhibition of art, historical objects, animals or other is located. Usually, it's composed of rooms connected together by doors. Finally, the Exit (F) can be located at the other side of the entrance or be the same door through which people enter.

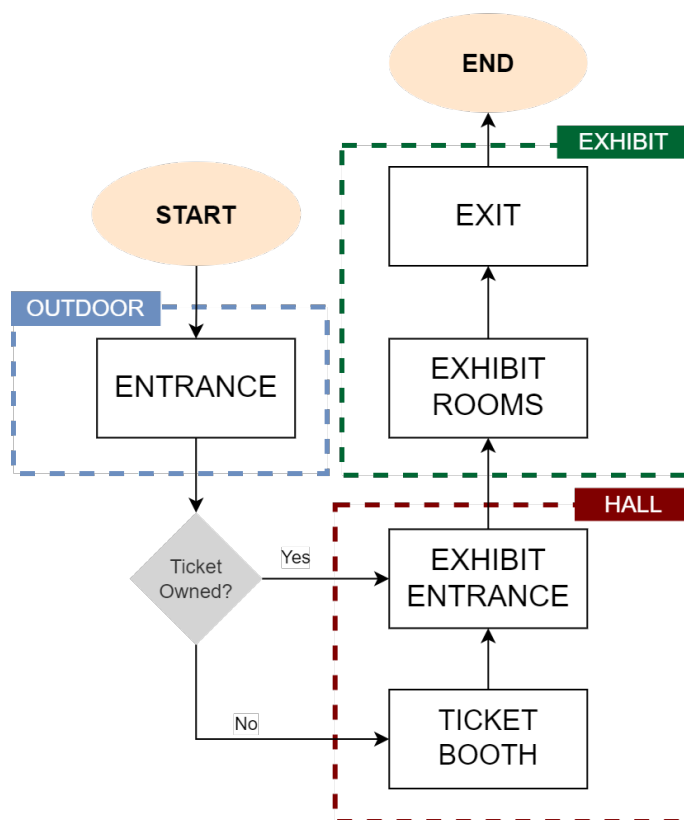


Figure 4.1. Key areas of a generic museum

4.1 Tasks

The identification of the main tasks along the journey started from the key areas. It was first determined the specific cases found in our general key areas. The Museum is located inside a neo gothic 19th century building found inside the Indro Montanelli Park near Porta Venezia. The museum is circled by a small private courtyard surrounded by fencing. The Entrance is located right in front of the main gate, at a distance of approximately 30 meters. The entrance doors are five medium sized doors located at the top of a small staircase (see 4.2a). Once entered, people find themselves in the Hall where the Ticket Booth is located. The booth is a square enclosure with glass on three sides and a sign indicating his function on top (see 4.2b). Based on the affluence of people, multiple ticket clerks at different sides can be present at the same time, however normally the front desk is used. To enter the rooms with the exhibit, people have to go left to the first door (see 4.2c). There, a member of the staff checks the tickets and allows people to enter. The first room is actually a store of goods and merchandise, so the actual Entrance to the exhibit is located at the end of this room. The rooms are labeled with a number and a descriptive title of the contents (see 4.2d). The itinerary explores the various fields of natural history, including archaeology, paleontology, extinction and human evolution, as well as having a dedicated first floor to dioramas of currently living organisms (see 4.2e). The itinerary ends by exiting the exhibit from the same doors of the Entrance. A schematic of the key areas of the museum was created, taking into account the actual disposition and floor plan of the building (see 4.3)

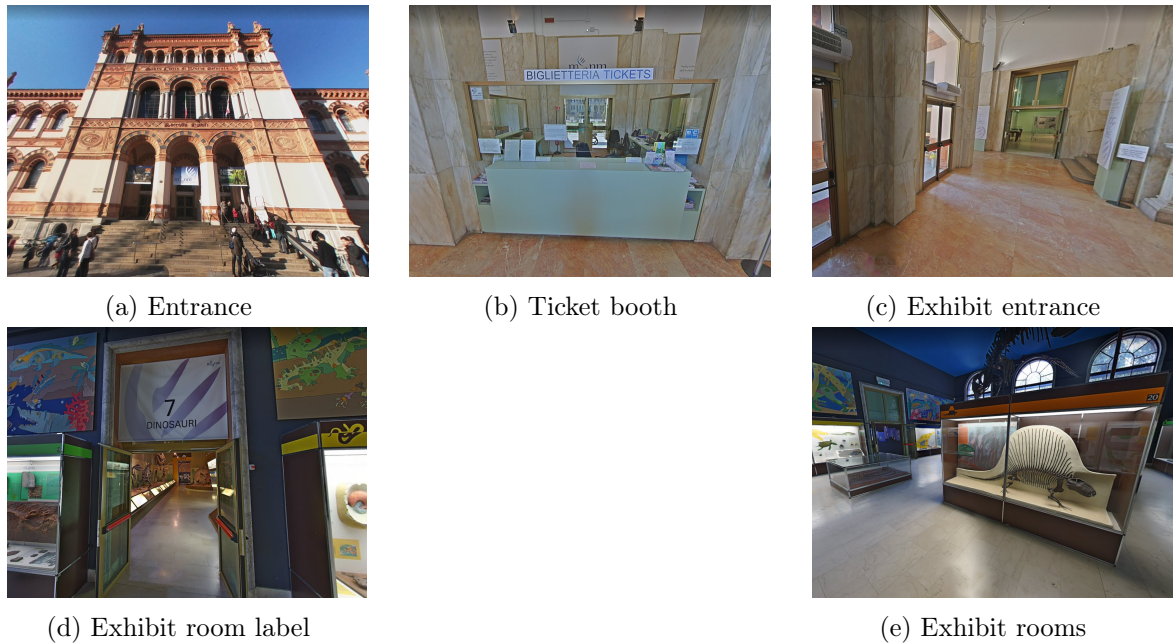
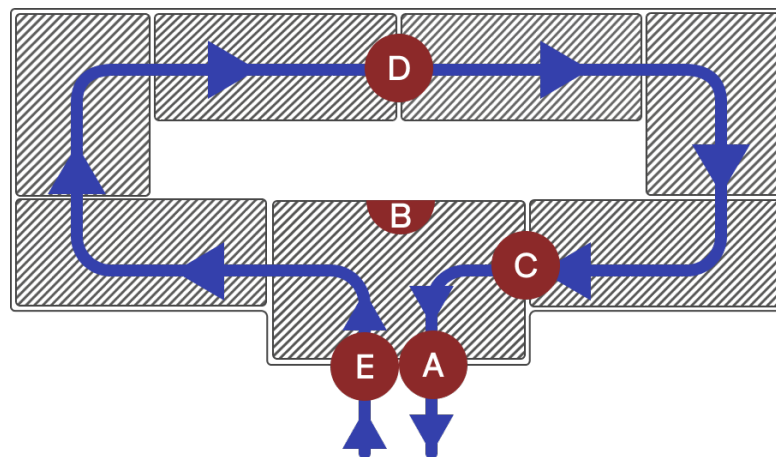


Figure 4.2. Pictures of the key areas

Figure 4.3. Key areas inside the *Museo di Storia Naturale*

From this, a typical user itinerary was created. Along the way, several tasks were identified that a person would incur inevitably into while navigating through the environment with the goals of moving forward and exploring the exhibits. Nevertheless, it was deemed permissible to make assumptions to streamline and smooth the experience, as well as simplifying it as it was thought that taking into account all the possible cases and variations would be too overwhelming and not appropriate for a first prototype of this application. Assumptions were:

- A1: User will start the experience at the gate of the museum
- A2: User will need to buy a single ticket
- A3: The ticket will be checked by a member of the museum staff
- A4: The experience will end when the user exits the building

The overall experience can be divided in the following main tasks:

- OUTDOOR
 - **Find the entrance to the museum:** the person has to find and recognize the correct entrance to the museum.
- HALL
 - **Enter inside the Hall:** enters the hall.
 - **Find ticket booth:** has to locate the ticket booth. If there are many people, they will have to get in line and wait; otherwise, they will go straight for the desk.
 - **Get in line and wait:** must recognize a queue and respect their turn and other people's needs.
 - **Speak with the ticket clerk:** must talk and communicate with the ticket clerk. They have to know what to ask and what to expect from the person behind the desk.
 - **Buy a ticket:** the user pays for a ticket and receives it as a small paper piece.
 - **Find Exhibit entrance:** the user has to locate the Exhibit entrance where they will find also the person dedicated to the checking of tickets.
 - **Check the ticket:** give the ticket to the museum staff that checks them, then is allowed to go through and enter the Exhibit rooms
- EXHIBIT
 - **Enter exhibit rooms:** find the entrance to the Room 1, then to the subsequent Room 2 and so on.
 - **Explore and learn:** once inside the Exhibit, explore, learn, read information panels and be curious about what is around them.

Figure 4.4 reports the UX flow diagram that sums up the main key areas and tasks. Following, all the possible helps to the user in the application were determined based on the specific needs of the various areas.

- OUTDOOR
 - User will need help in finding the way
- HALL
 - User will need help in knowing where to go next

- User will need help in finding the way
 - User will need help in knowing who to ask
 - User will need help in knowing what to ask
 - User will need help in managing movement through crowds
- EXHIBIT
 - User will need help in knowing where to go next
 - User will need help in finding the way
 - User will need help in managing movement through crowds

It is important to note that the user must always be free to start the experience in any of this moments and areas. This allows for a more focused training in the areas that the user finds more difficult, as well as resulting less boring to use. In VR, this means that they must have the possibility to choose the area were to start the application, and the problem can be fairly addressed by having a initial menu that allows them to choose the area. In AR, the approach is different. The user could theoretically open the application outside the museum and use the application in all its potential and be guided for all the experience inside the museum. However, they could also open the application while being inside the hall or near the entrance to the exhibit or even inside the exhibit. To address this problem, several possibilities were evaluated and ultimately the QR code method was chosen as the most reliable, cheap and accurate. You can see an evaluation and comparison in chapter 2. Therefore, a system of QR codes was located in key areas. This QR codes, once scanned, communicate to the device its actual position at that moment in time, allowing the app to respond with context-related responses.

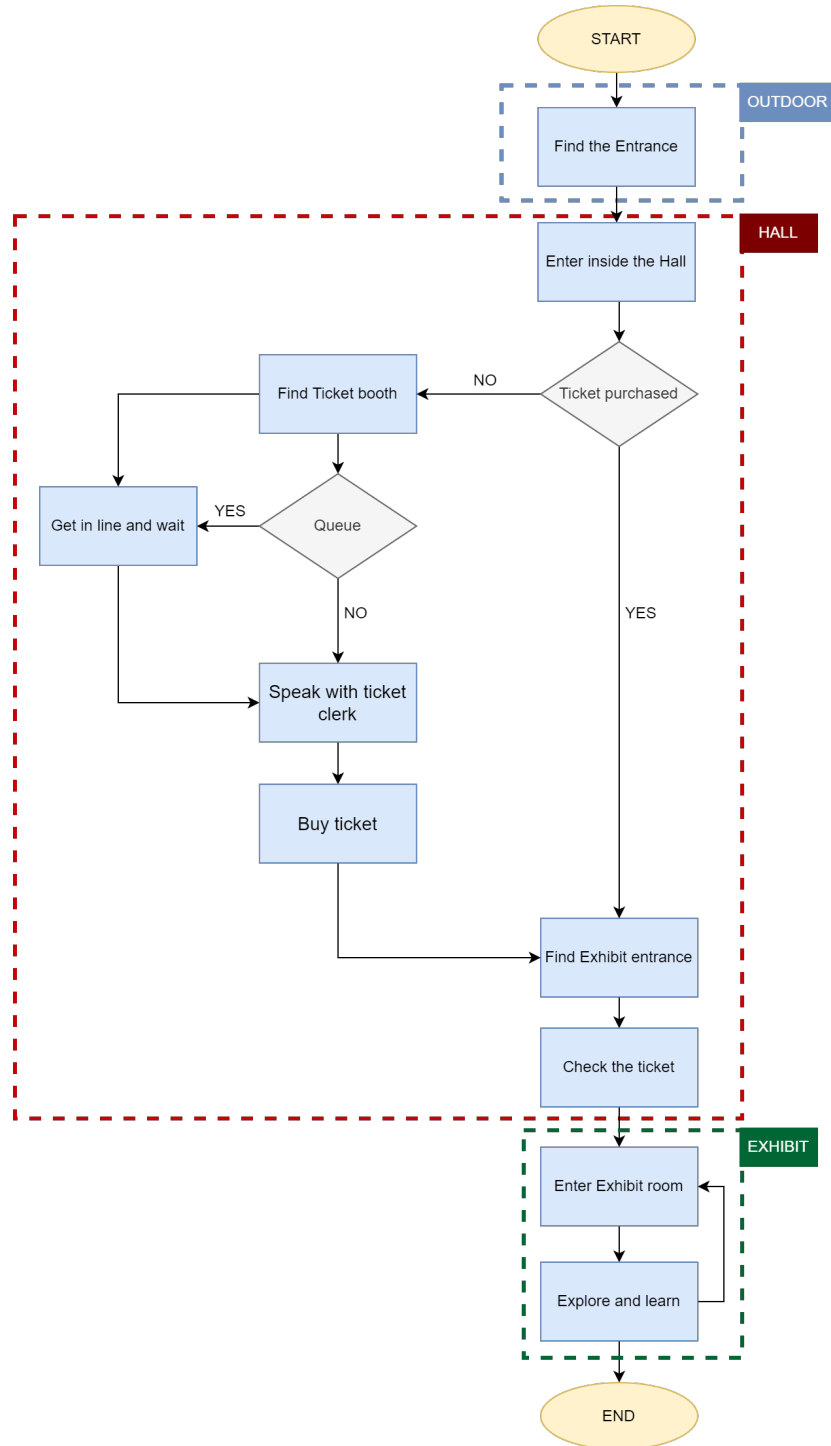


Figure 4.4. UX flow diagram

4.2 Transmedial elements

By "Transmedial Elements" are intended all those elements that can be found in both the VR and the AR applications. The presence of common elements is particularly relevant because of the target audience. As we previously said, the VR is mainly used as a training system to help users to familiarize with a public environment. To do this

several graphical elements are used. Then, once the user has understood the function of these elements, they can proceed to the AR experience, where they will find the same tools supporting them in a real public environment. Because people affected by ASD have closed and rigid ways of thinking, they find difficult to abstract on a conceptual level, and new experiences could bring them to a state of uncertainty and anxiety (see 2.1). Therefore, finding again familiar elements and procedures reassures them: they find themselves in front of environments, characters, interactions not totally new for them and so they know what to do. Here is presented an examination of this elements on a conceptual level shared by both applications. Their declination in two different mediums means that the specifics of visual presentation and interaction are slightly different between the two application, and as such this detail are discussed in sections 4.3 and 4.4.

4.2.1 Virtual Assistant

The main transmedial element developed to achieve this continuity was the use a virtual character common to the two experiences to create a guide element for the user, which could act as the Conversational agent and give indications, feedback, encouragement, and make them feel accompanied through the experience. It will guide the user first in the VR experience, and then provide help and assistance during the experience in the real world with Augmented Reality. It always follows the user in every place they go, taking the role of the user's own tutor, staying with him all the time. This was obtained in two different ways in VR and AR, and is explained throughout in sections 4.3 and 4.4. This requirements materialized in an avatar called Virtual Assistant VA, its functions to:

- make a seamless transition between the two experiences, offering a comparable visual and auditory stimulus that the user can interpret as the same steps in each one of the experiences;
- guarantee a greater human connection between the app and the user, stimulating the user to do better and maintain concentration;
- offering help and tips when the user is struggling in going forward, finding the correct path and interacting with items;
- give some clear payoff when a task is completed successfully;
- offer an engaging and fun visual character to present a less “serious” app that keeps a certain likeability during its use.

Design

To develop a convincing avatar, some key principles were used. It had to be simple, easily recognizable, easy to adapt to other apps for different experiences, easily translatable and iconic both in 2D and 3D. People with ASD have more difficulties in interpreting human language (verbal and non-verbal) [9]. These difficulties create a preference for non-human characters that don't present subtle changes in expressions that could cause frustration in the user [32]. The options were either to create an

animal character or go for a robot character. By going with an animal character, the dog would have been the chosen species, as it is a common pet in our houses and so easily recognizable by everyone. Moreover, it is a common thought that dogs are loyal, trustable and good friends of humans. The other option, the robot character, was chosen because it retained the main traits of high fidelity and trustability of the dog, adding also a sense of machine-like precision in delivering information. This value is highly important for people with ASD that often prefer to relate with screens and computers that give precise instructions than with people that could give vague statements [15]. Another fact made the robot character a preferable option: the simplicity of the forms that could be achieved, allowing for a smaller poly count, and the clean and brilliant look of metallic textures that are more easily handled by rendering engines (if compared to skin/hair of an animal).

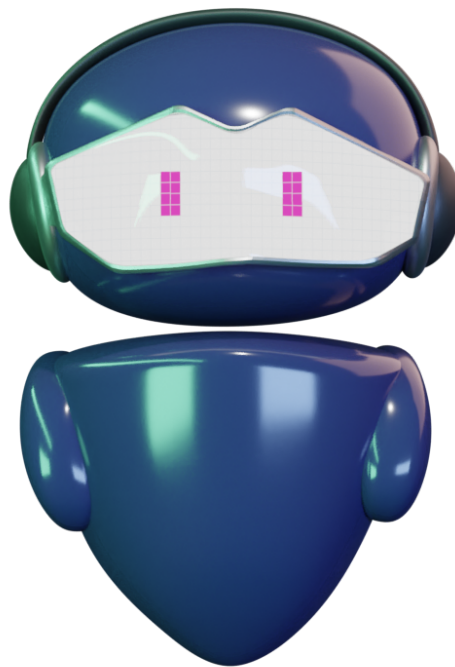


Figure 4.5. Virtual Assistant 3D model

In order to craft a good shape for the VA inspirations were taken from other conversational agents, as well as from animated movies and videogames. Figure 4.6 shows the analysis of design features in other popular animated characters. The robot as a mascot or companion is very popular in this mediums so it was easy to find a lot of different examples. In the end, we chose a simple body of oval shape. Limbs and appendages were reduced to a minimum, considering that they would have been difficult to animate and not very useful in communicating with the user, creating pointless extra visual noise on the screen. So, the arms were reduced to two stumps and the whole body to a floating oval. The head was to be the most important part, as it would have been the only mean to convey come character in the VA. The head is a floating oval and is treated as a visual continuation of the body. The most distinguishing features are the headphones and the screen board. Together they form

a sort of “headset” that is mimicking a VR headset, symbolic of the fact that the robot is communicating from the virtual world to the real world at the same time as the user, who is doing the exact opposite. Figure 4.7 shows the sketch of the volumes of the character; it was then translated into a three dimensional model made with the software Blender (figure 4.8).

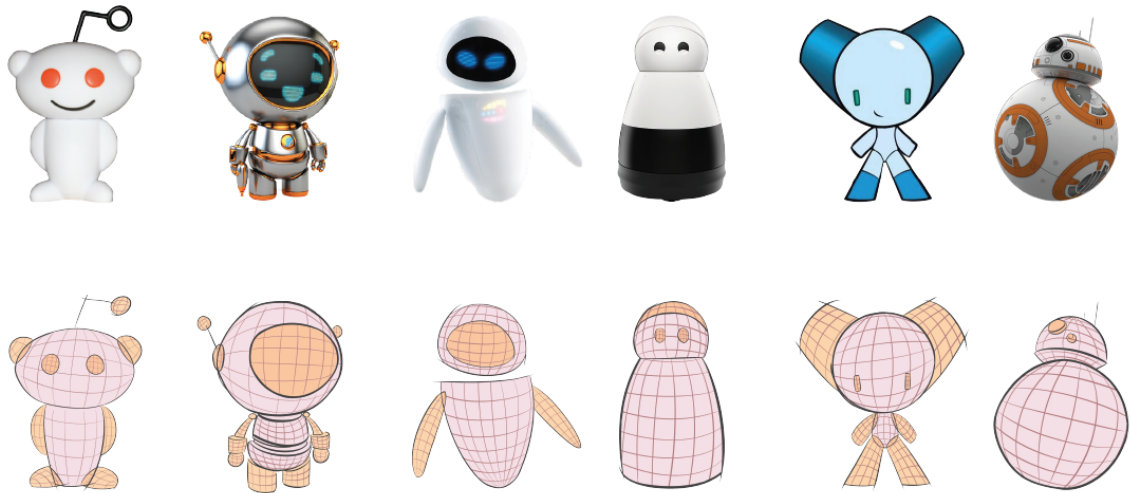


Figure 4.6. Analysis of the design of popular characters



Figure 4.7. Sketches

Colors and texturing

Regarding the colors, the chosen ones for the body is blue since it calls to mind feelings of calmness or serenity. It is a widely used color in logos and graphics that want to convey stability and reliability. Pink has been chosen for the conversational agent’s eyes since it creates a great contrast with blue and because it calls to mind feelings of

optimism, creativity and happiness. The eyes are placed behind a transparent glass visor and can change size and shape to give some human character to the Assistant. Figure 4.9 shows examples of the possible expressions obtainable with only the eyes.

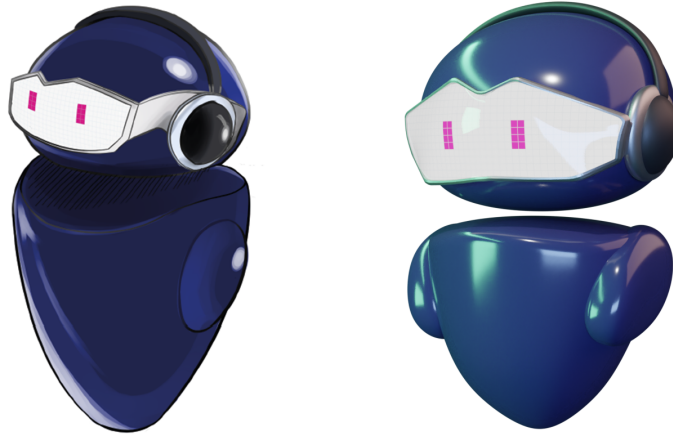


Figure 4.8. From sketch to 3D



Figure 4.9. Expressiveness of the VA. Some of this were unused in the final version.

4.2.2 Virtual Guide

It was deemed necessary to have a character, linked to the theme of the museum, that would welcome and engage the user inside the exhibit rooms. The VA already had the role of guiding the user and helping him with matters linked to the user's wellbeing, ability to find the way and in general more "pragmatic" needs. However, as reported previously in this chapter, inside the exhibit the user needs also to be engaged and interact in what surrounds them, using also game-like elements to keep high levels of attentions [37]. Giving this functions to the VA was deemed inappropriate for the confusion that could cause to the user: the VA should act as if it is the person's own accompanying therapist and not as a museum guide that illustrates the contents of the exhibit. Moreover, every museum has a character of its own, so a personalized character theme with the museum would give a strong feeling of the peculiarities, contents and items of the exhibit rooms. While the VA follows the user in every place they go, this new avatar linked strictly to the museum, was conceived as a static character that is found in each room in a certain place and acts as a museum guide. In fact, the name chosen was Virtual Guide.

The selected museum for the experience is, as previously said, the *Museo di Storia Naturale* of Milan. The museum has a permanent exposition of dinosaurs fossils as well as preserving a copy (the original is at Paleolab Museum in Pietraroja, Campania) of the famous little dinosaur Scipyonix (figure 4.10); nicknamed *Ciro*, is one of the museum's strongest assets and a valuable specimen for Italian paleontology. As such, *Ciro* was taken as the inspiration for the Virtual Guide Virtual Guide (VG). At first, a more true-to-nature design was conceived, but in the end the designed settled on a more cartoony and iconic design.



Figure 4.10. Reconstruction of the Scipyonix specimen nicknamed "Ciro"

Design

The figure 4.11 shows the design of the Guide is that of an anthropomorphized two legged dinosaur. Inspirations were taken from video game characters and mascots with

very strong features to be instantly recognizable. The three dimensional character was modelled with the sculpting tool on Blender software, in order to create soft organic shapes. He has a large head with big eyes, in this way the user knows instantly the direction it is facing. Limbs, feet and hands are soft and rounded shapes, and because the avatar does not need to carry anything in their hands, they were simplified. The figure 4.12 shows the initial sketch of the design, that was then translated into a three dimensional model (figure 4.13a, 4.13b).

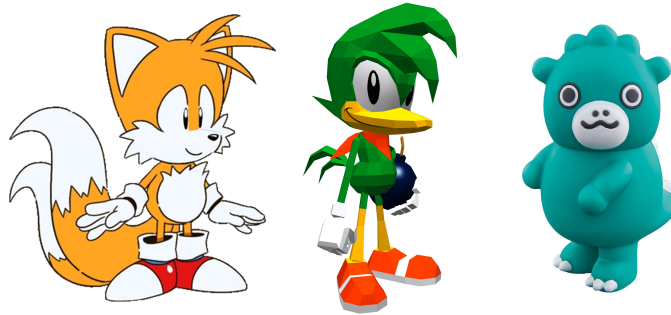


Figure 4.11. Analysis of the design of popular characters



Figure 4.12. Character design

Colors and texturing

Texturing and colors was created following some reconstructions of Scipionyx as well as adding some colorful imaginative touches on the head with some red feathers. The body is green and textured like a reptile's skin and the tail has a series of rings like in paleontologists reconstructions.

Animations

The characters stays fixed in one place, so a lot of animations are not needed. The two animations created for the game are:

- Idle: the relaxed state has a small rocking up and down;

- Jump: when the user completes the game successfully, the character jumps. Figure 4.14 shows the jump animation.

Animating the mouth during talking was considered but ultimately discarded for rigging difficulties and problems correlated to the amount of FPS obtainable in the VR application.



Figure 4.13. Virtual Guide 3D model



Figure 4.14. Jump animation

As said, the role of the VG is that of a museum guide that engages people by making them explore and acquire knowledge. This was obtained by making VG offer the possibility of doing context-related mini-games when the user approaches him. This will be explained in detail in the next section (4.2.3). Important aspect to clarify, the

Guide and the Assistant can and will be present at the same time in the same room, with the Assistant always present to give help and advice about the next thing to do to the user. Their roles are separate, as well as their dialogues with the user, to clearly set a line in dividing this two figures.

4.2.3 Minigames

Once the users are inside the Exhibit room by approaching the Guide they can play engaging context-related activities, or minigames. There are multiple possibilities for a big variety of minigames; for this first prototype a single minigame was developed called Treasure Hunt. The game is activated by speaking with the Virtual Guide, which gives the instruction to the user: as shown in images 4.15a and 4.15a it will give the user a list of items, in the case of this museum of animals extinct or alive, and they will have to find them in the room they are inside at the moment. Not all the items found in the room are part of the hunt, only the most visually appealing or that have an interesting history. To find an object, the user has to go near it and click it (for VR) or frame it (for AR). Once an item is found, information about it is displayed either in a panel under the item (for VR) or on the screen (for AR). This process is explained in detail in chapters 4.3 and 4.4. At the end of the game, the Guide informs the user that all the items have been found and thanks him. The user can then find all the information on the items they acquired in an Encyclopedia accessible in every moment. If the user interrupts the game or exits the room, the progress will be lost.



Figure 4.15. Treasure Hunt

The Encyclopedia is a collection of info and images of the most interesting items found at the museum. At the beginning of the experience, all the data are locked, and are unlocked by playing the minigame Treasure Hunt. In both the applications the user can access the Encyclopedia both during the experience by asking the VA to show it, or in the initial Menu, clicking the “Encyclopedia” button. Figures 4.16a and 4.16b show the Encyclopedia as it is shown in 5A VR and AR applications respectively.

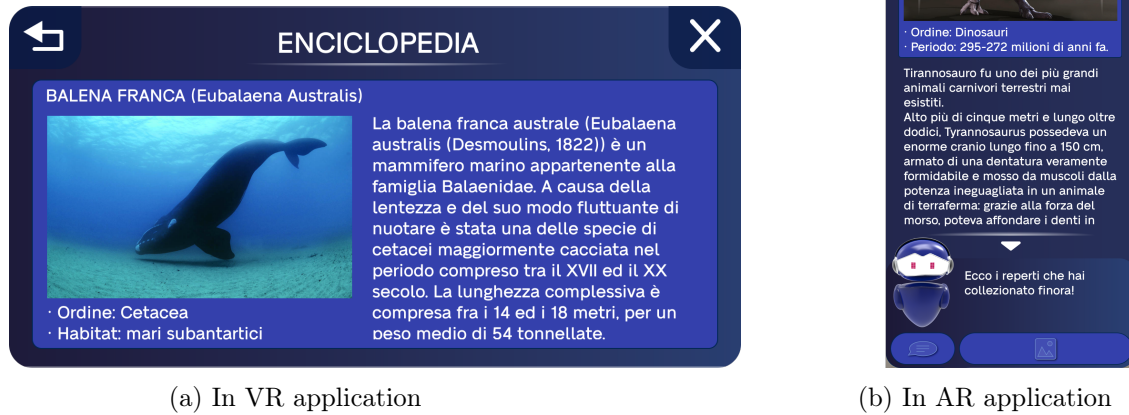


Figure 4.16. Encyclopedia menu

4.2.4 Pathfinding

The applications use different mediums to guide the user through the environment. One method is, however, a common element between the two. This element is the virtual arrow, that guides the user to the next place where the next task is placed. In the initial concept, a stream of arrows reaching from the user to the goal destination was investigated; in the end though it was developed in both the applications as an arrow that stay constantly in front of the user and points the direction of the place that the user is supposed to reach. The arrow is a three dimensional arrow, although it is quite thin in thickness; it has a large size to be always visible and is placed down in front of the user, at the same height as the floor they are walking on 4.17. While in VR the arrow is pointing directly to the goal objective, in AR it is also calculating the path around the obstacles that are encountered to have a more precise and true-to-life path. How this was implemented is explained in detail in section 4.4.

4.2.5 UI design

The design of the user interface of the two applications was created to have as many similarities and points in common as possible. Here are listed the aspects of the interface that can be found in both the applications and are dealt with in a similar manner.

Conversations

On a communicative level, the conversations that the VA and VG initiate with the user have the same tone and style. While working on the concept of the experience, two possibilities were evaluated:

- Have the VA and VG be more of an authority figure that know the answers to the user needs. The way they talk is more direct and command-like, but also



Figure 4.17. Arrow pointing to the next task in AR

encouraging;

- Have a dialogue built like a conversation between two friends were the VA and VG act like a person that proactively tries to let the user open up and talk more. This kind of conversation is built to encourage the user to use their own knowledge and requires a more active role.

After consulting the therapists, it was found out that even if the latter may seem more familiar and encouraging, it could be frustrating and confusing for people affected by ASD. Their restricted ability to comprehend complex sentences coupled with lack of long term attention span and unwillingness to speak and cooperate could make the application really annoying after some time. The former, on the contrary, has a point of strength in its ability to give clear answers. Of course, the conversations are more limited and often one-sided with the VA asking things to the user only to acquire information about the context and help them better.

This choice was selected and the dialogues were then shaped in more direct, effective way, keeping in mind however to always use encouraging and kind tones. In fact, this way of speaking is really similar to how a therapist would speak with a person with ASD. Figure 4.18 shows examples of conversations in both VR and AR.

Vocal and physical interactions

Another factor of considerable importance was to keep similar functions to similar interactions. The interactions in both the apps can be vocal or physical. Starting with the first, in both apps you have to activate a microphone to be able to talk: this was obtained in VR by clicking the *B* button on the controller, while in AR the user has to tap and hold the *Conversation* button to speak. In both the apps the feedback during this action involves a microphone. In VR a three dimensional model of a microphone is highlighted; in AR a 2D graphic of a microphone is shown: in each

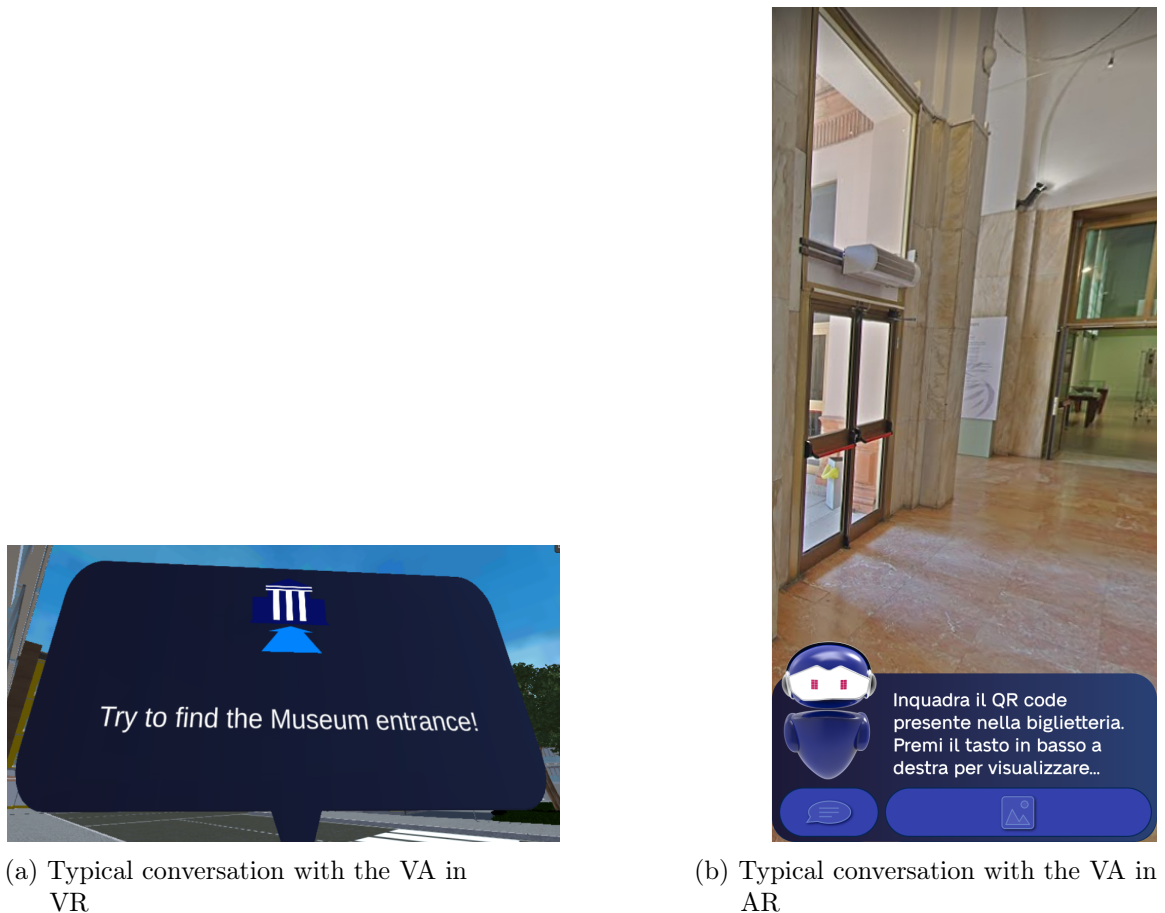


Figure 4.18. Typical conversations with the VA in the two applications

case the microphone shares the same color. Physical interactions with the object were maintained as similar as possible, in particular:

- touching the VA means always that the user is asking for help;
- touching the VG inside the Exhibit rooms triggers always the proposal to start the mini-game Treasure Hunt.

Dialogue boxes

The VA and VG speak to the user by text and voice. The text is shown in dialogue boxes that share a common color and geometry. Color coding is extremely important to communicate to the user he is in a particular state, area or kind of interaction. As seen from the introduction of Chapter 4, to main type of needs were individuated and translated in the two avatars. To reinforce this separation, a color coding of two main colors was conceived.

Colors are of extreme importance in an application built for people with ASD, as there is a high degree of cases of color blindness inside this group. In a study conducted on the 20 ASD individuals examined, 6 (30%) showed color vision losses. Elevated color discrimination thresholds were found in 3/9 participants with autism and in 3/11

participants with Asperger [26]. Some studies report that color blindness affects about 25% of all males affected by autism, with a much smaller percentage for females, but is still a debated topic [16]. The most common type of color blindness is *protanomaly*, a variant where browns, oranges and reds are easily confused, and pale shades in general are difficult to distinguish. Therefore, the colors chosen for 5A are mainly blues and greens. Blue (HTML code #242A52) is the color used for all the interfaces, graphics and dialogue boxes related to the VA and its guiding and help functions. It mirrors the Assistant's body colors and is perfect to communicate serenity and relaxing the eyes. Green (HTML code #065E39) is the the color of the interfaces related to the VG and so by default also to the mini-game. It relates also to the Guide's body color, it is a calm color like blue, but also reminds of wilderness and nature, a common theme in a Natural History Museum. This two colors were then used as basis for a wider range of shades, and different hues were used to achieve a less "blocky" effect in the interfaces. Regarding the shape of the dialogue boxes is in both the applications a rectangle with rounded corner, to communicate a more joyful feel.

Facilitator graphics

Picture 4.19a shows the use of Picture Communication Symbols Picture Communication Symbol (PCS). PCS is a system through which autistic people communicate without relying on speech. To communicate, people use cards with pictures, symbols, words or photographs to make requests or comments and answer questions [6]. This system was investigated and discussed with the therapists, however a full system of PCS able to put together entire phrases ultimately discarded for its complexity. Instead, inspiration was taken from PCS to create graphics that communicate to the user what they have to do in the clearest possible way. Figure 4.19b shows an example of graphics used inside 5A, the example shown in particular indicating to the user that he has to frame the museum entrance with his smartphone. The graphics were integrated both in VR and AR and share the same communicative style.



(a) PCS



(b) Facilitator graphics used in 5A

Figure 4.19. Comparison between PCS and the graphics used in 5A

4.3 VR design

In this section we will see all the specific design choice that have been taken concerning the VR application. All the decisions were made taking into account the user needs, the targeted hardware (Oculus Quest) and the therapists feedback. Before describing the design process, a little introduction on the Oculus device is mandatory. Oculus Quest is a VR headset developed by Oculus, a division of Meta Platforms. The headset is composed by a visor and two hand controllers. Both the visor and the controllers have movement sensors to map the user physical movements inside the virtual environment. Both the controller have an analog stick, three upper buttons, one trigger button and one grip button [4.20]. These controllers allow the users to feel a good sense of immersion and to actively interact with the virtual environment around them.

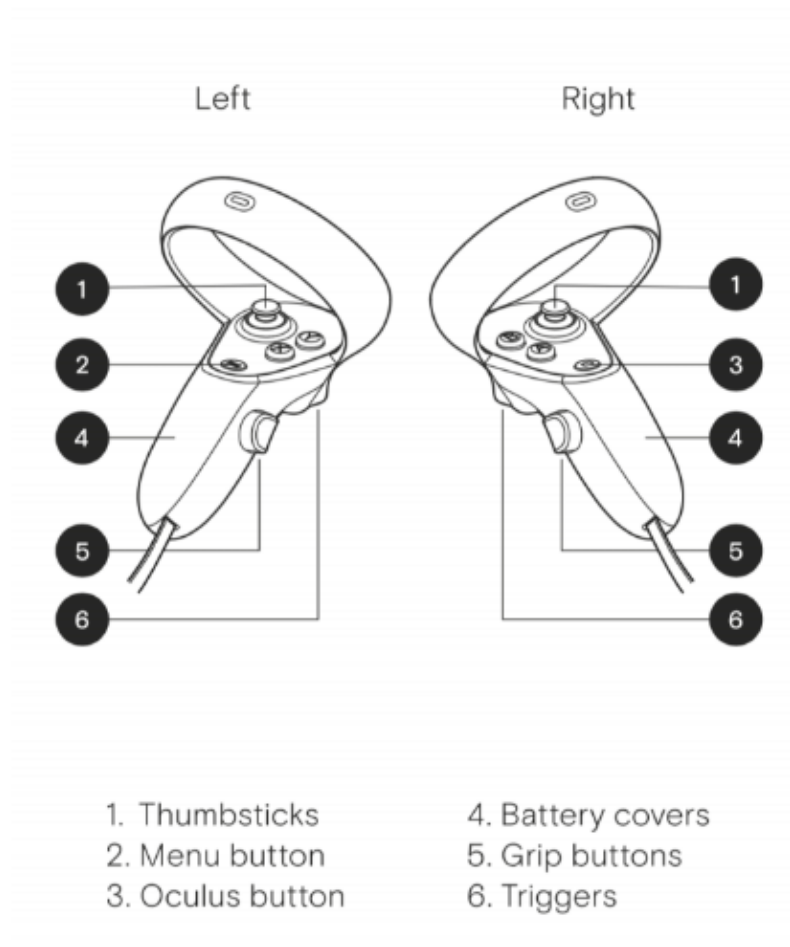


Figure 4.20. Oculus controllers

4.3.1 UX flow diagram

The user experience of the VR application has been designed taking into account the three main big areas that have been identified in a typical visit in a real museum. We can see a general flow chart with the main tasks of the application in the figure 4.21.

In the diagram the three areas (Outdoor, Hall, Exhibit room) are highlighted with a dashed rectangle. Then an expanded diagram for every area has been reported.

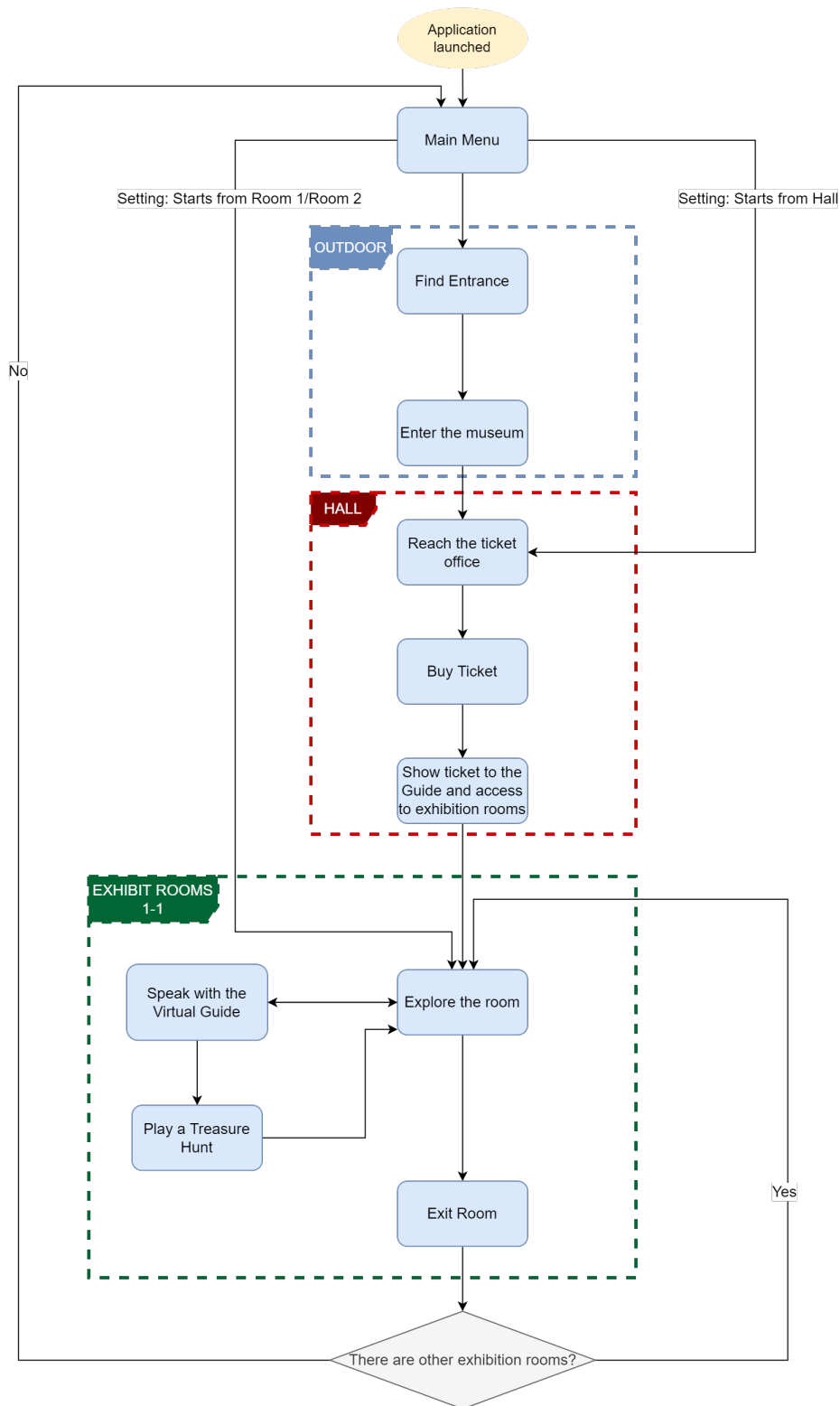


Figure 4.21. UX flow chart of the VR experience

Main menu

On the application launch, the first thing displayed is the settings menu. From here users can set the language of the application and the starting point of the experience. The default starting point is the Outdoor scenario, but users can choose to begin the game from the Hall or from one of the two Exhibit rooms. Another settable option from the menu is the difficulty level of the application. The following experience flow refers to a standard experience with all the difficulty parameters enabled. Figures 4.46, 4.45 and 4.44 show the menu screens.

Outdoor

The first scenario is the city outdoor the museum. The user spawns randomly in one of the three predefined spawn points. The Virtual Assistant welcomes the user and gives a short explanation of interaction commands. Then it assigns the first task to the user: finding the Museum Entrance. Users have to navigate around the city, paying attention to the cars and to walking Non-Player Characters (NPCs). When they reach the front of the museum the Virtual Assistant warns them that they have arrived to the museum entrance and invites them to interact with the door to enter the hall.

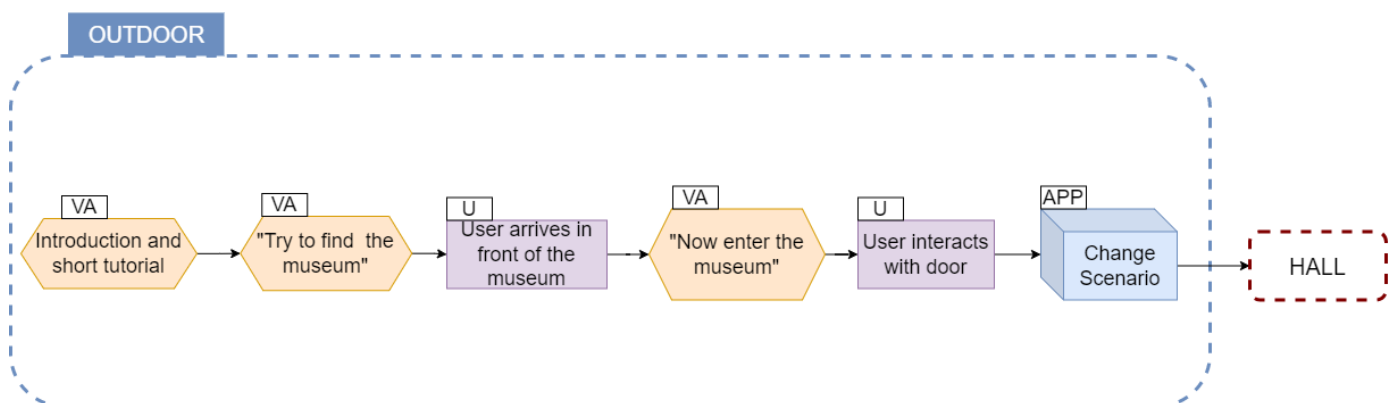


Figure 4.22. Outdoor scenario extended UX Diagram

Hall

Once a user interacts with the entrance door, the scenario changes and they enter in the hall of the museum. The first task assigned from the Virtual Assistant is to reach the ticket office and buy a ticket. The user can freely navigate in the museum hall, but they cannot interact with the door of the next room. The Virtual Guide is placed near to the next room door. If the user tries to interact with him without having a ticket, the Guide will suggest to buy a ticket to proceed to the next room.

When the user reaches the front of ticket office they can interact with the ticket clerk. The ticket clerk will show a dialogue box containing a multiple choice: *pick up the ticket* or *close*. The close button will simply close the dialogue box. To proceed in the experience, the user has to select the *pick up the ticket* button. Once this button is selected, a ticket spawns on the desk. The Virtual Assistant tells the user

to pick up the ticket from the desk. Once the user interacts with the ticket, this one disappears from the desk and appears in the player HUD (heads-up display). The Virtual Assistant suggests to interact with the Virtual Guide to show him the ticket and proceed in the Exhibit rooms. When the user interacts with the Guide while holding the ticket, this will disappear from the HUD and reappear in the Guide's right hand. When the Guide receives the ticket, he invites the user to proceed to the next room and the door becomes an interactable object.

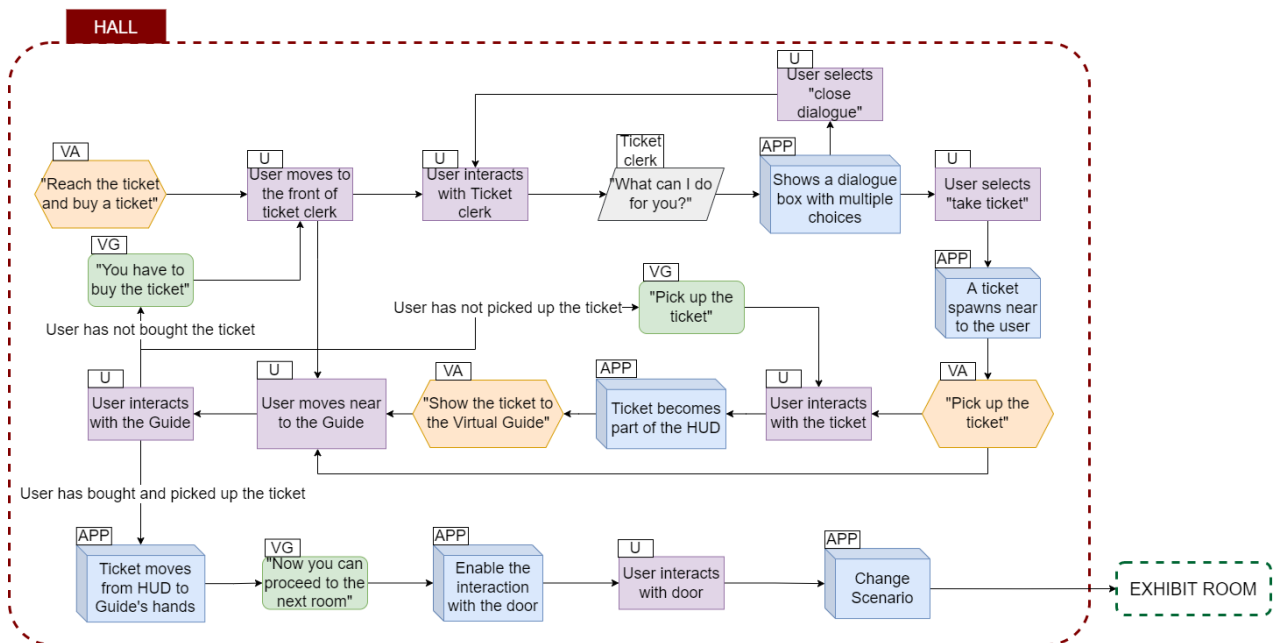


Figure 4.23. Hall scenario extended UX Diagram

Exhibit rooms

Once user enters in the first Exhibit room [4.23], they are free to move without a mandatory task, like they would in a real museum. The Virtual Assistant suggests three possible actions to do: walk around the room to study the items, speak with the Virtual Guide to start a Treasure Hunt game or simply go to the next room. The Guide is positioned approximately in the middle of both the first exhibit room and the second, so that is clearly visible in the path to the next room (or to the exit). Speaking with the Guide activates his dialogue box with two possible options: ask for information about the room or start the Treasure Hunt mini-game. In the first case the Guide will just give some general information about the actual room. In the second case the mini-game Treasure Hunt will start (see next paragraph). Once the user has explored the room and is satisfied, they can interact with the exit door to enter the second exhibit room (or to return end the experience and returning to the main menu if they already are in the second exhibit room).

This experience flow of the exhibit rooms is easily extendable and adaptable to the eventual future addition of new exhibit rooms.

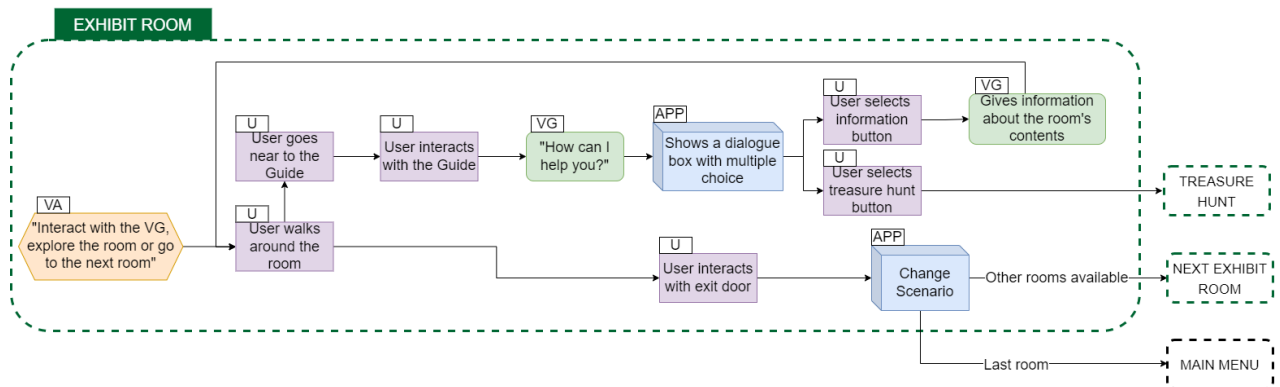


Figure 4.24. Exhibit room scenario extended UX Diagram

Treasure Hunt

The Treasure Hunt game [4.25] starts by selecting the Treasure Hunt button in the Guide dialogue box. When the game starts, the Guide dialogue box is filled with a list of names of four museum items: these are the items that have to be found in that room. Every name contained in the dialogue box is initially grey; when the user interacts with an item, if it is part of the list, the name of the item changes color and becomes light green, to give feedback to the user that they have found a correct item. If the item highlighting facilitator was selected at the start, the items to be discovered will go dark, providing a contrast with the others which will remain illuminated with back-lighting. When the correct item is found it will turn back to being back-lit (see 4.3.3) This procedure is repeated for each of the four items on the list. The dialogue box related to the mini-game can be closed and reopened by interacting again with the Guide. When the Treasure Hunt finishes, the found items are unlocked in the Encyclopedia. For a technical description of the Treasure Hunt game see section 5.4.

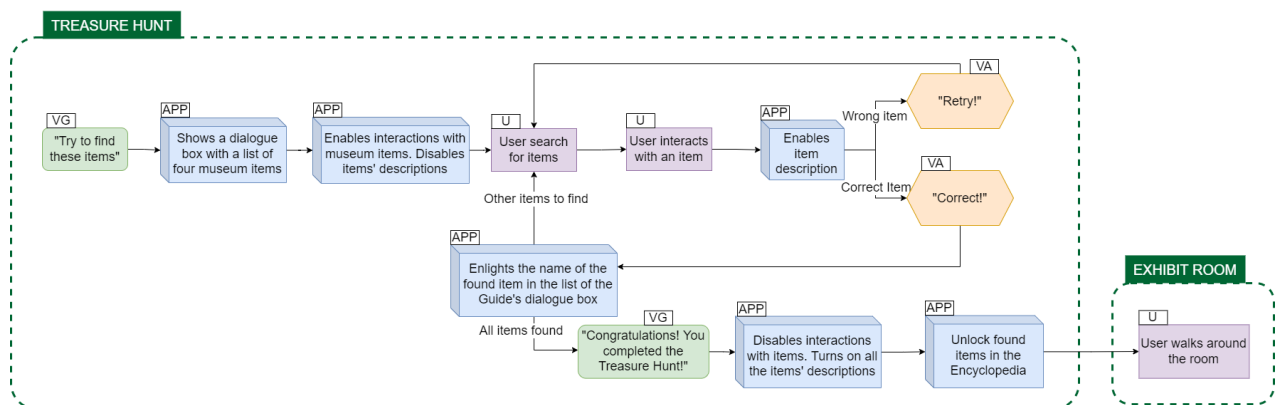


Figure 4.25. Treasure hunt extended UX Diagram

4.3.2 Environment

The city

The first big virtual environment that we can find inside the application is the city outside the museum. This environment is constituted by a road, with the shape of an "8" symbol, surrounded by several buildings and containing small green areas. Two cars are moving on the roads, and some pedestrians walk across the map (see figures 4.28).

The museum is the biggest building of the city and it's placed on the north side of the map. The building's design followed that of the real Natural History Museum of Milan. Figure 4.27a shows the real building: a neo gothic structure with a 80x60 meters plan located in a small private park. The façade is very complex, so it was simplified as shown in figure 4.27b. The colors and textures match the real ones, with big marble inserts, big windows all over the façade, and an entrance with columns and five doors. As it is seen from the outdoor environment, only the perimeter walls and the roof are modeled as it would have just lowered the overall performances of the app; the interior was made not visible by using reflective windows. Differently from the real building, a big sign saying *MUSEO DI STORIA NATURALE* is positioned on the Entrance to help the users identify the building. The park around the building was not modeled as it would have been too easy for the user to recognize a building so different from the others. 4.29 shows the rest of the buildings of the city: some skyscrapers like the one showed were modeled, while the rest of the houses were imported from previous 5A projects.

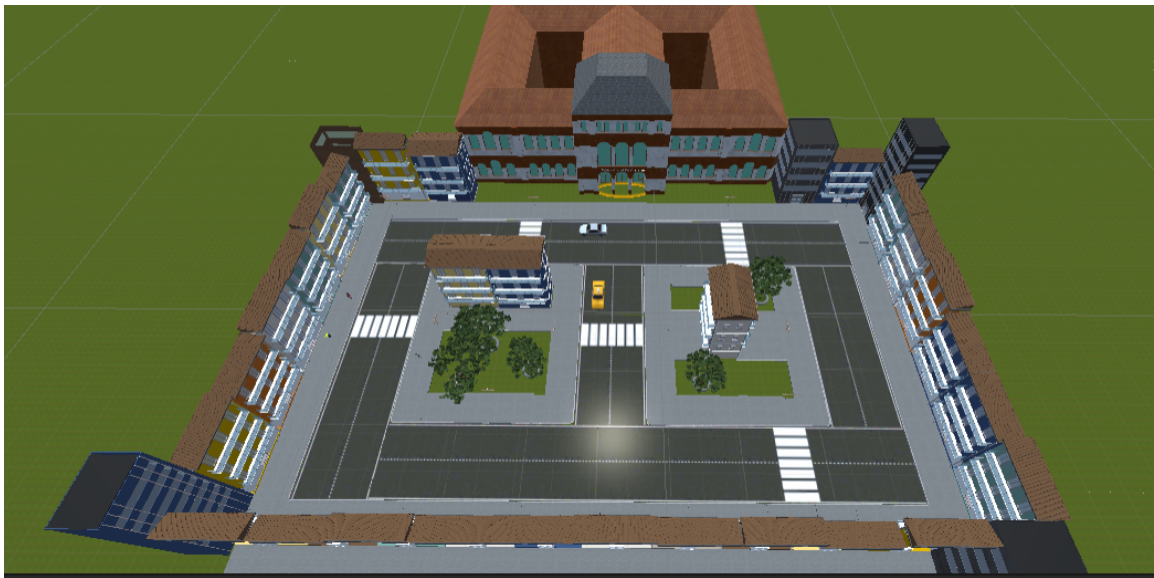
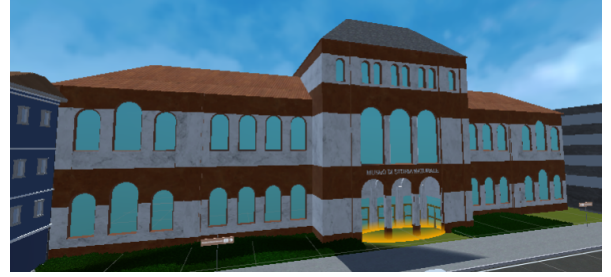


Figure 4.26. The city in the outdoor scenario



(a) The museum in real life



(b) Museum 3D model

Figure 4.27. Comparison between the real and virtual museum



(a) NPC walking



(b) Car

Figure 4.28. Moving objects and people on the city streets



Figure 4.29. Houses and skyscraper

The Museum Hall

The Hall is the first inner room of the museum. Once the user enters in this environment, the city environment is deactivated as it is not visible anymore to smoothen overall performance. The Hall is composed of a large room with marble walls and big columns, exactly as the real-life one. The ticket booth is placed in the middle of the room, at the sides there are low stairs and decorative pieces. The five doors on the south side are the same doors of the external facade, visible from the outdoor of the museum. The door located in the east side with the sign above indicating the name of the room is the one communicating with the first exhibit room. Near this door we can find the Virtual Guide waiting for the ticket (see figure 4.31).



(a) Hall in real life



(b) The museum hall in the virtual scene

Figure 4.30. Comparison between the real and virtual hall



Figure 4.31. Door communicating with the first exhibit room



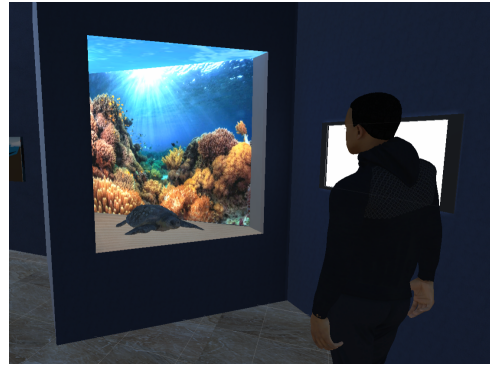
Figure 4.32. The first exhibit room

First exhibit room: marine animals

The first exhibit room contains several items on marine animals alive today. Most of the items are back-lit pictures, but there are some showcase containing 3D models of small animals. The structure of the room is intricate and small: this was a conscious choice made by the fact that several of the Natural History Museum's real rooms are made like this. In the middle of the room we can find the Virtual Guide.



(a) Backlit images



(b) Showcase with Marine Turtle

Figure 4.33. Different methods of item display inside the museum

Second exhibit room: Dinosaurs

The second exhibit room is the dinosaurs room. Contrary to the first room, this one is an open space shaped like a big hallway. On the left sides are positioned three 3D models of different specimens of these extinct creatures. On the right, a big glass showcase shows the evolution of this animals from dinosaurs to birds with the help of a Velociraptor and a an Eagle. Finally, in the end, the model an Allosaurus and a Pteranodon in a small habitat with plants are showed. The Guide is found once again waiting for the user and is located in the middle of the room.



Figure 4.34. The second exhibit room



Figure 4.35. Allosaurus reconstruction



Figure 4.36. Showcase of the evolution of dinosaurs

4.3.3 Interactions

In the virtual environment users can interact with several objects placed in the scenario. First of all it must be stated what is meant by the term *interaction*: by interaction is intended every single action made by the user that can modify the state of an object in the virtual environment. These includes moving around the map, moving objects, activating objects, and many others. In our case, we have two main type of interactions: player movements and objects activation.

Moving in VR

Usually in VR we have three type of possible movements:

- **Teleport Movement:** users can teleport on some special platforms placed around the map;
- **Free Movement:** users can move normally using the joystick like in a standard video-game;
- **Physical Movement:** this corresponds to the users' movement in the real environment that are detected by the motion sensors of the worn device.

The physical movements of the head in a VR application are always enabled. while other movement can be restricted.

Concerning this type movement, the most utilized one in VR is generally the teleport movement. This is due to the fact that free movement can sometimes cause motion sickness symptoms in the playing user.

However, the design decisions on the movement system were made taking into account the therapists opinion based on past 5A applications in scenarios like taking the subway or going to the grocery store. The therapists thought that moving only by teleport using restricted platforms placed by the developers could result in a too simplified experience, reducing the training benefits on the final user. In light of this, we decided to maintain both the free movement and the teleport always active. Moreover, the teleport moves are done on big platforms that cover all the walkable ground in the scenarios. In this way, even with the teleport, the user can move across all the environment, without being constricted by the position of the teleport platforms.

Player movement inside the environment can generate changes in other moving objects, in particular in vehicles and pedestrians. All the cars and NPCs will stop moving if the user passes in front of them. They will restart their movement when the user is no longer blocking their way.

Other interactions triggered by movements are the activation of tasks. When the user reaches the target goal, the new task is activated and the Virtual Assistant will explain it.

Ray interactors

As we said in this chapter introduction, Oculus Quest device is constituted not only by the visor, but also by two hand controllers with movement sensors. In game these

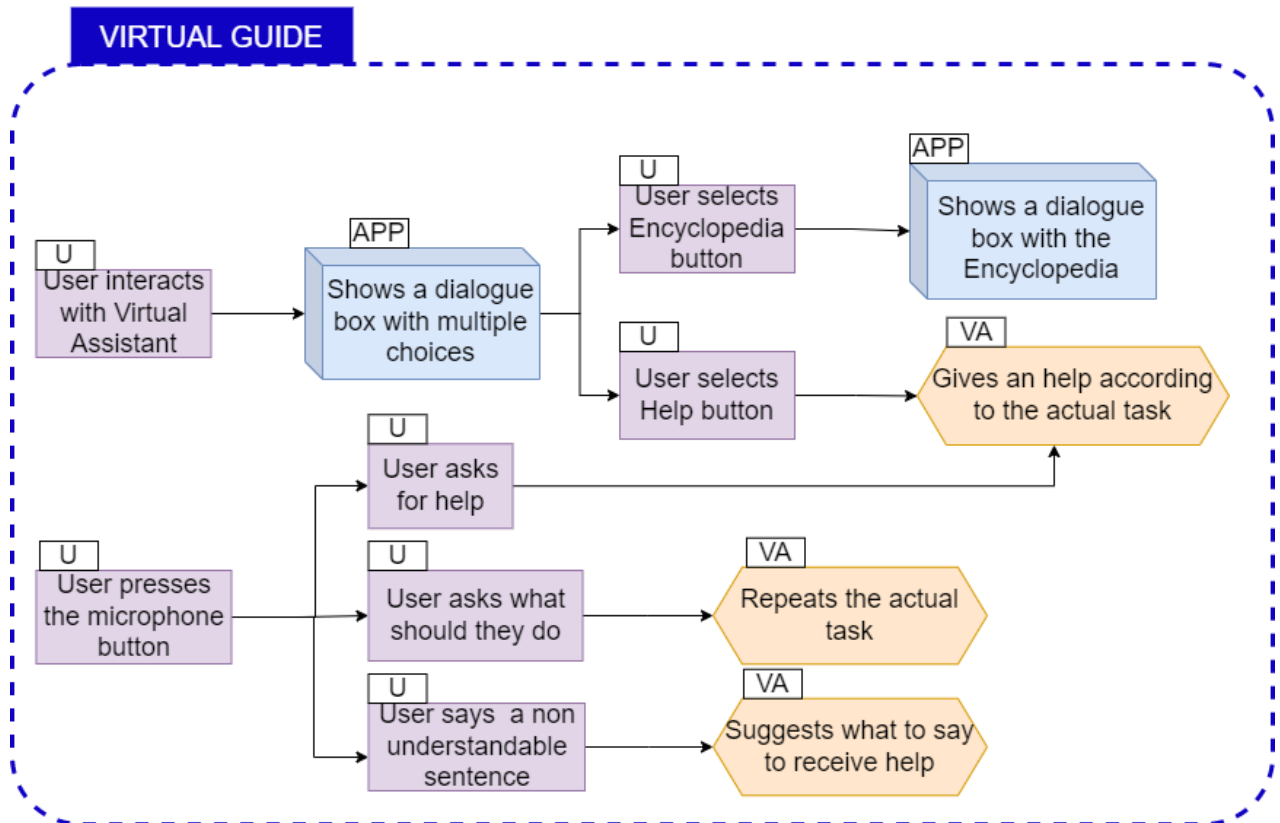
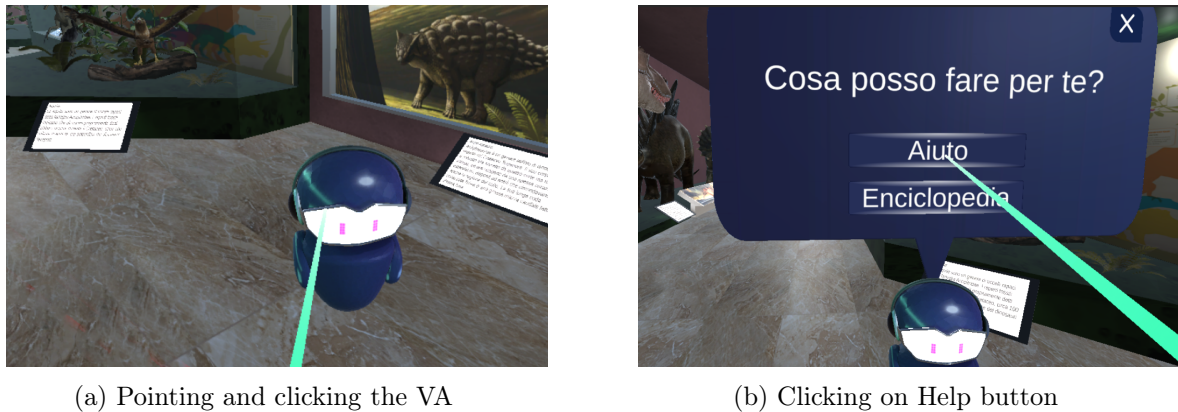


Figure 4.37. Virtual Assistant extended UX Diagram

controllers are represented by approximated 3D models with a red ray casted from the top of it. When a ray is blocked by an object marked as *interactable* and it changes color, becoming teal. When the ray is in this state, users can press the interaction button (mapped on the trigger button and grip button [4.20]) on the controller to "activate" the object. There are several interactable objects and characters inside all the scenes:

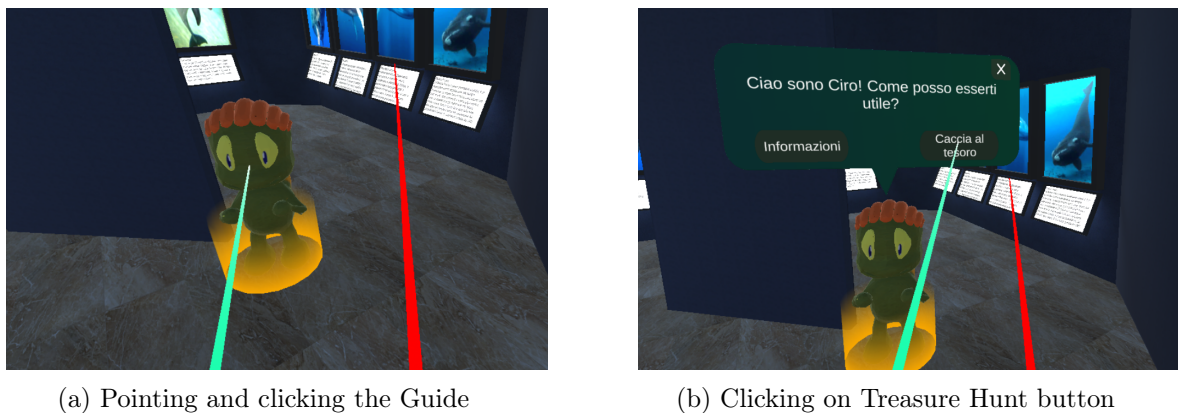
- **VA:** the little robot assistant is normally positioned on the front-right of the user, hovering over the floor; however if the user is really close to a wall or item the Assistant moves to the nearest free position available to be always visible by the user. Also, the VA always faces the user. Interacting with it is always possible and will open a dialogue box showing an *help* button and a *Encyclopedia* button [4.37]. By selecting the *help* button the VA will give an hint on how to complete the current task. By selecting the *Encyclopedia* button the Encyclopedia dialogue box will be showed.
- **VG:** the Virtual Guide does not move and is located in key areas. The interaction with him can happen in these different areas, and comes with two different interaction sequences. If the user interacts with the Guide inside the Hall, he will simply tell them to take the ticket or to proceed in the next hall if they have already bought it [fig: 4.23]. If the user interacts with the Guide inside the Exhibit room, he will show a dialogue box with the *Information* button and *Treasure Hunt* button [fig: 4.24].



(a) Pointing and clicking the VA

(b) Clicking on Help button

Figure 4.38. Interactions with the Virtual Assistant

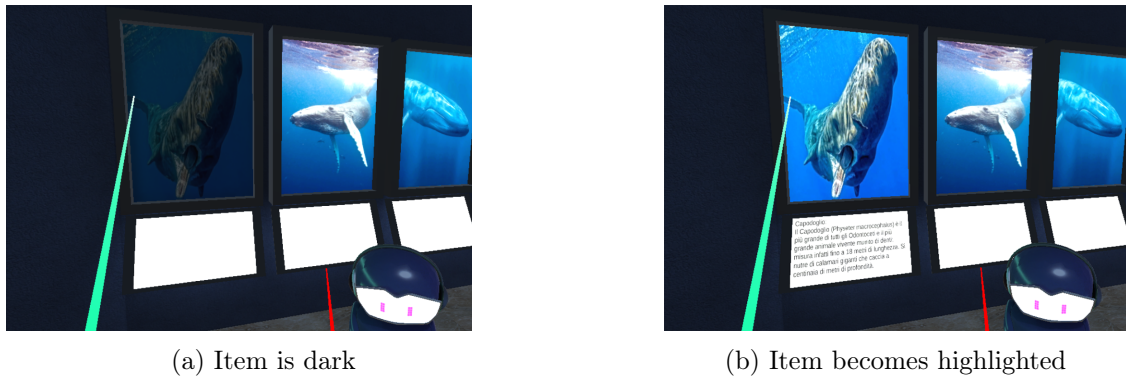


(a) Pointing and clicking the Guide

(b) Clicking on Treasure Hunt button

Figure 4.39. Interactions with the Virtual Guide

- **Museum items:** museum items are the objects exposed in the exhibit rooms. They become interactable only during the Treasure Hunt game. Interacting with them will turn on the description panels and will continue the Treasure Hunt game (if the item is correct). If the user has activated the facilitator related to the item highlighting, the items to be found will go dark like shown in 4.40a creating a contrast against all the others that will remain illuminated with back-lighting. Once the correct item is found, it pops back to being highlighted by back-lighting as shown in 4.40b. This process repeats for all the four objects in the list.
- **Ticket clerk:** interactions with the ticket clerk are enabled when the user is in front of the ticket office. These will trigger the dialogue box of the ticket clerk and will allow the user to buy a ticket from the ticket clerk.
- **Ticket:** the ticket can be picked up from the desk and given to the Guide. When user interact with the ticket near the ticket office, this object becomes part of the Head-up display (HUD). Then, interacting with the VG, the ticket will be given to him.



(a) Item is dark

(b) Item becomes highlighted

Figure 4.40. The process of clicking on a Treasure Hunt item

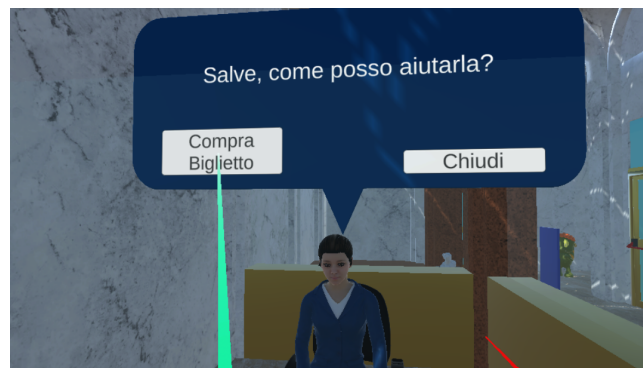
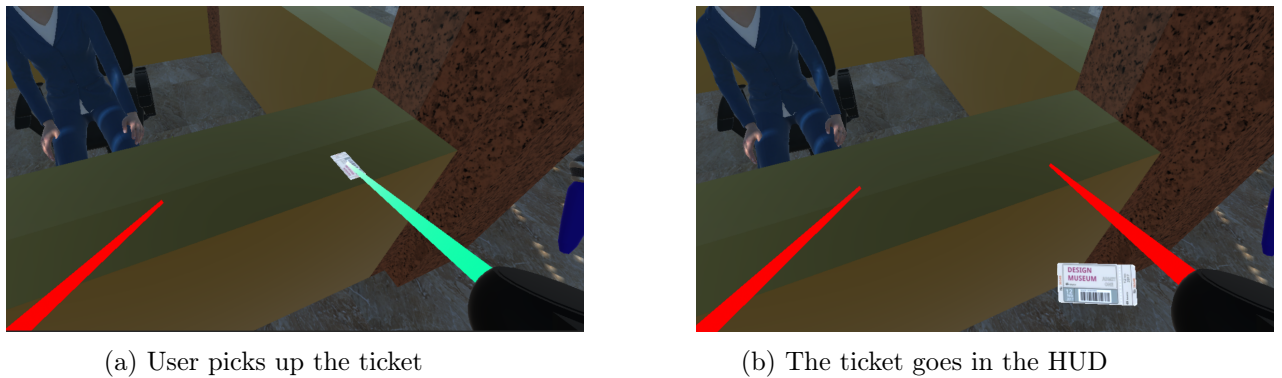


Figure 4.41. Ticket clerk asking the user if they want to buy a ticket



(a) User picks up the ticket

(b) The ticket goes in the HUD



(c) User bring the ticket to the Guide to check it

Figure 4.42. User bring the ticket to the Guide to check it

Vocal interactions

The application also support vocal interactions with the VA. Vocal input is activated by pressing the *B* button on the right controller. This will turn on the little microphone attached to the right controller, that becomes bright orange. The application will stop to record the vocal input automatically when the user stops speaking (or after 15 seconds if the user doesn't start to speak). The vocal interaction can have three possible outcomes [4.37]: the Assistant gives an help, the Assistant repeats the task, the Assistant suggests what to say.

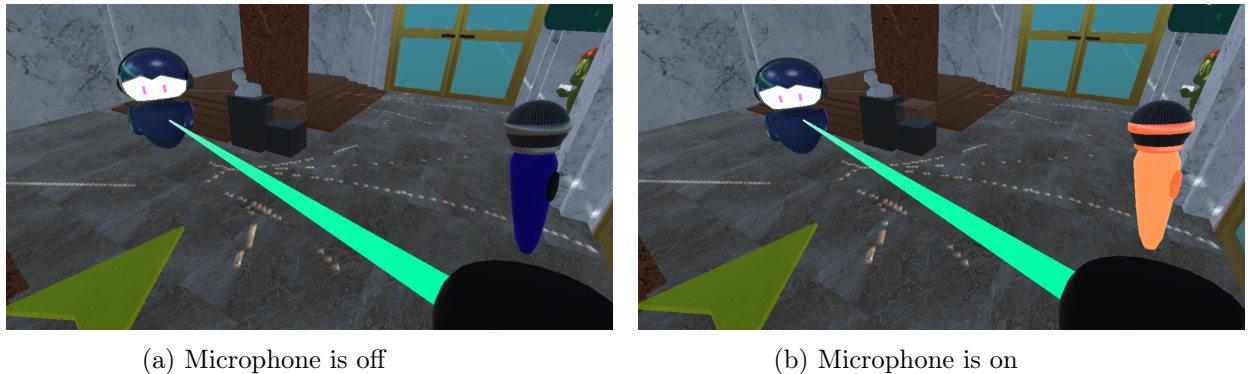


Figure 4.43. Activating the microphone

4.3.4 User Interfaces

Almost every user interface element in game is constituted by a dialogue box that appears above an interactable item. This dialogue boxes are simplified if compared to the ones in the AR application, for the fact that subtle color gradings in the VR environment didn't sort a visible result. Therefore, solid colors of greens and blues were then chosen. Another difference is the horizontal development of the dialogue boxes that being positioned in a free space have much more freedom in the horizontal direction. Similarly to the AR application, the overall shapes of the dialogue boxes and buttons is a rounded and regular style. All dialogue boxes close after 12 seconds or if the user click on the *X* button on upper right side, and can be reopened by clicking on the VA or VG respectively.

Main menu

The first big User Interface (UI) component is the initial main menu showed on the start of the application. This menu is composed by a dialogue box containing two big panels. The first one is always visible and contains four buttons: *Language*, *Starting zone*, *Difficulty* and *Start*. Selecting one of the first three buttons will activate or change content of the right panel.

VA dialogue boxes

The dialogues of the VA are composed by three different dialogue boxes. The default one, that shows only text supported by Facilitator graphics showed on the upper side



Figure 4.44. Main menu with language panel



Figure 4.45. Main menu with starting zone panel

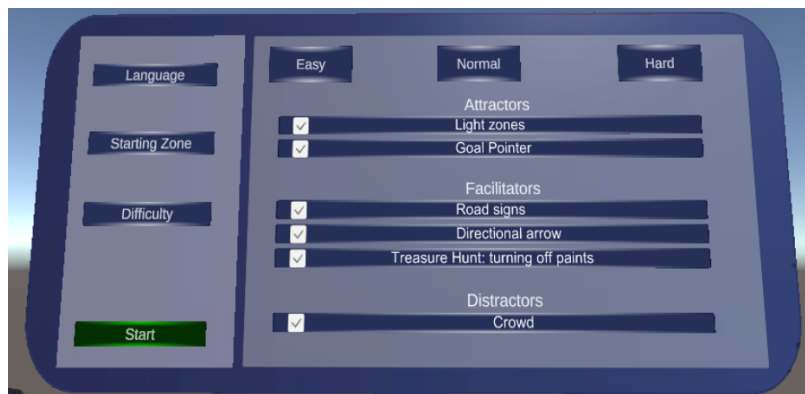


Figure 4.46. Main menu with difficulty panel

of the box, is activated every time that the Assistant talks, while the Help dialogue box and the Encyclopedia dialogue box are activated on interaction. Under the text are positioned the buttons that show the possible options. The full list of Facilitator graphics is shown in figure 4.49.

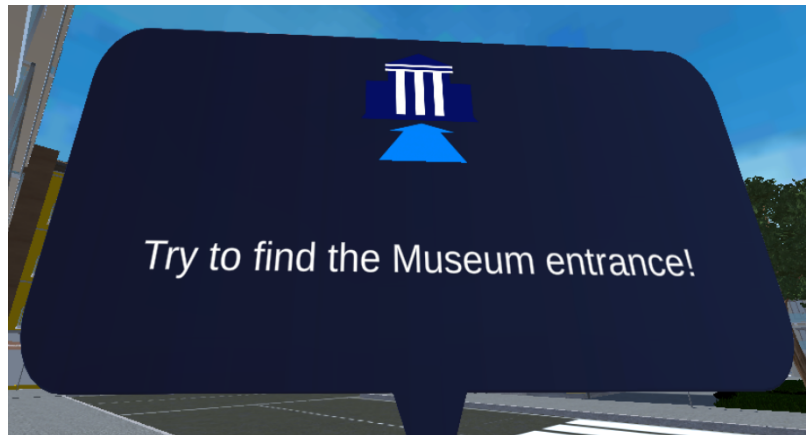


Figure 4.47. Virtual Assistant's default dialogue box

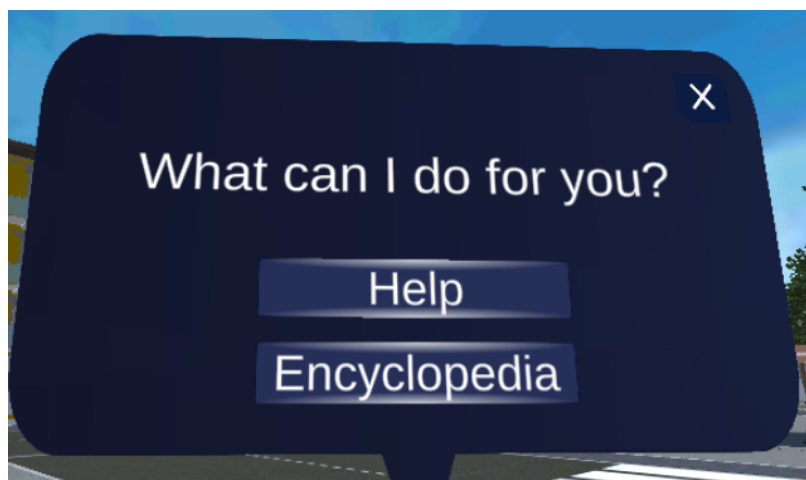


Figure 4.48. Virtual Assistant's help dialogue box

Encyclopedia

The Encyclopedia is constituted by two dialogue boxes. The first one lets the user choose the room they want to see (figure 4.50a), then presents the list of items unlocked in that particular room (figure 4.50b). It has a "page" structure, where every page contains a maximum of four items. Two navigation buttons are used to scroll through the pages. The items found during the Treasure Hunt are selectable, while the ones that have not been unlocked are unselectable. The second dialogue box contains an image of the selected item, the name and the description (figure 4.50c).

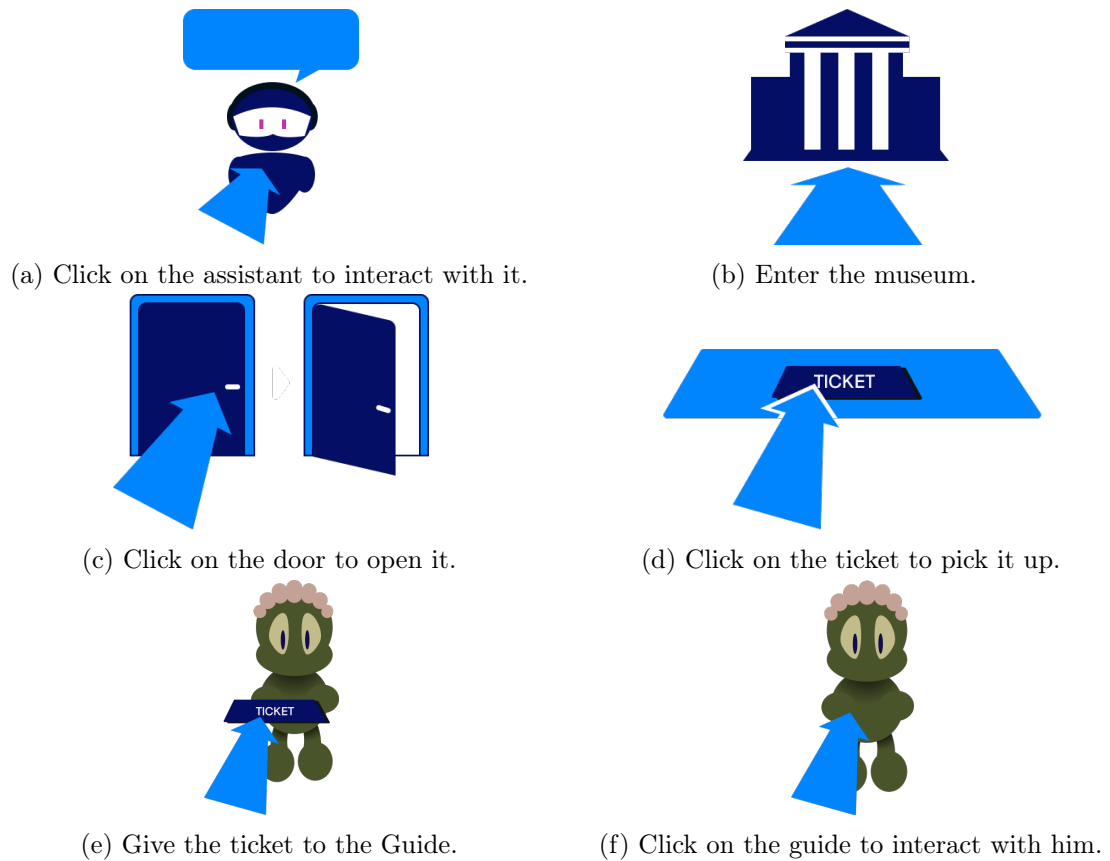


Figure 4.49. List of Facilitator graphics.

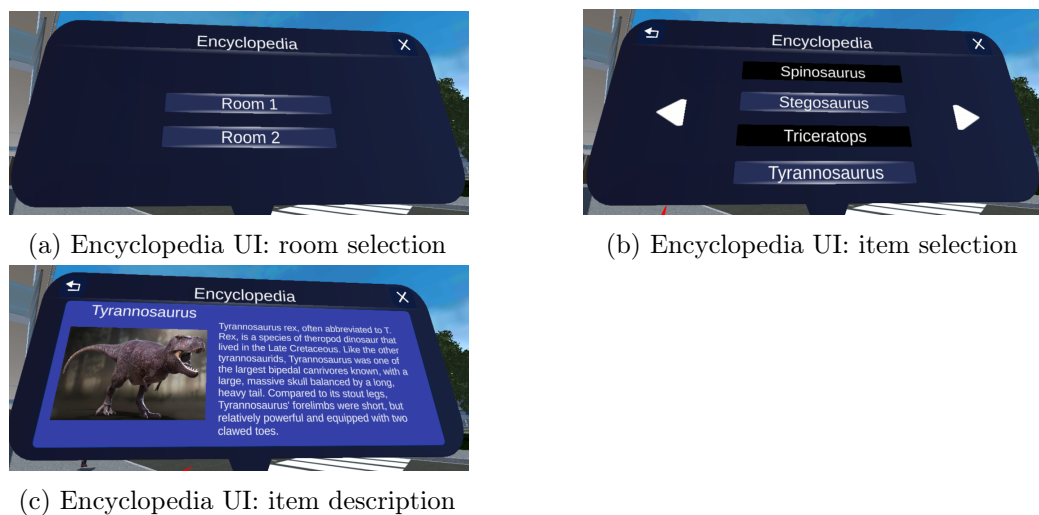


Figure 4.50. Encyclopedia UI

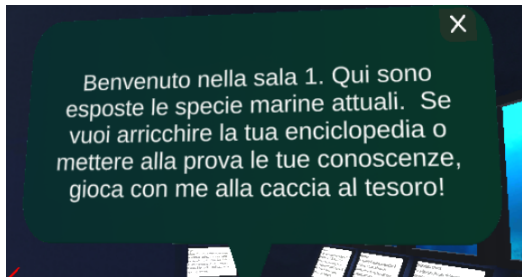
VG dialogue boxes

The Virtual Guide, like the Assistant, has three different dialogue boxes:

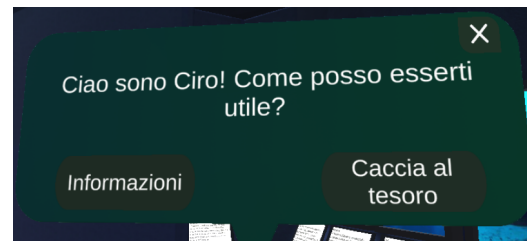
- the information dialogue box: it contains only text;
- the interaction dialogue box: contains the Information button and the Treasure

Hunt button.

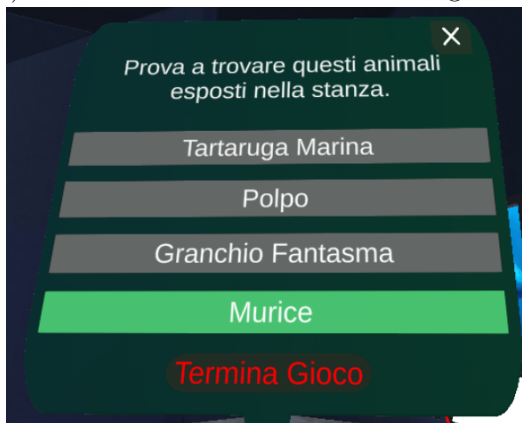
- the Treasure Hunt dialogue box: contains the items that have to be found during the Treasure Hunt. Main components are four caption containers that become bright when the user interact with the right object in the room.



(a) Virtual Guide UI: Information dialogue box



(b) Virtual Guide UI: Interaction dialogue box



(c) Virtual Guide UI: Treasure Hunt box

Figure 4.51. Encyclopedia UI

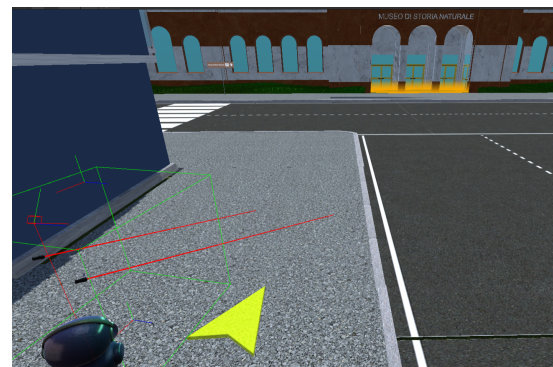
4.3.5 Customization of the Experience

Like in the other two 5A projects, a system of Facilitators, Attractors and Distractors have been used to face the needs of users affected by ASD. Facilitators are graphical elements that have the function to help the user in finding the right path to follow or the correct object to interact with. Inside the VR application we have three Facilitators:

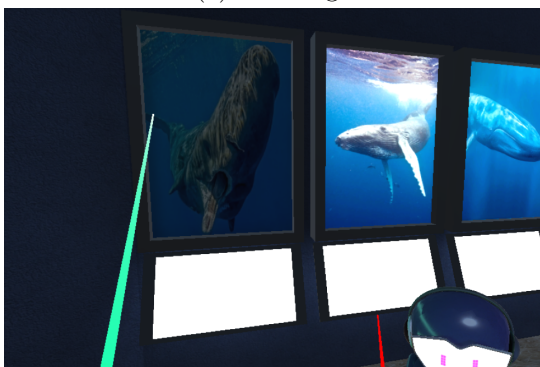
- Road signs: the signs meant to simulate actual road signs in the outdoor scenario that point toward the museum.
- Directional arrow: it's the arrow in front of the player that points always to the current goal.
- Treasure Hunt - turning off items: this is a facilitator regarding the Treasure Hunt game. If this facilitator is enabled, when the treasure hunt starts, the items to be found will turn off the back-lighting. These items will return bright when the user interacts with the item.



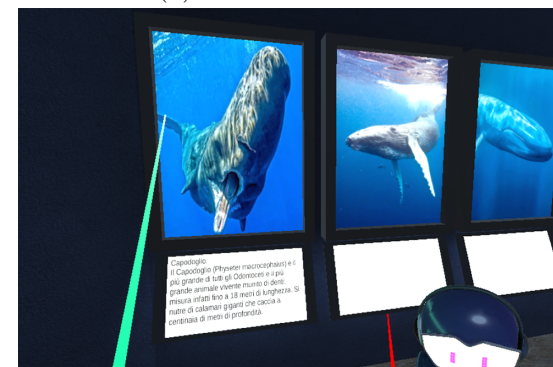
(a) Road sign



(b) Directional arrow



(c) Item turned off



(d) Item with backlighting

Figure 4.52. Facilitators

The Attractors are a subcategory of Facilitators. Their function, as the name says, is to attract the user towards them or simply attract the user sight. In the application we have two Attractors:

- Highlighted zones: it's represented by the orange beacon light that is around every goal.

- Goal pointer: the orange 3D arrows that spawns above the Virtual Guide



(a) Highlight of the area in front of the ticket booth (b) Vertical 3D arrow pointing at the Virtual Guide

Figure 4.53. Attractors

Finally the Distractors have the task to increase the difficulty of the game, simulating external stimuli that can be found in a public space. The Distractors are the walking and talking NPCs inside and outside the museum, and the cars driving on the road outside. The crowd talking sound is a crowd stock sound that serves the purpose of increasing the perception of a big number of people.

Every difficulty option can be activated or deactivated from the main menu, before starting the experience. The application also gives the option to select from three predefined difficulty levels:

- Easy: all Facilitators and Attractors are enabled Distractors are disabled;
- Medium: highlighted zones and treasure hunt facilitator are enabled. Distractors are enabled. All the other option are disabled.
- Hard: Distractors are enabled. All Facilitators and Attractors disabled.

4.4 AR design

In this section we will go into details about the design choices that were taken before and during the development of the AR application of 5A-Museums. These decisions, like the ones from the VR application, have been made and refined taking into account user needs, therapists feedback and strengths and weaknesses of the targeted hardware (Android and iOS mobile devices).

4.4.1 UX Flow diagram

The user experience of the AR application has been designed in order to trace the one of the VR application as closely as possible, in order to be more familiar and less distracting for the user. To do so, the AR experience has been divided into three macro sequences that define a typical museum visit: finding the entrance of the museum (Outside), buy the ticket and show it at the entrance of the exhibit rooms (Hall) and finally the exhibit rooms tour (Exhibit).

As we can see from the figure 4.54 the three macro sequences are highlighted with colored dashed rectangles. A further remark must be made for the two states at the sides of the flow diagram that are reachable from and lead to multiple states:

- **Scan QR Code:** In any moment of the experience, a user should be able to scan a QR code and change the state of the application depending on the contents of it. For example, if a user launches the application while in the hall of the museum, he must scan a QR code in order to jump directly to the first state of the *Hall* section and avoid the states from the *outside* section. In order to explain this behaviour, the *Scan QR Code* state must be reachable from each other section and then, depending on the contents of the QR code, it will lead to the beginning of another section.
- **Ask for Virtual Assistant help:** As in the previous case, in any moment of the experience, a user should be able to interact with the VA and ask him for help changing the state of the application. For example, if a user asks for help about the museum exhibit, the application will immediately jump to the *exhibit* section assuming that the user is situated in that part of the museum. Since no QR codes have been scanned in order to reach the section where the user needs help, the application must ask to scan the nearest one in order to recover the data about the room and to check that the user has not asked for help concerning the wrong section of the experience. This explains why the *Ask for Virtual Assistant help* state always leads to a *Scan* state, excluding the *outside* section, where data and position are recovered exploiting GPS functionalities.

In the next sections it will be given an in-depth analysis of specific sequences of the application flow, supported by more specific parts of the flow diagram. The images that support the analysis are related to the guided mode, because the guided experience is the most complete and it will be also the one used for the experimentation. A more exhaustive description of the difficulty levels and their differences will be given in the *Levels of difficulty* section (4.4.3).

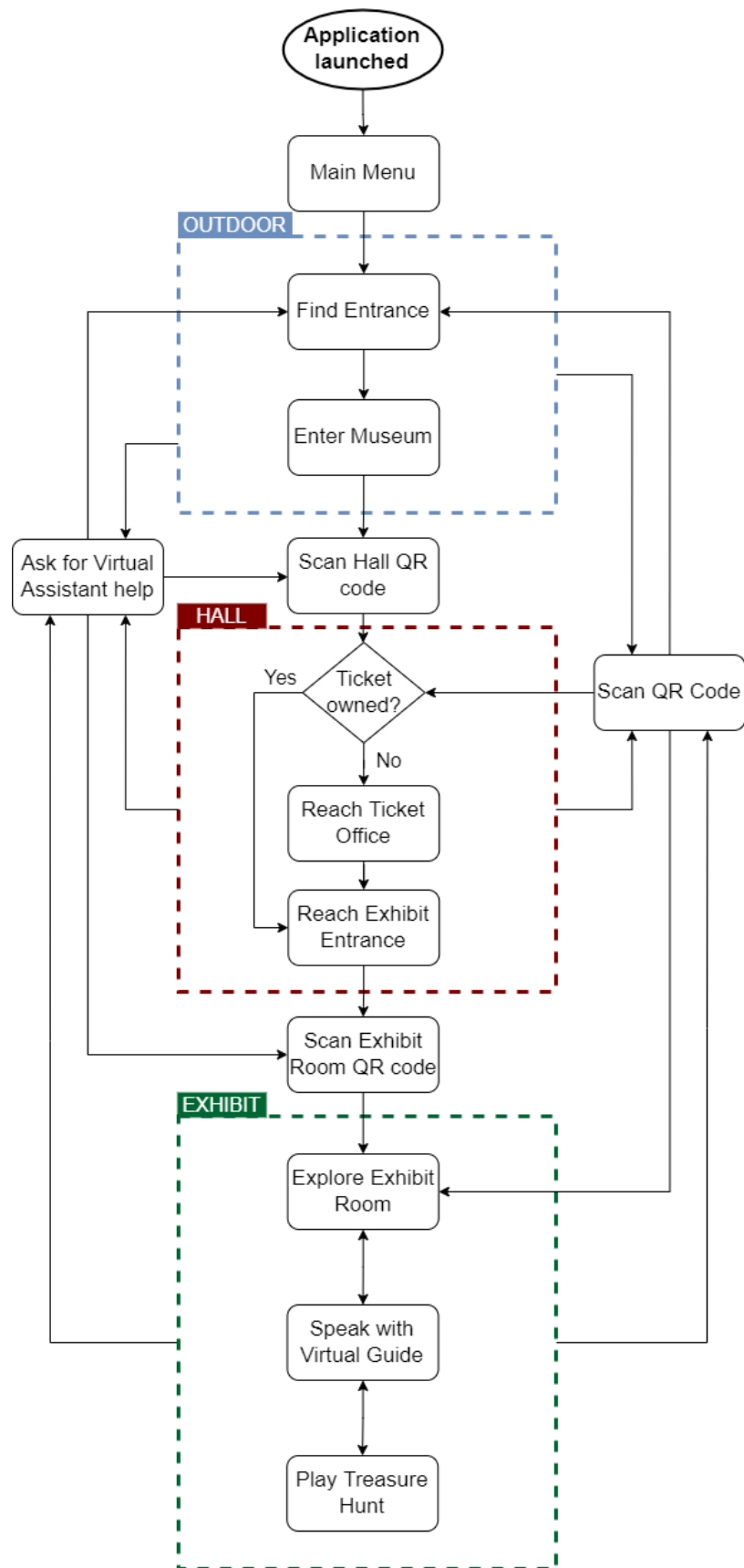


Figure 4.54. Temporary

Main Menu

When the application is launched, after a splash screen with 5A-Museums placeholder logo, the user lands on the main menu. From the menu it is possible to select the language in which the application will be displayed and choose the difficulty level between autonomous and guided. The difficulty levels are described in more detail in section (4.4.3).

After choosing language and difficulty level, a short tutorial that will explain the main functionalities of the application will be shown. In order to advance with the tutorial, the application will ask to try some features in order to be sure that the user understood how to use them. The tutorial screens explain how to ask for help, where to read text from the Assistant, how to use the vocal inputs, how to display the Facilitator graphics, how to scan a QR code. The app also provides a button that can be used to skip the tutorial pages, in order to give to a user that is already familiar with the features the possibility to avoid this phase.



Figure 4.55. Examples of menu and tutorial screens

Outdoor

Right after the tutorial, the state of the application switches automatically to the outside section. As shown by the image 4.57, at first the application recognizes the

museum and recovers all its data exploiting the user's GPS location, without having them look for a specific QR code. After this phase, the user is asked to say whether he is inside or outside the building and in the first case the application asks the user to scan a QR code in order to locate them, while in the latter the flow continues normally. This design choice was taken in order to allow users that launch the application while inside the museum to avoid the *outside* section.

After the application has made sure that the user is outside the museum, the flow goes on in two different ways for **Guided** and **Autonomous** mode:

- **Guided mode:** In the guided mode, as shown by the figure 4.57, the application shows a virtual arrow on the ground slightly ahead of the user and the VA tells them to follow its direction in order to reach the entrance of the museum. The arrow is blue and always point the shorter path from the user to its target, in this case the entrance, taking into consideration the obstacles along the way.

Once the entrance is reached, the arrow disappears and the VA congratulates the user and gives them its next task, while the application waits for the user to scan a QR code or ask for help to the Assistant.

- **Autonomous mode:** In the autonomous mode, on the other hand, the application won't show the virtual arrow and the VA asks the user to try to find the entrance and frame it with his device's camera. Once the user frames the part of the museum where the entrance is, the application recognizes it using image recognition, then the VA congratulates the user and gives him its next task, while the application waits for the user to scan a QR code or ask for help to the Assistant.



Figure 4.56. Example of a screen displayed while outdoor

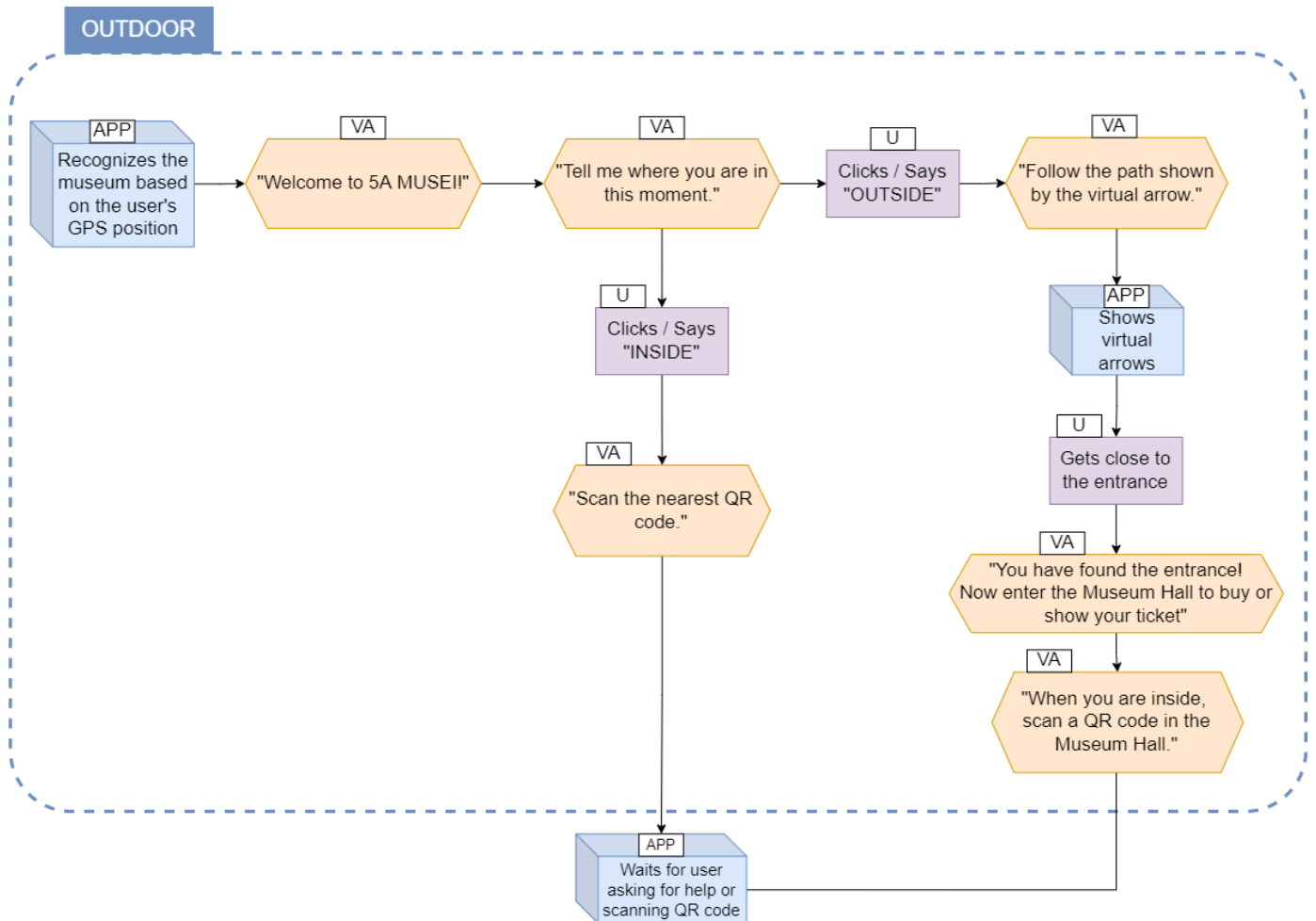


Figure 4.57. Outdoor AR UX Diagram

Museum Hall

The museum hall experience can be divided in two separate tasks, that are reaching the ticket office in order to buy a ticket and reaching the entrance of the exhibit, where the ticket must be shown to a museum employee.

As shown by the figure 4.59, the beginning of the application flow inside the hall, the VA asks the user if they already bought a ticket for the exhibit: if the answer is affirmative, the application avoids the sequence that guides the user to the ticket office.

The flow of the two tasks is very similar, and the pathfinding works in the same way as in the previous section. The differences from the *outside* section are mainly two. First of all at the beginning, since the application cannot be sure that a QR code has been scanned, the VA must ask the user to scan the nearest one. Secondly, once the user reaches its target, the application does not know how long it will take for them to get its ticket or to show the ticket to the museum employee, so it displays a button saying *I have bought the ticket* that, when clicked, starts the following task.

In **Guided** mode, since the two main tasks of the *hall* section are linked, it will be possible to continue from the ticket office to the exhibit entrance without scanning a second QR code before starting the second one. In **Autonomous** mode, on the

other hand, the two tasks are completely separate and the continue buttons are never shown. For example, if a user needs help in reaching the ticket office, they can ask the VA for assistance, and it will guide them in the same way as the Guided mode but instead of showing the continue button, the application would wait for the user asking for help again or scanning a QR code. At this point, if the user needs further help, they can ask the VA for assistance, and get guided to the Exhibit entrance.



Figure 4.58. Example of a screen displayed while indoor

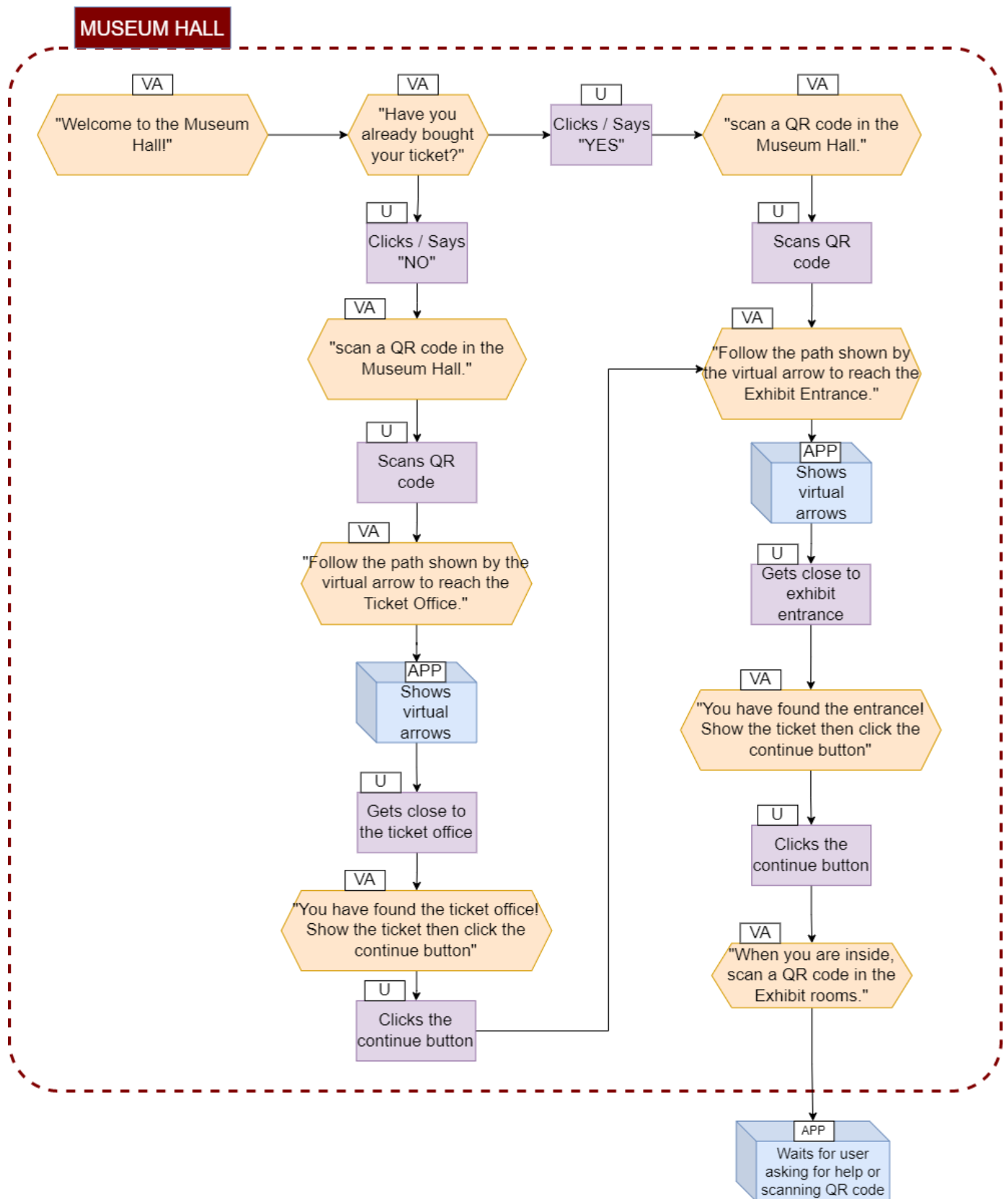


Figure 4.59. Hall AR UX Diagram

Exhibit Rooms

There is no fixed application flow inside the exhibit rooms, because the exploration inside them is completely free and not restrained by specific tasks to complete. That said, it does exist an application flow for the exhibit, but it is completely optional and can be accessed through the interaction with the VA or the VG. The specific tasks given by the VA and the VG are described respectively in the two sections *Exhibit Rooms Help* (4.4.1) and *Virtual Guide and Treasure Hunt* (4.4.1).

Virtual Assistant Help

As already said, the application allows the user to ask the VA for help in any moment of the experience. As shown by figure 4.60, the UX flow of this section is based on locating the user inside the museum and then changing the application state in order to assist the user in the most accurate way.

At the beginning, the VA asks the user where they are giving them four possibilities, three concerning the macro areas of the museum and the last one if the user is completely lost. If the user selects one of the first three, the application changes state according to the macro area selected. If otherwise the user selects the fourth option, the VA simplifies the question changing the possibilities to outside or inside the museum. In the first case the application jumps to the *outside* section exploiting GPS functionalities in order to recover data and position of the user, while in the second case the user gets asked to scan a QR code in order to be located inside the museum, and then the state of the application changes according to the QR code content.

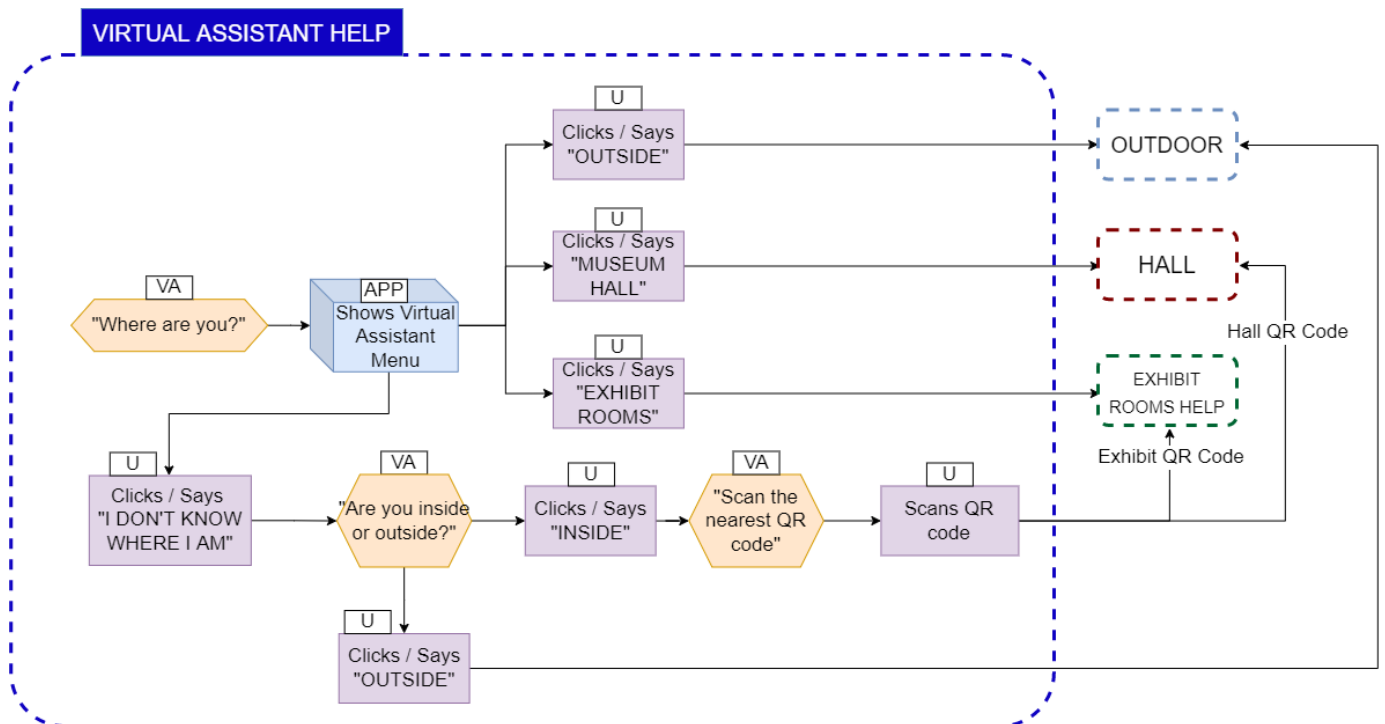


Figure 4.60. Virtual Assistant help AR UX Diagram

Exhibit Rooms Help

This section of the application flow is reached when interacting with the assistant asking about the exhibit rooms. As shown by the figure 4.62, at the beginning, the VA asks the user what they need giving them four answer options:

- **Encyclopedia:** When the user selects this option, the application shows the encyclopedia menu. The flow of the encyclopedia is described in more detail in the *Encyclopedia* section (4.4.1).
- **Next Room:** As this option is selected, the VA asks the user to scan the nearest QR code, in order to make sure that the application knows the latest position of the user. As soon as the QR code is scanned, the application shows the virtual arrow and guides the user to the entrance of the following room where the arrow disappears and the VA gives a feedback to the user.
- **Interactive Activities:** When the user selects this options, as in the previous case, the VA asks the user to scan the nearest QR code, and in the same way, the application guides the user near the VG. The flow of the application in the moment when the user gets close to the VG is described in more detail in the *Virtual Guide and Treasure Hunt* section (4.4.1).
- **More Info:** If the user selects this option, the VA gives them an advice on how to get more information about the subject of the room or a specific artifact from the exhibit.

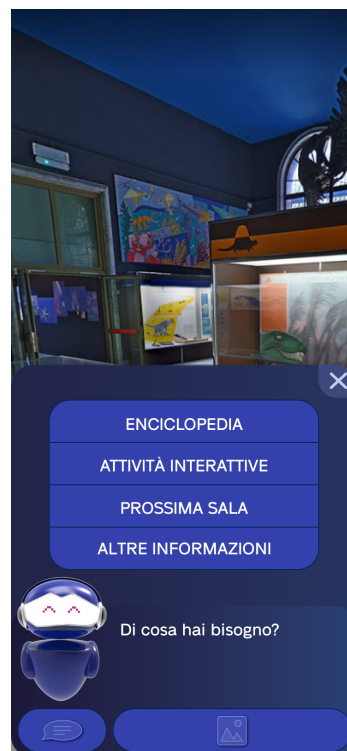


Figure 4.61. Example of help screen inside an exhibit room

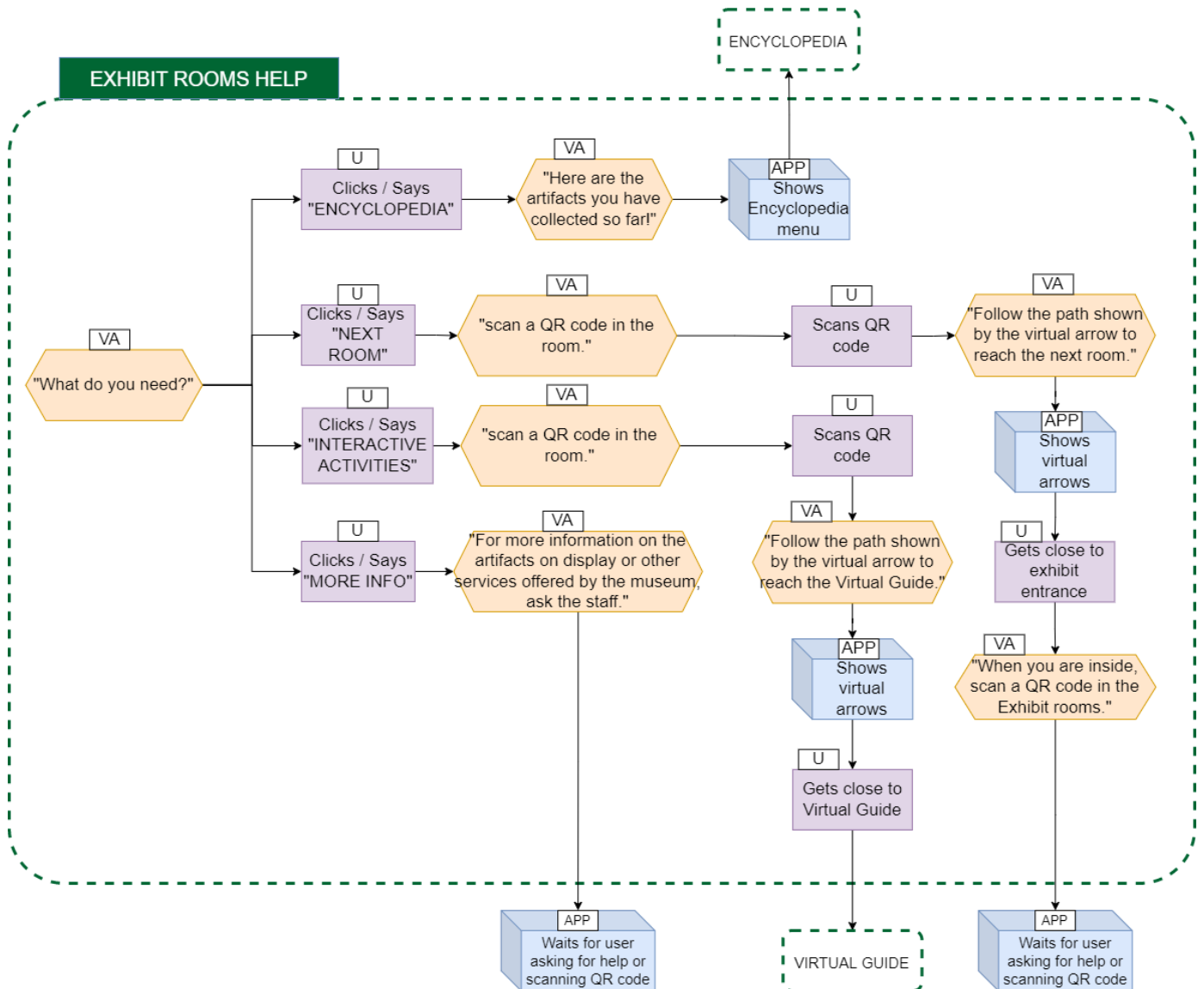


Figure 4.62. VA help AR UX Diagram

Encyclopedia

This section describes the the encyclopedia menu, accessible through the interaction with the VA inside the exhibit rooms. As shown by the figure 4.63, at the beginning of the application flow, a list of buttons with the names of the exhibit rooms of the museum is shown. When the user selects a room from the list, a second list of buttons is shown, this time containing the names of the artifacts that the user discovered through the treasure hunt contained in the selected exhibit room. This list can be empty if the user does not have discovered any artifacts in a room. Finally, if an artifact is selected, the application shows a page with data about the artifact: its name, an image and description. In each state of the encyclopedia menu it is possible to click a *back* button that can be used to return to the previous state or to close the menu.

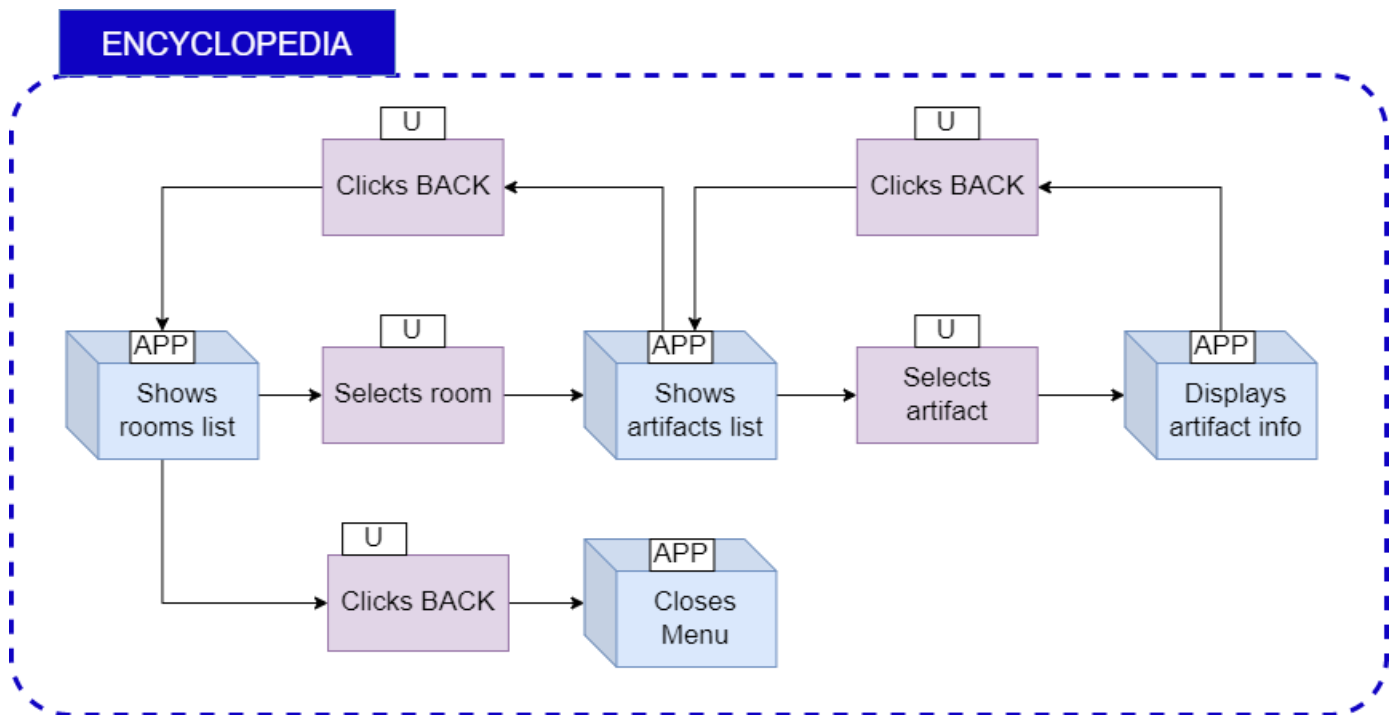
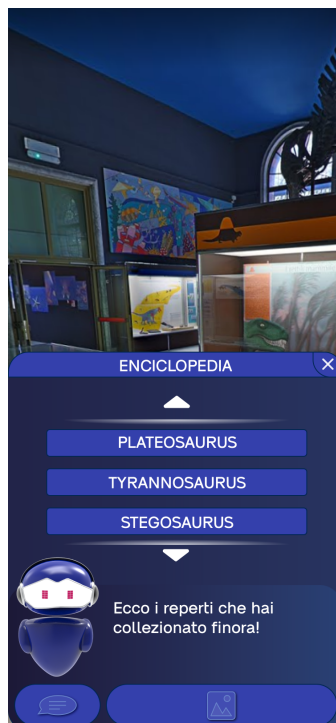
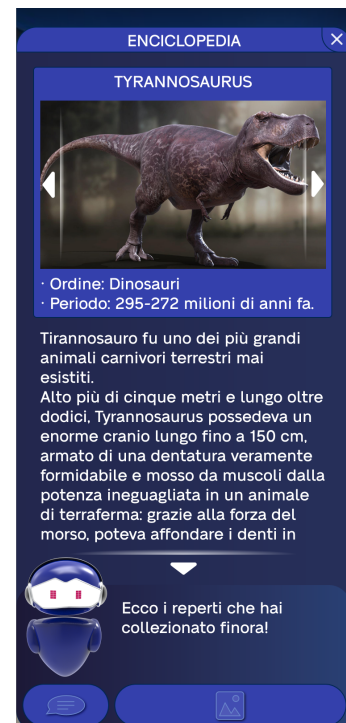


Figure 4.63. Encyclopedia AR UX Diagram



(a) User can choose between the animals they have unlocked inside a room



(b) An animal item is selected: images and informations are shown

Figure 4.64. Screens showing how the Encyclopedia works

Virtual Guide and Treasure Hunt

Whenever a QR code from an exhibit rooms get scanned, the model of the VG gets placed in a specific point of the room. Whenever the user gets near to its model, the interaction with the VG gets triggered, and a clickable banner with a text appears at the top of the screen, as shown by the figure 4.66. The banner disappears as soon as the user moves away.

When the user clicks the banner, the application shows a bigger dialogue box where the VG explains that it is in charge of the Treasure Hunt minigame. From this dialogue box, the user can select the *instructions* button, that shows a textual guide for the minigame supported by explanatory pictures, or he can select the *Treasure Hunt* button.

If the *Treasure Hunt* button is clicked, the application shows a list of the artifacts inside the room where the user is located, and as soon as the user chooses the artifact they want to look for, the hunt starts.

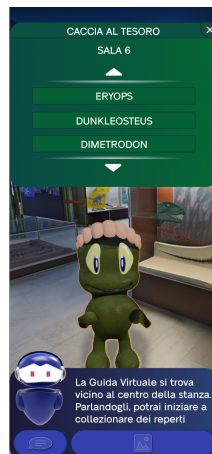
The flow of the application for the Treasure Hunt minigame is slightly different for the *Autonomous* and the *Guided* modes:

- **Guided mode:** In the guided mode, as shown by the bottom part of the figure 4.66, a clue in the form of an image is shown inside the VG dialogue box, while the VA gives feedbacks as the user gets closer or farther from the artifact. During this phase of the treasure hunt it is possible to hide or show the hint image, in order to clear a part of the screen.

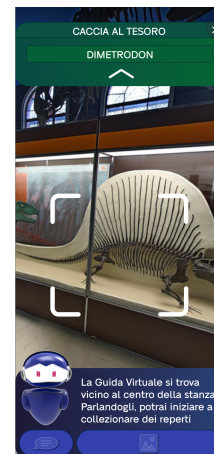
When the user is really close to the position of the artifact, the VA asks him to look around, recognise the artifact and frame it with his device's camera. At this point, if the user frames the right artifact, the application detects it through image recognition, then the VA congratulates the user, the artifact is unlocked in the encyclopedia, and the application shows the details of the artifact with its description. If the user is framing the wrong item, the image recognition cannot happen and the app still asks to scan the selected item.

- **Autonomous mode:** In the autonomous mode, on the other hand, the flow is similar but the user is free to explore the room without the VA giving hints about the location of the artifact. After the hint image is shown, in fact, the VA only asks the user to frame the artifact with his device's camera once they find it in the room. For the remaining part, the flow is identical to the guided one.

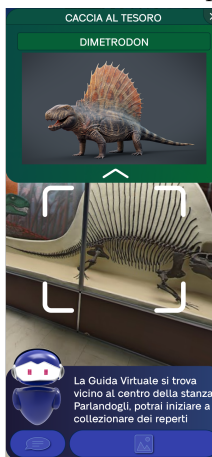
In each state of the VG menu it is possible to click a *back* button that can be used to return to the previous state. If the *back* button is clicked while the menu is displaying the details of the artifact just discovered, the application returns to the state that allows to choose an artifact for the Treasure Hunt.



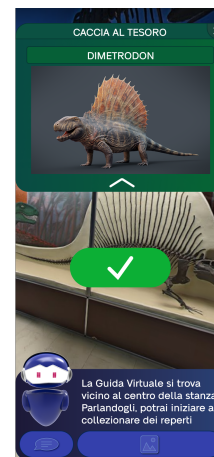
(a) A list of four items is presented



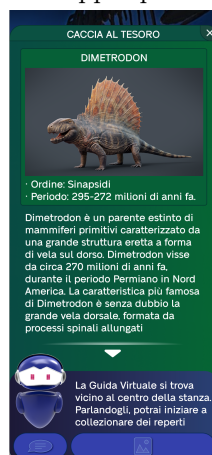
(b) An animal from the list of items is selected



(c) An helping image can be viewed by clicking on the arrow in the upper part of the screen



(d) The item is framed correctly and a check mark is shown



(e) Informations about the animal is displayed. The informations are also saved in the Encyclopedia

Figure 4.65. The Treasure Hunt minigame

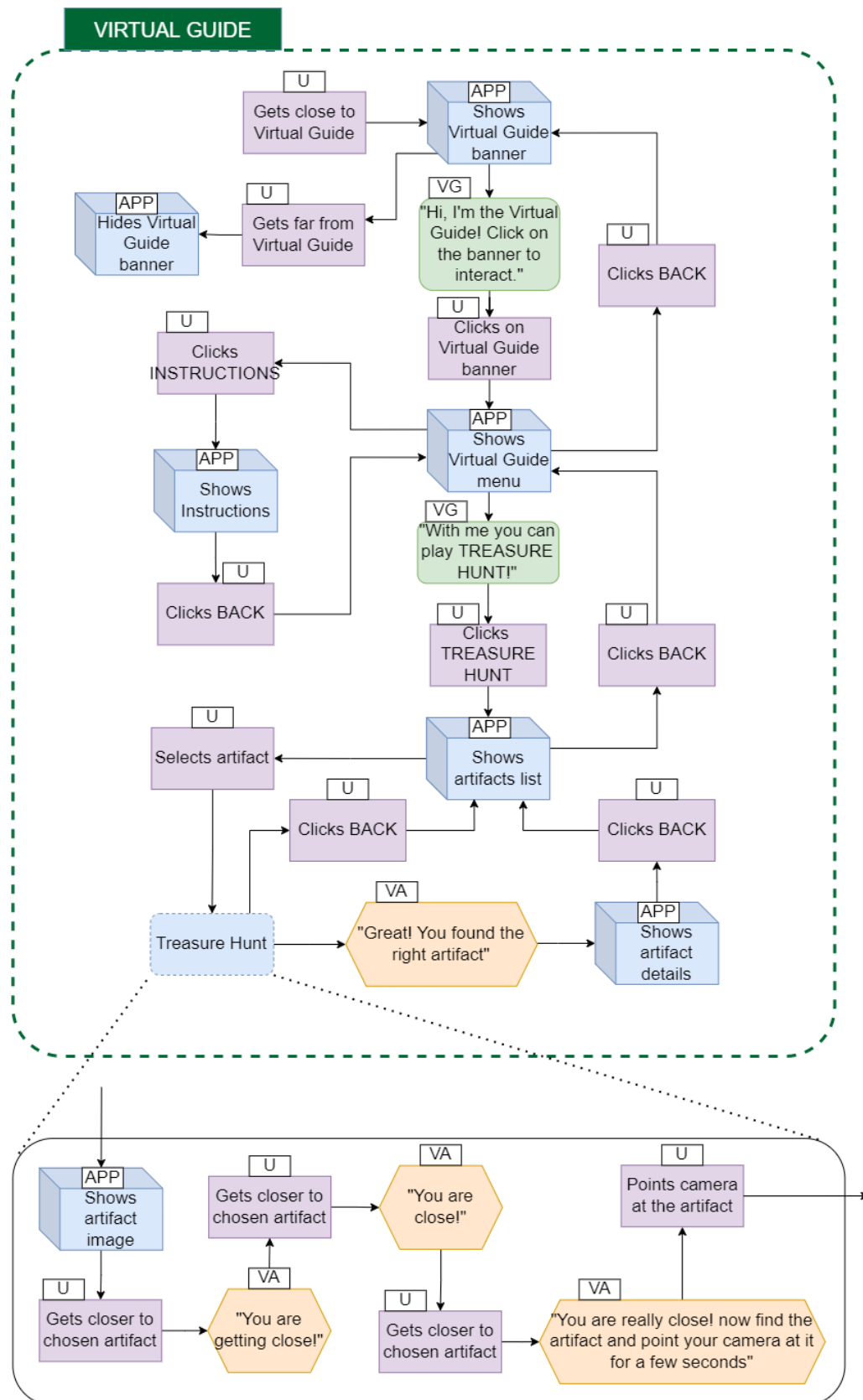


Figure 4.66. Virtual Guide AR UX Diagram

4.4.2 Interactions

In the AR environment, the user has the possibility to interact with virtual and real objects in multiple ways. So, unlike the VR experience, in the AR one, it is possible to modify the states of the objects in the virtual environment both by interacting with the real and the virtual worlds.

The basic interaction, that is also the most common one in the application, is tapping on the screen to select buttons or to interact with the VA or the VG. Each time the conversational agents ask questions with several possible answers, the application shows a list of buttons that can be selected by tapping on them. Beyond that, in order to interact with the VA or the VG, the user needs to tap on the 2D image of the Assistant in the first case and tap on the guide banner in the second.

Vocal Interactions

The other main interaction with the virtual world are the vocal interactions with the VA. Vocal input is activated by pressing and holding the button under the VA. While holding it, the button enlarges and become bright and the Assistant's eyes turn into a microphone shape to communicate the fact that the user is now speaking to him. The microphone is also a callback to the microphone used in the VR applications. Figure 4.67 shows the user holding the Vocal input button and saying they are "outdoor". The application responds if the user asks:

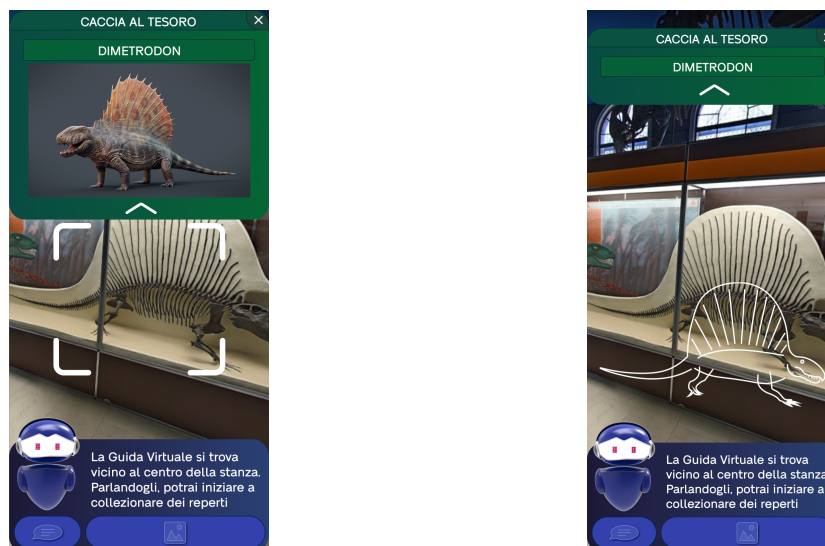
- A general *Help*, to which it opens the general Help menu;
- *Where they are*, to which it first ask if the user is outside or inside, then directs the user to the Entrance with GPS (the former case) or asks him to scan the nearest QR code (for the latter);
- *What they have to do*, to which it responds in the same manner as the last one, because the application can't be sure if the tracked position of the user is still accurate.

Image Recognition

Interaction with the real world is made possibly through Image-Based Recognition technology. This technology is used to recognize the museum Entrance and while the user is playing Treasure hunt and is requested to frame the correct items. The functioning of image recognition is explained in 5.5. Image recognition works only by the angle the image has been submitted into the application, albeit still having a certain degree of freedom that allows the recognition to work from a restricted degree of angles. To suggest to the user the correct angle, when the image recognition is active, the user can view an image of the item from the correct angle. Other methods to help the user in framing correctly the item were evaluated as valid but not implemented for time constraints (see Chapter 7). A few seconds are necessary to recognize it, then when the process is done, a check mark is shown as a feedback to the user.



Figure 4.67. Holding the Vocal input button



- (a) As developed in 5A: an image from the correct side of the animal serves as a suggestion for the user
- (b) Option for future work on 5A: a silhouette of the animal serves as a suggestion for the user (see section 7.3)

Figure 4.68. Used and proposed solutions to give suggestions to the user

Movement Interactions

Finally, the last way to interact is by moving in the real environment. When the user moves, in fact, the movements of their device gets tracked and are used to translate a user indicator on a virtual map. Sometimes during the experience, the user is asked to reach a certain place of the museum, like the ticket office or the exhibit entrance; in these cases, the movement of the user is followed by the movement of the indicator on the map that, when it reaches the target, it triggers a change of state in the

application.

4.4.3 Levels of difficulty

As already introduced in the previous sections, the AR application provides two different levels of difficulty, which differ in terms of how much support the user receives and how strict the experience flow is. Even if the details of their differences in each macro area of the experience flow have already been given in the *UX Flow diagram* section (4.4.1), here we want to discuss the main choices that were made during the design of the two difficulty levels:

- **Guided mode:** In guided mode, the user is assisted step by step during the whole experience by the VA. The assistant gives the user the tasks and waits him to complete them before moving to the following ones. The only exception is the exhibit rooms section, where the user is given more freedom in order to let him explore the room on its own, and the additional information and the treasure hunt are accessible through the interaction with VA and VG.

Since the VA gives the user all the possible help during the basic flow of the experience, asking for help to the Assistant does not give any additional support (except in the exhibit rooms). However, if the user has problems following the flow, for example skipping a part, lagging behind, or simply not needing help in a specific section, he can ask for help to the Assistant which will give him context-related guidance.

- **Autonomous mode:** In autonomous mode, on the other hand, the VA supports the user during the experience giving him hints on the following steps he needs to take. For example, when the user scans a QR code inside the hall of the museum, the VA tells them that the first step is to find the ticket office in order to buy a ticket. However, if the user asks the VA for help, they get access to a more deeply assisted guidance, similar to the one for the guided mode.

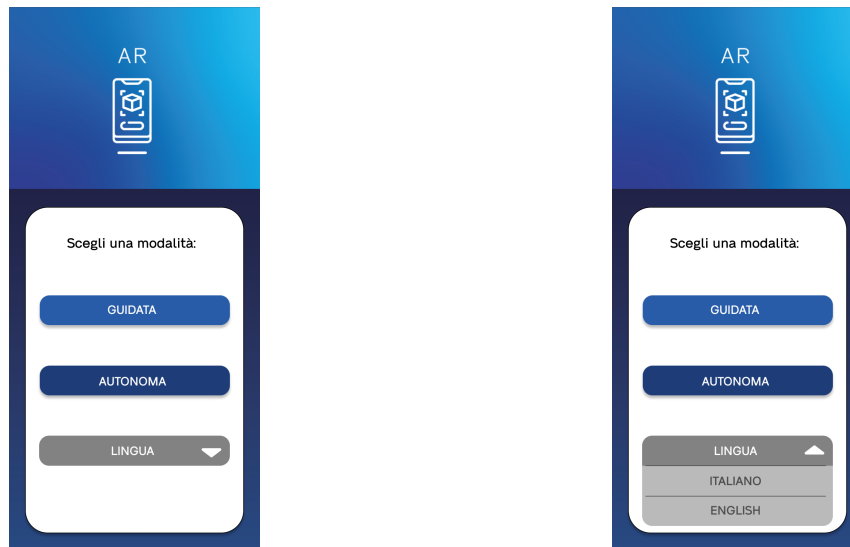
4.4.4 User Interfaces

Main menu

When the user opens the application, they are welcomed by a splash screen with a placeholder logo for the application, followed after some second by the Main Menu. The Menu is fairly simple and presents two blue buttons of different shades, one for the Guided Experience and one for the Autonomous Experience; a third grey button on the bottom opens a dropdown menu that allows the user to choose the language (in this case only between Italian and English) On the upper part a vector graphic displays a simplified image of a AR application on smartphone. The overall colors are a degree of the shade of blue (HTML code #242A52) used in both VR and AR applications as described in section 4.2.

Dialogue boxes

The dialogue boxes are the place where buttons and text from both the VA and VG are showed to the user. Regarding the VA, it itself is part of the interface on the



(a) The menu showing the two buttons for Autonomous and Guided experience

(b) The dropdown menu for choosing the language (english or italian)

Figure 4.69. Menu screens

dialogue box. In fact, this is very different from 5A on VR where the Assistant is a three dimensional model following the user. This decision was made for a number of factors:

- the difficulty in positioning the Assistant in a place where it would be easily visible and present during the experience, while at the same time being in a comfortable position in respect to the user
- the possibility that at certain moments it would cover too much of the screen disrupting the experience;
- due to inherent problems regarding the precision of the AR objects in the screen in respect to the real world, the possibility that the Assistant would disappear from the screen or be positioned in an unwanted area difficult to reach, in turn creating discomfort for the user
- as the technology required to position the Assistant would be plane recognition, we understand that finding an open plane without disruption, especially in a crowded place like a museum, could be difficult for the user.

For this reasons the VA in the AR application is a 2D sprite, complete with animations and a three dimensional style. It was, in fact, derived directly from pre-baked animations of the same 3D model of the character used in the VR application. The VA sprite is positioned at the left side of the screen and seems to hover thanks to a idling animation that goes up and down. Two other animations are present:

- when the user taps on the Assistant, it moves forward on the screen and the eyes inside its visor change shape to assume a friendly expression;

- when the user chooses the vocal interaction by pressing the button under the Assistant, its visor shows a microphone to let the user know they can now talk to him.



Figure 4.70. The three animations of the VA sprite, respectively: Idling, Helping and Vocal input

The colors used for the background are a degree of the shades of blue (HTML code #242A52). The dialogue boxes have two sizes: a smaller one for the standard text, and a vertically enlarged one when the app shows Facilitator graphics or when the user asks for help and the additional buttons about the areas they are in are shown. This solution of two sizes was chosen because having always a big dialogue box on the screen would impede the user's vision through the camera.

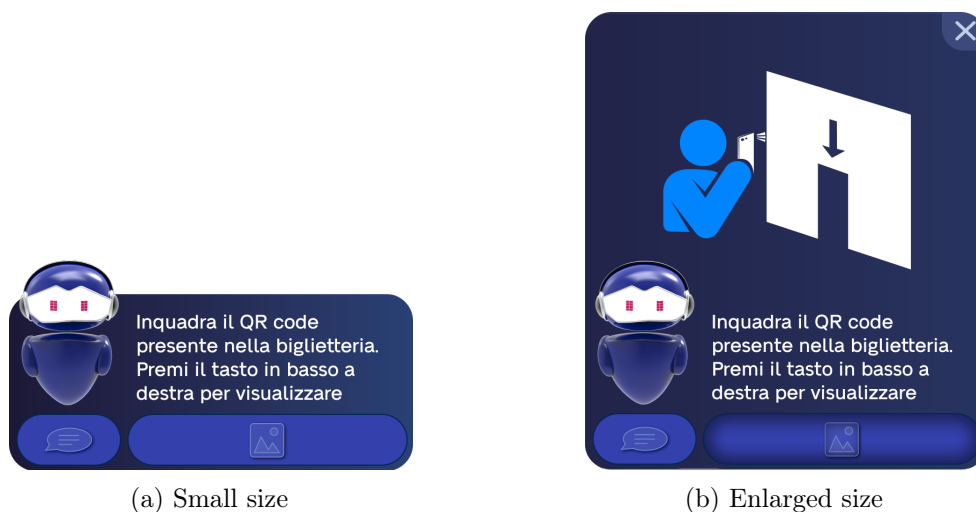


Figure 4.71. The two dimensions for the VA dialogue boxes and menus

The text is white for better visibility and is positioned at the right side of the box; it has medium-sized characters (18 pts) to be readable for everyone. Longer phrases are handled by having a scroll function on the text, that the user can scroll up and down by sliding with their finger on it. Buttons are positioned at the bottom. The one under the sprite of the Assistant is for activating the microphone, the one under the text is for showing the Facilitator images. The icon used are descriptive of the button function, and were revised with the help of the therapist. In one occasion, when the user is buying the ticket, a third button appears at the center of the screen: its function is to communicate that the ticket has been bought.

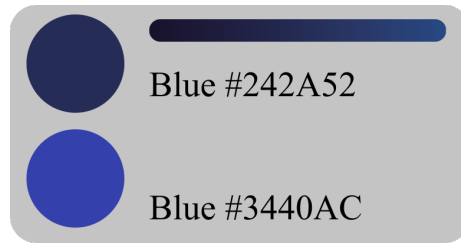


Figure 4.72. Color palette used for the VA dialogue boxes

The VG also talks through dialogue boxes. To differentiate them from the VA's, they are positioned on top of the screen and their main colors are degree of shades of green (HTML code #065E39). They also come in two sizes, a smaller one for simple text and an enlarged one for when the Guide is giving the instructions about the minigame or when the list of items of the Treasure Hunt is showed.

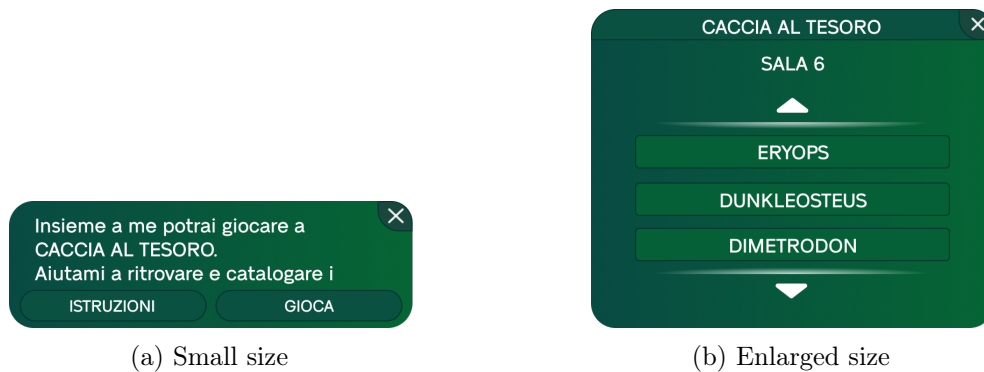


Figure 4.73. The two dimensions for the VG dialogue boxes and menus

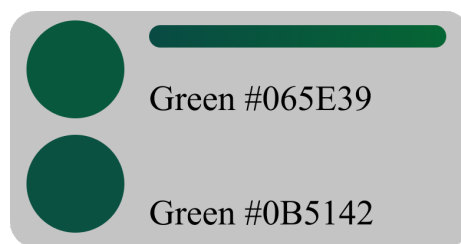


Figure 4.74. Color palette used for the VG dialogue boxes

The overall shapes of the dialogue boxes and buttons is a rounded and regular style. The colors are solid for the buttons while the background of the dialogue box has different shades of blue.

Facilitator graphics

Facilitator graphics are used in the app to assist the user in a visual way, as described in section 4.2.5. In the AR application are used in two types of interactions:

- during tutorial or instruction screens, as a complementing part of the written text;

- during conversations with the VA or the VG;

In the first case they are shown alongside the text, while in the latter they can be opened using the button on the bottom right and closed by an *X* button. Once opened, a pop-up extension of the dialogue box appears on the upper side, showing a context-related image that helps the user to know the task they have to do.

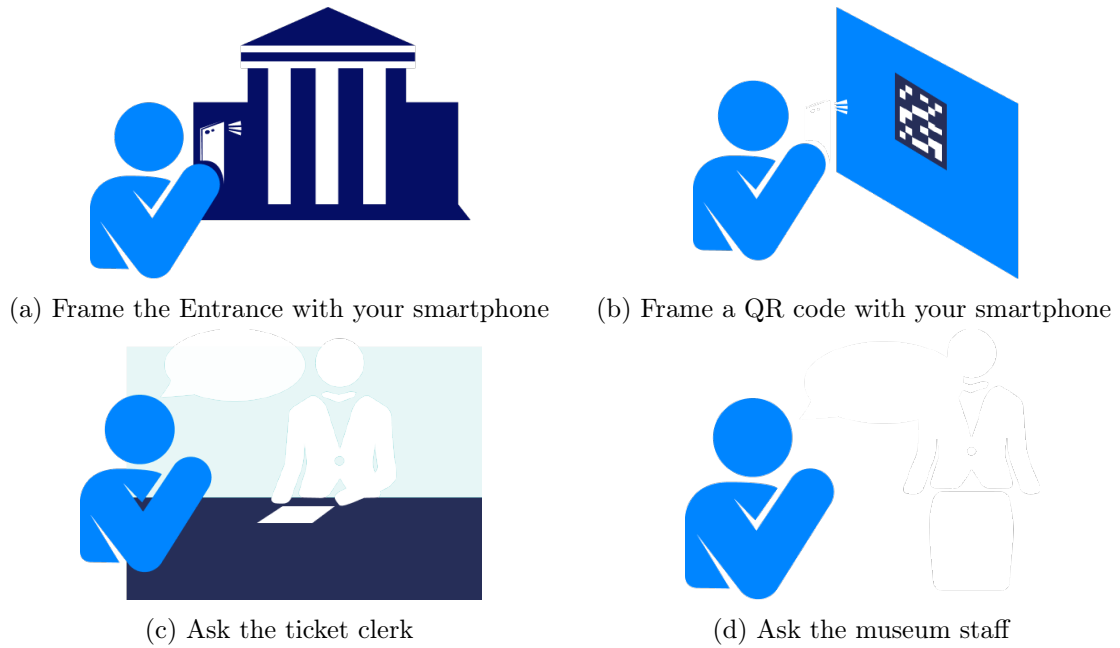


Figure 4.75. Facilitator graphics used in the AR app

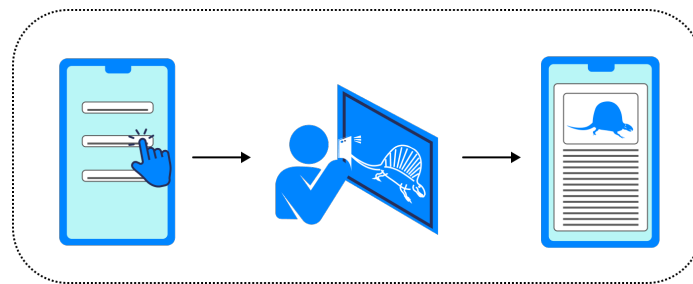


Figure 4.76. Instructions to play the Treasure Hunt



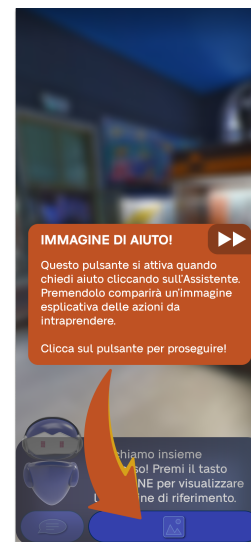
(a) How to ask for help



(b) Where to see the text advices



(c) How to talk to the Assistant



(d) How to display the Facilitator graphics



(e) How to frame the QR codes

Figure 4.77. Tutorial screens

Tutorial

Tutorial screens are a bit different from the rest in the color choice. To clearly state that the image shown are only indications on what the user will have to do, and not part of the actual UI, the color chosen is orange for its high contrast with blue. The style used is also a bit different a less regular if compared with the rounded but regular style of the rest of the UI. To go forward the user is requested to tap or use the current feature of the app that is being explained, in this way training the user to the movement. In any case, all the screens can be skipped by tapping on the arrows on the upper right corner. The tutorial screen are shown in figure 4.77

4.4.5 User Tracking

In order to give support to the user in the most accurate way possible, the application must be able to keep track of the user location throughout the whole experience. In order to find the best technology to solve to this problem, we have explored several options:

- **GPS positioning:** a positioning technology that relies on GPS features in order to locate and track the user's device. It can be used freely and it is available on every modern smartphone, but since it relies on signal coming from satellites, it is not extremely accurate, especially indoors, and its usage drains device's battery fast.
- **Bluetooth low energy Beacons:** Beacons are battery-driven devices that rely on Bluetooth connection in order to communicate with the user's device. The device detects the Beacon's signal and can roughly calculate its distance from it, allowing it to estimate a user's indoor location. Beacons are cheap, small and have a long battery life but they are not extremely accurate, so they must be installed in large numbers inside each museum in order to be able to overcome this flaw. Furthermore, this technology can be difficult to be used outdoors because the Bluetooth signal is not particularly strong. Figure 4.78 shows a Beacon device.



Figure 4.78. Beacon technology

- **WiFi Positioning System:** WiFi can be used in the same way as beacons and its signal is even stronger (even if its accuracy is a bit lower), but it must be connected to a power source and it needs an internet provider and more

expensive hardware to be set up. However, it can be assumed that nowadays almost all museums would have their own public WiFi network already installed.

- **Ultra wideband:** it is without doubt the most precise indoor positioning system on the market, but it needs a more expensive equipment in order to be installed compared to the other technologies described. The main problem of this technology is that the number of smartphones that support it are really few and also really expensive. A solution to make the technology accessible by everyone is to use locator tags, small electronic devices that, when given to the users, get located by the ultra wideband system and share their position with the users' smartphones. However this solution also raises the cost needed to implement this technology.
- **QR code positioning and device tracking:** this solution involves the exploitation of two different technologies used in parallel. The first one is the QR code, a type of barcode that contain information as a series of pixels in a square-shaped grid, and can be easily read by all modern smartphones. The second one is device tracking, a feature implemented by AR platforms like ARCore and ARKit, and supported by almost all modern smartphones, that allows to track the device's position and orientation in physical space.

The main idea behind this solution is that the QR codes would be used to contain their own position and orientation, and communicate it to the user's device when they get scanned, while device tracking would update user's position and orientation as they move around the museum.

- **Image recognition positioning and device tracking:** really similar to the previous solution, but instead of using QR codes, the information relative to the position and orientation of the user would be recovered through image recognition, a feature implemented by ARCore and ARKit like the device tracking. The main issue of this solution is that all the images to be recognised and the information related to them must be contained in the application a priori, a problem that severely limits the scalability of the application.



Figure 4.79. 5A-Museums QR codes

After an in-depth research of all the pros and cons of these technologies we decided to use two different solutions for the outside and the inside of the museum. For the

outdoor environment it was opted for the GPS solution, because a high accuracy was not needed and because of its affordability and accessibility. For the inside of the museum, on the other hand, it was opted for the QR code positioning, because since QR codes are used to store data, it would have been possible to recover all the information about the museum rooms, besides the location, directly from scanning them, greatly boosting the scalability of the application. Another reason is that, if the position and rotation of the user gets often refreshed by QR code scanning, device tracking is able to provide a high level of accuracy inside the rooms of the museum.

An in-depth analysis of the components providing QR code scanning, device tracking, indoor mapping and indoor navigation is provided in the *Implementation* chapter, under the *AR architecture* (5.5). Figure 4.79 shows the style of the QR codes used. To make them more visible, their background was edited to match with 5A graphical themes.

Chapter 5

Implementation

5.1 Unity

Both the VR and AR experience have been developed using Unity. Unity is a cross-platform game engine released in 2005 by Unity Technologies. It is widely used in professional video-game industry and it's generally the most chosen engine for indie developers, especially when they have to develop for mobile devices. Unity gives users the ability to create games and experiences in both 2D and 3D, and the engine offers a primary scripting API in C#. Moreover is surprisingly easy to use because of its drag and drop system for game objects inside the environment.

It was decided to use Unity mainly for three reasons:

- Support for extended reality technologies: Unity offers some default plugins to easily work with virtual and augmented reality. These plugins give an already modeled interface to interact with XR devices in a pragmatic way, offering a complete control on input systems and sensors. These Plugins will be described in detail in the *Libraries and Plugins* section (5.1.1).
- Large range of supported devices: the engine gives the possibility to build an application on several devices with distinct hardware and operating systems, making an eventual porting process easier and faster. In this way we can avoid to re-implement from scratch a new version of the applications for new systems. This can be very useful, especially considering the enormous variety of Oculus machines actually available on the market and the big differences even between the two main mobile operating systems (iOS and Android).
- Reusability of 2D and 3D assets: Unity allows the re-use of assets in different applications build on it. Because 5A VR application and AR application share transmedial elements, this is an especially important point.
- Free plan: Unity offers several plans according to the project revenue. In our case the free plan, available for all products that don't overcome the threshold of 100K \$/year, covered all our needs in terms of development tools.



Figure 5.1. Unity IDE

5.1.1 Libraries and Plugins

XR Interaction Toolkit

XR interaction toolkit is a high-level, component-based, interaction system for creating VR and AR experiences. It provides a framework that makes 3D and UI interactions available from Unity input events. The core of this system is a set of base Interactor and Interactable components, and an Interaction Manager that ties these two types of components together.

Used version: 1.0.0-pre.4 - May 19, 2021 [40]

XR Plugin Management

Unity Package that provides simple management of XR plugins. It manages and offers help with loading, initialization, settings, and build support for XR plugins.

Used version: 4.0.7 - July 12, 2021 [41]

AR Foundation

AR foundation is a framework purpose-built for augmented reality development that allows to work with AR platforms in a multi-platform way within Unity. This plugin presents an interface for developers to use, but does not implement any AR feature itself. In order to use AR features, AR Foundation needs separate packages for the target platforms officially supported by Unity. In our project we integrated the functionalities of ARCore and ARKit, in order to build the application both on Android and Apple devices.

Used version: 4.1.7 - April 08, 2021 [42]

ARCore XR Plugin

This plugin is used by AR Foundation to enable ARCore support via Unity's multi-platform XR API. ARCore is Google's platform for building augmented reality experiences, designed to work on a wide variety of android devices running Android 7.0 and later. This version of ARCore XR Plugin uses ARCore 1.24 and supports functionalities like device localization, device tracking, plane detection, anchors and 2D image recognition and tracking.

Used version: 4.1.7 - April 08, 2021 [43]

ARKit XR Plugin

This plugin is used by AR foundation to enable ARKit support via Unity's multi-platform XR API. ARKit is the platform designed by Apple for the development of Augmented Reality applications for iOS devices running iOS 11.0 or later and having an A9 or later processor. This version of ARKit XR plugin supports the same functionalities listed previously in the description of ARCore XR Plugin.

Used version: 4.1.7 - April 08, 2021 [44]

Oculus XR Plugin

Provides display and input support for Oculus devices.

Used version: 1.9.1 - May 11, 2021

Localization

The localization package provides tools for adding support for multiple languages and regional variants to the application, for example supporting text in multiple languages or culture-specific assets such as audio or textures.

Used version: 1.0.5 - October 28, 2021

TextMeshPro

The TextMeshPro package replaces Unity's UI Text. It employs Advanced Text Rendering techniques, as well as a collection of custom shaders, to improve visual quality while providing users more control over text design and texturing.

Used version: 3.0.6 - April 22, 2021

5.1.2 Packages and Assets imported from external sources

ZXing.Net

ZXing.Net is a library which supports decoding and generating of barcodes (like QR codes, PDF 417, EAN, UPC, Aztec, Data Matrix, Codabar) within images.

Imported from:

<https://github.com/micjahn/ZXing.Net>

Speech And Text

Speech And Text is a package that provides a simple interface to interact with Android and iOS native Text-To-Speech and Speech-To-Text functionalities.

Imported from:

<https://github.com/jimmyto9/Speech-And-Text-Unity-iOS-Android>

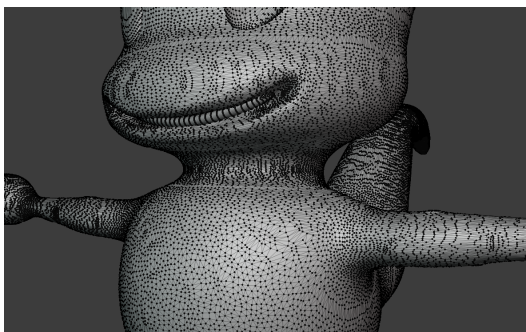
Imported from:

<https://github.com/Unity-Technologies/XR-Interaction-Toolkit-Examples>

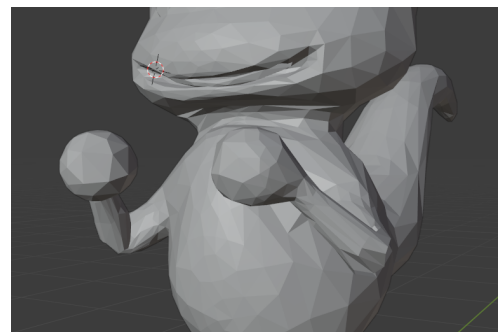
5.2 Blender

Blender is a free and open-source 3D computer graphics software. It's used widely in the gaming, animation and visual effect industry. It was chosen for its flexibility, ease of use, wide variety of exporting formats that are highly compatible with Unity. Blender supports different types of modeling, the main ones being polygon modeling and sculpting. Polygon modeling is a more precise, albeit slower, method of modeling. It's ideal for geometric shapes but can also be used for complex shapes. Sculpting allows much more freedom during its execution and works similar to shaping a piece of clay. For this, it's perfect for organic shapes and free forms.

Figure 5.2 shows an example of sculpting.



(a) Sculpting modeling used on the Guide



(b) The same model after applying subdivision surfaces

Figure 5.2. Guide's body is sculpted

Polygon modeling was used for geometric shapes like the museum environment and the body of the Assistant, while sculpting was chosen for the more rounded and complex shapes like the Virtual Guide. Figure 5.3 shows the clearly different body parts of the Assistant after they were merged and subdivided. Each model was initially composed of different parts, that were then merged together to allow animations and exportation on Unity.

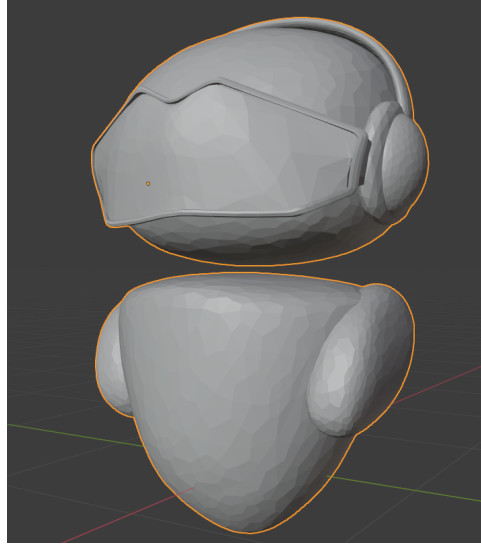


Figure 5.3. Assistant's body is poly modeled

Subdivision surfaces were also used to keep the poly count low and facilitate an easy export to Unity, as well as guarantee an acceptable amount of FPS on the application. As shown in figure 5.4, textures were initially applied on Blender and then baked prior to export on Blender in order to create a single texture file: this permits a smooth and less cumbersome output on Unity. In regards of the animations, they were also created on Blender by rigging the characters (Human avatars, Virtual Assistant, Virtual Guide) and creating the necessary poses and key frames. Each character has a set of animations that include idling, walking, jumping. Animations were also created for the NPC that walk through the environment. Figure 5.5 shows an example of rigging and export the animations on Unity. The export format used from Blender to Unity was *.fbx*, a format that acts as a library that saves information about the polygon mesh, textures and animations.



(a) Original textures applied on the Guide

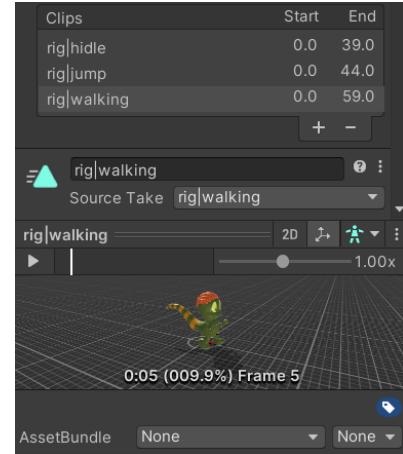


(b) Baked texture optimized for exporting on Unity

Figure 5.4. Poly modeling vs. sculpting in shaping a head.



(a) The rig on the Guide



(b) Animations list on Unity

Figure 5.5. Animations were made by rigging characters and then exporting to Unity

5.3 Speech-to-text and Text-to-speech

Both the VR and AR application use STT (Speech-to-text) and TTS (Text-to-speech) to improve users' communication ability and ability to listen, as well as improving the feedback of the system. TTS is used to transform strings in vocal audio, while STT is used to record users voice and recognize vocal input. The main problem in the implementation of the two technologies was the availability of a prefigured engine for the chosen device. These devices (Oculus Quest and smartphones) run both on Android but Oculus doesn't support the native Android vocal synthesis functionality. So two different vocal synthesis engines have been used for the two applications.

5.3.1 Microsoft Azure STT and TTS

For the VR application, running on Oculus, Microsoft Azure services have been chosen. Azure is a cloud computing service operated by Microsoft for application management via Microsoft-managed data centers. It provides software as a service (SaaS), platform as a service (PaaS) and infrastructure as a service (IaaS) and supports many different programming languages, tools, and frameworks, including both Microsoft-specific and third-party software and systems. We opted for Azure services mainly for its scalable price plan and for the easy integration with Unity [45].

5.3.2 Android and iOS native STT and TTS

Both Android and iOS have native Speech-To-Text and Text-To-Speech functionalities that are supported by Android and iOS operative systems. Since the service is native to both systems, its use is completely free and unrestricted as long as the application is built on Android or iOS devices.

In order to implement our functionalities of STT and TTS, we used a third-party plugin (5.1.2) that provided us a tool to interface more easily with these services. As regards Text-To-Speech, both systems offer a voice whose language, pitch and rate can

be changed at any time. On the other hand, as regards Speech-To-Text, the language to be recognised can be selected and changed in any moment.

5.4 5A-Museums - VR architecture

The VR application is composed only by a fat client running on the Oculus device. Even if the application executes all the main processes locally, it needs a stable connection to communicate with Microsoft Azure servers. This is due to the Speech service free plan that have to count the used characters in the application and keep it updated on Microsoft servers.

5.4.1 The modular structure

Scenes

To maximize the scalability of the system and to guarantee good performances (especially considering the Oculus big limits in terms of hardware), the application has been implemented with a modular structure. Each scenario (described in the design chapter) corresponds to at least one Unity Scene. So we can distinguish five Unity Scenes:

- *MainMenu*: it's the starting scene of the application. From here users can customize the experience by setting its language, difficulty and starting point;
- *OutdoorScene*: this is the default starting point of the experience. It corresponds to the Outdoor scenario described in the design chapter;
- *MuseumHall*: the first indoor room of the museum, where we can find the ticket office (see Museum Hall Scenario);
- *Room1*: where the exhibition starts. In this room we can observe some items related to marine specie (see Exhibition room Scenario);
- *Room2*: a secondary exhibition room. This is the dinosaurs exhibition room (see Exhibition room Scenario);

With this structure, extending the application with new rooms or scenarios becomes really easy: in fact, the extension process only requires to add new Scenes to the project.

Checkpoints

Every scene is also characterized by one or more checkpoint, that are used from the application to establish the next user goal or to load the right Assistant dialogue. This means that not all the checkpoints correspond to a real goal. Some of them are only used to load the right dialogue at the right time All these checkpoints can be found in the enumeration class `Checkpoint.cs`

5.4.2 Main components

We can identify several main components in the application. Most of them are in common for all the scenes.

MyXrRig2.0 and XR Interaction Manager

These are the main objects of all the application. The first one contains the main camera of the application and transforms all the Oculus input in user actions inside the game. We can classify two different types of script inside the object:

- scripts that manages the Oculus sensors and movements (Move Providers, Turn Providers, ecc...)
- scripts that manages the controller inputs, contained inside the LeftHand and RightHand objects

The XR Interaction Manager contains a reference to the input command scheme, that is the buttons that the users have to press to interact with the virtual environment. We can find an instance of the XR Rig and the InteractionManager in every Unity Scene.

Game managers

The application flow is managed through several type of game managers:

- **GameFlowManager**: this is a singleton class containing all the data needed from the beginning to the end of the experience such as the difficulty parameters, the chosen language, the actual active scene and checkpoint. Every game manager inside the scenes retrieves data from this class.
- **GameManager**: the game manager is an abstract class with three different implementation:
 - **OutdoorSceneManager**: manages the outdoor scene.
 - **HallSceneManager**: controls the hall events.
 - **InternalRoomGameManager**: supervises the actions inside the exhibition rooms. Every scene containing an exhibition room has its own manager. Since the interactions and the events inside all the exhibition rooms are the same, there is no need to create a specific manager for every room.
- **AudioManager**: every scene contains an audio manager. As the name says, the AudioManager component manages all sound events of the scene that are not related to the speech or dialogues of the assistant.

Virtual Assistant

The virtual assistant is the first transmedial element of the application and the conversational agent of the application. It's contained in the *MyXrRig2.0* object so that it can follow the user movements everywhere. If the user goes near to a wall or to another obstacle, the assistant rotate around the camera, always staying in a predefined distance range from it. Since it's an interactable object, it's main controller script implements the *XRBaseInteractable* abstract class (contained in the XR Interaction toolkit package). This allows to use the interaction methods, like the *OnSelected*, *OnActivate* and many other, directly inside the controller.

Virtual Guide

The Virtual Guide is the second transmedial element of the application. Every internal room scene has a Virtual Guide. Even if the guide is an interactable object just like the virtual assistant, its structure is a little bit different: in fact its main controller, the *Guide* script, does not implement directly the *XRBaseInteractable*. Instead, a supplementary script has been added to the Virtual Guide object, the *XRSimpleInteractable*, giving to the Guide the same interaction methods of the Virtual Assistant.

DialogueManager and Speech

The *DialogueManager* component works on the resources directory, loading the dynamic dialogues when needed. The *Speech* script is a static class that manages all the speech-to-text and text-to-speech operations. [5.4.4]

Museum exhibition items and Encyclopedia

Inside the two exhibition rooms we can find several interactable objects: the exhibition items. These items are composed by an interactable object and a descriptive panel. Every item has a controller that implement the abstract class *MuseumItem*. We can distinguish between:

- Paints objects: these are the most common items in the first room. They're managed through the *Paint* class.
- 3D Models: managed by the class *Item3D*

All the items data are contained in the encyclopedia of the application. For more information about encyclopedia and museum items see section 5.4.3.

Vehicles and crowd

Vehicles and crowd are the main distractor [4.3.5] elements of the experience. They both are managed by the *Vehicle* script during their movements (the only difference between the two categories is the movement speed). They implement really simple AI algorithms. Inside the scenes some invisible cubes have been placed: the *PathPoints*.

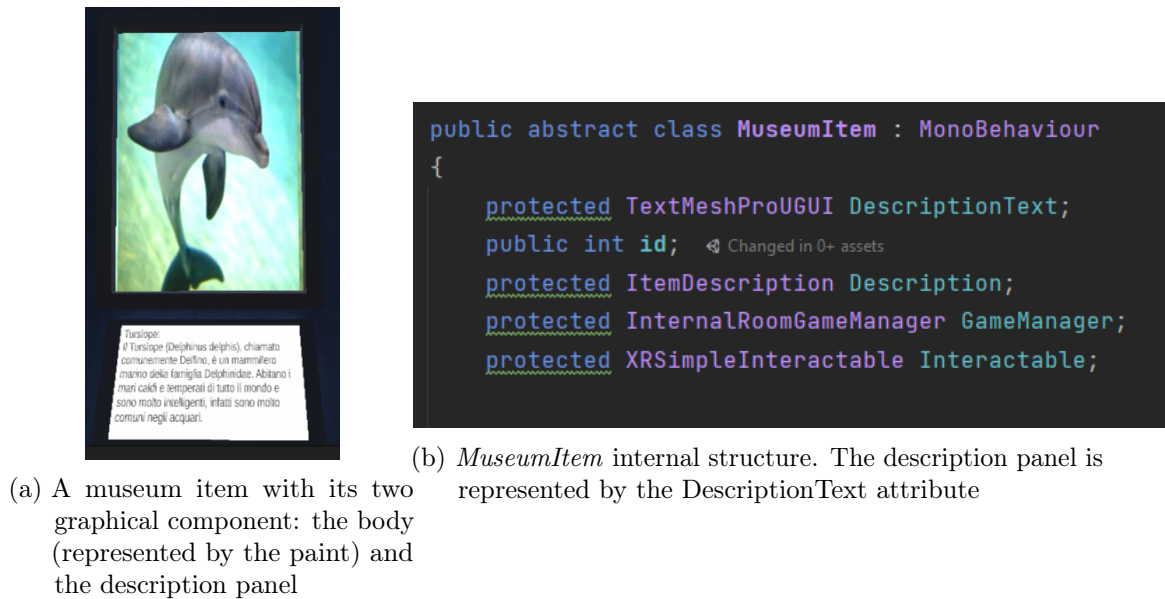


Figure 5.6. Museum items

Every *PathPoint* contains the reference to all the next possible *PathPoints*. All vehicles and walking NPCs (Non-Player Characters) contain a reference to a target *PathPoint*. When a vehicle or a walking NPC arrives to its goal, a new random *PathPoint* (among those referenced by the target *PathPoint*) is chosen and the vehicle/NPC will start moving toward the new goal.

Moreover, in the *FixedUpdate* method a *Spherecast* is launched to check that no obstacles are placed in front of the moving object. If the *Spherecast* retrieves an obstacle, the moving object simply stops its movement until the obstacle is outside the *Spherecast* range.

5.4.3 Gamification process: Treasure Hunt

One of the main feature of the application is the presence of a mini-game inside the exhibition rooms: the Treasure Hunt game. The game begins when the user selects the Treasure Hunt option from the Virtual Guide dialogue panel.

The museum items structure

The "treasures" of the Treasure Hunt game are the exhibit items inside the rooms. As already mentioned in 5.4.2 the items objects are basically composed of an interactable body and a descriptive panel. The interactable body always contains the *XRSimpleInteractable* script component and an implementation of the *MuseumItem* script. The main functionality of the *MuseumItem* is to enable and disable the description panel and the emission of the materials related to the body. The main difference between the *Paint* objects and the *3DItem* objects is the quantity of related materials. Every item has several data stored in a resource folder in different JSON files:

- Id: a unique increasing integer number to identify the item.
- Name: a string common name

- Description: a short description of the item. This is the description showed by the item panels and in the Encyclopedia
- Image: a picture of the item used used by the Encyclopedia.

Starting the game

Before the game begins, all the interactable components of the objects are disabled and all the description panels and the material emissions are enabled. When the game starts the *Guide* communicates with the *InternalRoomGameManager* to load 4 random items. The manager, after the random choice, loads the data about the four elements and sends them to the *Guide*. Moreover it disables all the panels and, if the emissivity facilitator is enabled, it also disables the emissivity of the four chosen items. The Virtual Guide then shows a dialogue panel containing the items that the user have to find inside the room.

During the game

The user can now interact with all the objects in the current room. When an object is selected, the panel is enabled again and, if the emissivity of the object was turned off, the object materials become emissive again. Every interaction with an object is followed by a feedback given by the Virtual Assistant. Moreover, if the item was correct, a "positive" feedback sound effect is played and the name of the found item on the Virtual Guide panel lights up, otherwise the user only hears a "negative" feedback sound effect. The game ends when all the correct items are found.

Game ending

After all the items have been retrieved, the Guide starts a little jumping animation. Meanwhile, the game manager saves the data on the retrieved objects for the Encyclopedia in a JSON file. This JSON simply contains two lists of integer called "unlocked" and "locked". The integers are the reference to the museum item ids. The save file is called "unlockedItems.fun" and is stored in the application *persistentDataPath* folder. The location of this folder changes according to the operating system that is running the application. In case of Android devices, the *persistentDataPath* generally points to `"/storage/emulated/0/Android/data/<packagename>/files"`.

Encyclopedia

The Encyclopedia is constituted by a dialogue panel showable from the Virtual Assistant help menu. The encyclopedia is composed by an object with two different dialogue box:

- EncyclopediaBox: contains four selection buttons and two navigation buttons. These buttons are managed through the *Enciclopedia* script. When the box is firstly opened the *SetMainPage* method is called. This method disables the first and fourth selection buttons and changes the text of the two remaining button to "Room 1" and "Room 2". When one of the two is selected, the script loads all the items' data calling the *DescriptionLoader* and the *ImageLoader*, two static

classes that works with Resources directory. Then all the buttons are re-enabled and their string content are changed to the items name in alphabetic order. Also the OnClickListeners of the buttons are always reset by the *Enciclopedia* script.

- **DescriptiveBox:** when an item is selected, the descriptive box is showed. The DescriptiveBox contains an image and some text containing the item description.

5.4.4 Conversational Agent and dialogues

Virtual Assistant

The conversational agent of the application is the Virtual Assistant. This one will follow the user during the whole experience and will give tasks by speaking and showing dialogue panels to the user. The Virtual Assistant is always aware of the actual state of the application, the current checkpoint and the goal to achieve. We can distinguish between three type of dialogue panels:

- **VADialogueBox:** this is the panel that shows all the tasks, helps and warnings. It only contains Facilitator Graphics (FG) and a TextMeshPro object. The string content of this is changed at runtime by the TriggerDialogue method in the *VirtualAssistant* script.
- **HelpDialogueBox:** this one is the dialogue box opened when the user interact with the Virtual Assistant. It contains the help button and the encyclopedia button.
- **Encyclopedia:** contains the two Encyclopedia panels [5.4.3]

Loading Dialogues

The Dialogue system is composed by several component:

- *Dialogue:* an abstract class containing a string text and the Checkpoint linked to the dialogue. It's extended by three subclasses:
 - *VADialogue:* contains dialogues relative to the Virtual Assistant. It can have four type: Default, How, What,Warning. The "Default" dialogues are those ones activated every time a new Checkpoint is reached. They describe the next goal to reach. "How" type dialogues are loaded when the user asks for help to the assistant. "What" dialogues basically contains the tasks to execute. The difference between the "Default" and "What" dialogues is the moment in which they are triggered: in fact, the "What" dialogues are triggered only when users ask to repeat the task they have to do. Basically "What" dialogues are short versions of the "Default" ones. Finally the "Warning" dialogues are triggered for general warning situations (e.g.: users cross the roads outside the zebra crossing)
 - *GuideDialogue:* are those dialogues related to the Virtual Guide.
 - *TicketManDialogue:* dialogues related to the ticket collector.

- *DialogueBuilder*: a class that follows the builder pattern for dialogues.
- *DialogueManager*: a static factory class for dialogues.

Every time that a component needs to load a dialogue it calls a Load method of the *DialogueManager*. There are three different Load methods according to the dialogue owner: LoadVaDialogue, LoadGuideDialogue, LoadTicketManDialogue. Every method receives as parameter the language and the checkpoint related to the needed dialogue. In the case of the *VADialogue* also a parameter to identify the type of needed dialogue is requested. The main function of the loader methods is to construct the right path to the resource text file with the dialogue content related to the passed Checkpoint. Then the loaders create a new *DialogueBuilder* to construct the new *Dialogue* container. Once the *Dialogue* has been constructed, it is returned to the first component and it is displayed in a dialogue box inside the virtual environment.

```
public Dialogue LoadVaDialogue(Language language, CheckPoint intent, VaTypeDialogue vaTypeDialogue)
{
    string fileName = "";
    string path = basePath + "VADialogues/";
    switch (intent)
    {
        case CheckPoint.StartingPoint:
            fileName = "starting_dialogue";
            break;
        case CheckPoint.EntranceFound:
            fileName = "entrance_found";
            break;
        case CheckPoint.FindEntrance:
            fileName = "find_entrance";
            break;
    }
}
```

Figure 5.7. Example of a loader method in the DialogueManager

Reading dialogues and talking to assistant

Once a dialogue has been loaded, a Read method is called from the static class *Speech*. There are two Read methods: ReadVaSentence and ReadGuideSentence. The difference between the two method is the voice that have to be used to read the dialogue. The voices are constituted by two *SpeechSynthesizer* attributes initialized during the launch of the application and stored in the *Speech* class. Every time that the language changes, the voices have to be reset according to the chosen language. All the algorithms to transform the string dialogue in a vocal output are managed by the Azure Speech SDK components.

Talking to the assistant is a more complicated operation. First of all the application has to listen for the microphone activation input. The *SttInputManager* registers the InputActionReference in the system and links the action to the "FromMic" method contained in the *Speech* class (N.B: the application uses the new Unity input system). Once the user activates the microphone, the application starts recording and it stops when no more world are sensed. Then the audio is analyzed by the Azure *SpeechRecognizer* component and a string containing the vocal transcription is returned. This string is passed to the *SttInputManager*. This manager have to

identify the type of request.

The application can recognize two types of request:

- **Help request:** help requests are triggered by sentences containing words like "help". When an help request is recognized a "How" dialogue is loaded according to the actual Checkpoint and actual Scene
- **Repeat task request:** these requests are triggered when the analyzed string contains sentences like "What should I do?" or "Where should I go?". In these case a "What" dialogue is loaded (always according to actual Scene and Checkpoint).

5.5 5A-Museums - AR architecture

The AR application is composed by a fat client running on the user device (Smartphone) while sending GET requests to a thin server that acts like a database. A stable connection is required for the application to work, because even if all the processes are run on the device, in order to support a great level of scalability, data and information about each museum are stored in the web server and accessed via URLs contained in QR codes.

5.5.1 AR Session

In order to use AR Foundation functionalities, Unity environment have been set up with the *ARSession* and *ARSessionOrigin* GameObjects. The first one is a Gameobject implementing an *ARSession* component used by the device to track features in its environment and an *ARInputManager* component required in order to enable world tracking. The second one implements several components, the most relevant are:

- *ARSessionOrigin*: responsible for transforming trackable features, such as planar surfaces and feature points, into their final position, orientation and scale in the Unity scene.
- *ARAnchorManager*: a type of trackable manager that allows to define certain points in space to be tracked with more effort by the device, called anchors. Additional work is performed by the system in order to update position and rotation of the anchor. Inside our project it is used to guarantee a better positioning of the VG in the environment.
- *ARTrackedImageManager*: a trackable manager that performs 2D image tracking. It is described more in details in the image recognition (5.5.5) section.

5.5.2 Modularity through QR codes

Like the VR application, also the AR one is modular, but its modularity is given in a different way. Each QR code is specific to a single place of the museum, that can be the outside part, the museum hall or one of the exhibit rooms. Whenever the user

scans a QR code, the state of the application changes according to the data recovered from it.

Scenes

Unlike the VR application, since AR aims at bringing virtual objects inside the real world, the AR application does not need to model a complex environment that can impact its performance or scalability. In this case, the Unity Scenes are only two:

- *MainMenu*: this is the first scene, needed to select the desired settings and learn how to use each app feature. it contains the splash screen of the application, the difficulty and language selection page and the tutorial on the app functionalities;
- *ARScene*: This is the scene where the whole experience will take place. the camera of the device will show on screen the environment while the virtual objects will be placed according to the user position and rotation;

In the AR part, there will be no need to add new scenes to expand the application, because the experience is personalized by the data obtained from the QR codes inside the rooms of each museum.

Checkpoints

The experience is divided in three main checkpoints: outside, museum hall and museum exhibit. The checkpoints are used to differentiate the contents of QR codes and allow the virtual assistant to load the appropriate dialogue.

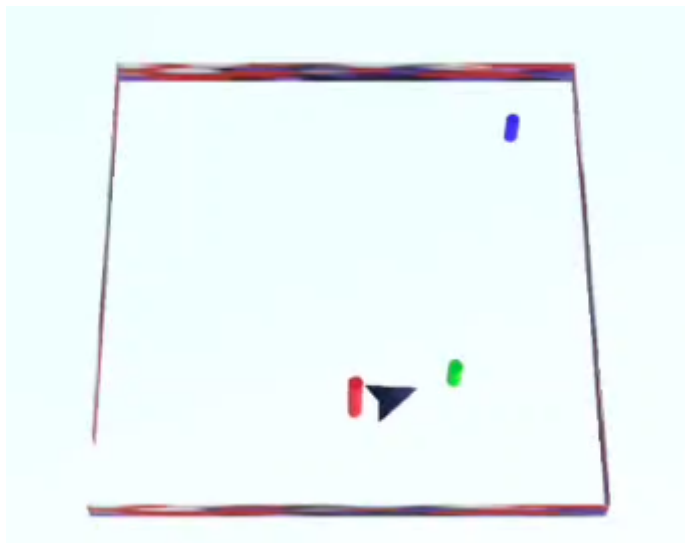
5.5.3 Pathfinding

Pathfinding is implemented exploiting Unity's NavMesh components features, that allows to define a walkable surface to a GameObject chosen as floor and compute the shortest path between two points, avoiding obstacles between them. In our case, the AR scene contains a GameObject called *DynamicMap* implementing the *MapGenerator* component that builds a new map every time a new room is discovered through QR codes scan. As shown by the figure 5.8a, these maps are composed by:

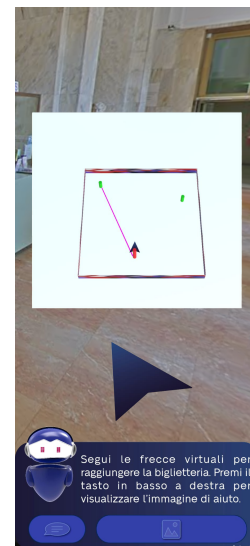
- A square GameObject used as floor that gets scaled according to the room's dimensions.
- A set of square GameObjects used as walls with a fixed height and a width scalable according to each wall's length that are placed orthogonally to the floor. Each wall GameObject contains a *NavMeshModifier* component that sets their area type as *Not Walkable*. They are needed by Unity's NavMesh in order to mark them as obstacles.
- A set of indicators shaped as green cylinders that represent the points of interest inside the room.

- A set of indicators shaped as blue cylinders that represent the artifacts of the museum and are used for the treasure hunt game. They are exclusive to the exhibit rooms of the museum and contain a *ArtifactData* component that contains all the data about the artifact they represent.

The *DynamicMap* GameObject contains also an indicator for the user (the red cylinder in the figure 5.8a) and a NavMesh GameObject that are shared between all the maps. The first one follows the movements of the user's device and gets translated and rotated according to the map that gets generated. In the figure, it is visible also an arrow that shows the direction the user is facing. The second one contains a *NavMeshSurface* component, needed to bake the walkable surface on floor GameObjects, that computes the walkable floor every time a map gets generated or loaded. In the figure 5.8b it is shown the minimum path between the user and a point of interest calculated by Unity's NavMesh component and also the virtual arrow that traces the path to follow. A final remark is that the map will be visible on the screen only in debug mode, so it will be used only to calculate minimum paths and distances, while final users will never be able to see it.



(a) Example of exhibit map



(b) Example of hall map with minimum path and virtual arrow

Figure 5.8. Example of virtual maps generated by the AR application

5.5.4 GPS Positioning

While inside the museum, the user's movements are tracked studying the position of their device in the physical space; outside this solution was not viable because the environment is way larger and it is not possible to put QR codes everywhere. In fact, QR codes are not only used to recover data about museum rooms, but also in order to reset the location of the user and eliminate any position or rotation offset due to the limits of the device tracking technology.

Since outside the museum using QR codes was not a viable option, we opted for a GPS based solution that, even if it is not extremely accurate, it offered us the possibility to refresh user's position in real time without depending on the QR code technology. Unity already implement components that allow to interface with the GPS features of the mobile devices in order to recover the geographical coordinates which describe the point where the user is. The main issue about geographical coordinates is that they are expressed in latitude and longitude, two angular values that describe position around the globe as the difference between its angles with the equator (the 0° parallel of latitude) and the prime meridian (the 0° meridian of longitude).

Since all distances on the application map are calculated in meters, the angular values of latitude and longitude had to be transformed in order to adapt with our system. The main idea behind the algorithm that transforms latitude and longitude to meters is to use a reference point in space (the origin of our axes) for which we know both the geographical coordinates and the coordinates in our system. Having a known point, it is then possible to use the *Haversine* formula (1) in order to find the distance in meters between the point and the geographical coordinates of the user's device. The distance is calculated along two axes, one that follows the direction of the longitude and the other follows the latitude.

Algorithm 1 Haversine algorithm

```

R ← 6378.137           //Radius of the earth in km
ΔLat ← lat2 * (π/180) - lat1 * (π/180)   //Difference of latitude
ΔLon ← lon2 * (π/180) - lon1 * (π/180)   //Difference of longitude
a ← Sin2(ΔLat/2) + Cos(lat1 * π/180) * Cos(lat2 * π/180) * Sin2(ΔLon/2)
c ← 2 * Atan2(a1/2, (1 - a)1/2)
distance ← R * c * 1000           //Distance in meters

```

Since these distances are calculated on the axes that describe latitude and longitude, the system of coordinates must be rotated around its origin until the axes correspond to those on which our map is built. In order to rotate the axes around their origin, we used the *Rotation of axes* algorithm (2).

Algorithm 2 Rotation of axes algorithm

```

angle ← angleDeg * π/180   //Angle between the systems
x' ← x * Cos(angle) + y * Sin(angle)
x' ← -x * Sin(angle) + y * Cos(angle)

```

5.5.5 Image Recognition

AR Foundation includes a component called *ARTrackedImageManager*, a trackable manager that performs 2D image tracking. Before an image can be detected, the manager must be given a set of reference images to look for in the real world. These images must be inserted in a reference image library that is then set as serialized library inside the *ARTrackedImageManager* component.

Since the images needed during the museum experience are not known a priori, the reference image library must be generated at runtime each time a QR code is scanned. In order to do this, the *ARTrackedImageManager* component implements a method called *CreateRuntimeLibrary* that allows to create a reference image library at runtime. It is possible then to use the newly created reference library as a *MutableRuntimeReferenceImageLibrary* and add images to it through its *ScheduleAddImageWithValidationJob* method.

5.5.6 Main Components

The AR app is managed by several script components. Here we report the main ones:

Application Flow

- **GameFlowManager:** This is a static class that keeps track of the application's state during the entire experience. Several data, such as the current checkpoint, language, game mode, name of the actual room, and name of the museum, are used to indicate the state of the application.
- **GameManager:** The GameManager component is in charge of initiating the flow of the application when the state of the program changes, initializing the GPS services on AR scene initialization, and recognizing the museum the user is in based on the his location.
- **HelpManager:** This is one of the application's most significant classes, as it totally controls the flow of the Guided experience and manages the Autonomous experience when the user requests the Virtual Assistant for help. It's also in charge of displaying the VA menu and adjusting the configuration and text of its buttons depending on the question the user has to answer. We decided to change configuration and text of the buttons, activating and deactivating them when needed, because it lowers the burden that is placed on the CPU when having to rapidly create and destroy GameObjects.

Conversational Agents

- **AssistantController:** The assistant controller is responsible for changing the line of dialogue of the Assistant and setting the sprite of the facilitator graphic in the Assistant menu. It also calls functions from the *TextAndSpeech* component in order to read Assistant's lines of dialogue.
- **GuideController:** This is the component that manages the positioning and removal of the Virtual Guide's GameObject within each room of the museum exhibit as well as updating its position when needed.
- **EncyclopediaMenuVA and EncyclopediaMenuVG:** These are the components in charge of managing the treasure hunt and the encyclopedia, the main contents inside the museum's exhibition rooms. The first one is responsible for handling the encyclopedia menu within the VA interaction box, while the second one completely controls the VG menu.

- **DialogueManager, DialogueBuilder and DialogueModel:** These are the components responsible for loading, formatting and managing the dialogues of the Virtual Assistant. In more detail, the first one generates a folder path depending on language, checkpoint and game mode from where it loads a txt file that contains the dialogue. Then the second one parses the txt file in order to generate a DialogueModel, a class whose instances contain the information about specific dialogues.
- **PCSLoader:** It contains a list of methods that can be called to load specific sprites for the facilitator graphic image in the VA menu.

QR Code Scanning and JSON Parsing

- **ScanStart:** ScanStart is the component that primarily manages the detection and scanning of QR codes as well as parsing the data they contain. Since QR codes for our experience are designed to contain an URL that returns a JSON file, this component is also able to send GET requests and decode JSON files.
- **ScanResult, MuseumArtifact and MuseumsList:** These are three serializable classes whose instances are created from decoding the JSON files. The first contains information about the room as well as the user's position and orientation (Figure 5.9a). The second contains information about a particular museum artifact (Figure 5.9b), and the third contains a list of all the museums on the web server, as well as their geographical location and JSON data about their outside (Figure 5.9c and 5.9d).

Pathfinding

- **MapGenerator, NavigationController and UpdateNavigation:** These are the components in charge of creating the maps of the rooms of the museum and managing the navigation inside those rooms. In more detail, when a JSON is decoded, MapGenerator receives data about the position of walls, user and points of interests inside the room and creates a stylized map. Previous rooms are not deleted but only set inactive, so, if the user scans the QR code of a room he already visited, the map does not have to be generated a second time, but only activated. The GameObject that represents each map also contains a NavMesh component that is then used by the NavigationController class to compute the shortest path to the points of interest when needed, while the UpdateNavigation class uses the path to update the direction and position of the arrow that guides the user.
- **UserManager:** It is the class that is responsible of keeping track of the position and rotation of the user. While the user is outside the museum, its position is computed using the GPS coordinates that, since they are described in latitude and longitude, they are also converted in meters and rotated in the same system of axes as the rest of the museum's measures. On the other hand, inside the museum, UserManager computes the user's position and rotation from the fluctuations in position and rotation of the user's device. In addition, this

```
[Serializable]
9 usages Davide Perugini 1 exposing API
public class ScanResult
{
    public string entranceUrl; ◀ Serializable
    public float entranceImageContentWidth; ◀ Serializable
    public string roomName; ◀ Serializable
    public List<float> userPosition; ◀ Serializable
    public float userOrientation; ◀ Serializable
    public List<float> museumEntrancePosition; ◀ Serializable
    public List<float> ticketOfficePosition; ◀ Serializable
    public List<float> exhibitEntrancePosition; ◀ Serializable
    public List<float> guidePosition; ◀ Serializable
    public List<float> nextRoomPosition; ◀ Serializable
    public List<MuseumArtifact> artifacts; ◀ Serializable
    public List<float> wallsX; ◀ Serializable
    public List<float> wallsY; ◀ Serializable

    //angle between y axis and north direction
    public float mapAngle; ◀ Serializable

    //coordinates of a chosen point used as reference
    public List<float> pointCoords; ◀ Serializable
    public List<float> pointLatLong; ◀ Serializable
}
```

(a) ScanResult class variables

```
[Serializable]
3 usages Davide Perugini
public class MuseumArtifact
{
    public string name; ◀ Serializable
    public List<float> position; ◀ Serializable
    public string description; ◀ Serializable
    public string imageURL; ◀ Serializable
    public string artifactPhotoUrl; ◀ Serializable
    public float artifactImageContentWidth; ◀ Serializable
    public string info; ◀ Serializable
}
```

(b) MuseumArtifact class variables

```
[Serializable]
4 usages Davide Perugini 1 exposing API
public class MuseumsList
{
    public List<Museum> museums; ◀ Serializable
}
```

(c) MuseumsList class variables

```
[Serializable]
5 usages Davide Perugini 2 exposing APIs
public class Museum
{
    public string name; ◀ Serializable
    public List<double> coords; ◀ Serializable
    public string dataURL; ◀ Serializable
}
```

(d) Museum class variables

Figure 5.9. Serializable classes instantiated through QR code scanning

component also keeps track of the distance between the user and several points of interest, in order to trigger specific dialogue or continue with the application flow.

- **UserIndicatorController:** It uses the position of the user tracked by the *UserManager* component in order to change position to the user indicator on the virtual map.

Image Recognition

- **ImageRecognitionManager:** Image recognition is used both to recognise the entrance of the museum during the "Outside" part of the experience and to recognise the artifacts during the "Treasure Hunt" minigame. ImageRecognitionManager is the component in charge of managing the application flow when an image from the library is recognized.

Sounds

- **SoundManager:** It exploits unity's AudioClips and AudioSources in order to implement simple methods that are called from other components to play sounds or effects.

Text-To-Speech (TTS) and Speech-To-Text (STT)

- **TextAndSpeech and SpeechToTextParser:** These are the components that manage the functionalities of TTS and STT. TextAndSpeech exploits the plugin "Speech and Text" (5.1.2) in order to present simple methods to read strings of text, transcribe what the user says or adjust the settings of TTS and STT. After the user's words have been transcribed, the result is then parsed by SpeechToTextParser and the application's flow is altered accordingly.

Chapter 6

Exploratory Study

After the design and implementation of the two applications that define the project of 5A-Museums, an exploratory study has been planned in order to test the potential of the implemented technologies directly on the targeted final users. Because the actual test is in progress, results are not reported. The purpose of this study is to evaluate and investigate the usability of immersive technologies for autistic subjects. The exploratory study consists of a phase of familiarization with the technology, a testing phase through the use of two applications, one in VR and one in AR, and an evaluation phase after the testing. The chosen scenario for the study is a museum visit.

6.1 Goal of the Study

The goal of the study is to ensure that:

- the training in VR is effectively translated in a better performance in a real-life scenario;
- the experience in AR is effective in reducing social anxiety in a public environment;
- the experience in AR is effective in increasing orientation skills;
- the Virtual Assistant does provide help and does not create discomfort in the users;
- the minigame activities are considered interesting and stimulating.

6.2 Variables

- **Effectiveness of the app:** how much is the system effective for training and guiding?
- **Involvement of users:** how much are the users immersed and enjoy the app and its functions? The application focuses enough the attention of the users?

- **User satisfaction:** how satisfied are the users with the application? Are they willing to use it again?
- **Ease of use:** do the users find the interface and interactions difficult to be used? How much does the system accommodate the users and facilitate navigation inside the app?
- **System reactivity:** how fast and effectively does the system react to the user's needs?
- **System reliability:** is the experience completely playable from the beginning to the end?

6.3 Hypotheses

- the pathfinding works correctly and helps the users to reach the main areas of the museum;
- the AR objects are shown correctly in relation to the real environment;
- the conversational agent results efficient in responding to the users needs;
- the users do not take too much time in accustoming with the interfaces and interactions;
- the users show little to no difficulty in understanding the tasks requested;
- the applications focus the attention of the user enough and reduce their social anxiety;
- the users acquire new knowledge about the contents and the artifacts of the museum through the applications.
- the familiarization phase is useful and helps the users to get quickly used to the proposed technologies and their controls.

6.4 Subvariables and their Measurement

- **Effectiveness of the app:**
 - Is the system providing correct information on the position?
 - Is the Virtual Assistant helping effectively?
 - Is the app responsive in giving the necessary help?
- **Involvement of users:**
 - Is the Treasure Hunt minigame enjoyed thoroughly?
 - Does the user feel immersed in the environment? Is the user's immersion affected by any occurring technical issues?

- **Satisfaction:**
 - Is the graphics of the menu pleasant?
 - Are the design of the Guide and Assistant pleasant?
 - Is the application understanding what the user is saying?
 - Are the sound effects pleasant?

- **Ease of use:**
 - Is the position of the buttons comfortable for use?
 - Is the vocal function easy to understand and use?

- **System reactivity:**
 - Is the user able to move at ease between the various sections of the app?

- **System reliability:**
 - Is it recognizing the items?
 - Is it positioning correctly the AR objects?
 - Does the application flow without interruptions?

All these parameters will be measured by monitoring and taking notes about the users' behaviour during all the experience. All the users feedbacks and mistakes will be annotated. For some tasks we are also going to time the users performances. Also all eventual bugs or crashes will be wrote down.

6.5 VR

The experience will take place within the therapy center for people with ASD and will consist of a virtual reality museum tour for each individual user. The chosen hardware systems for experimentation are two Oculus Quest headsets (version 1 and version 2). Virtual reality is a technology that is getting more and more used for therapeutic purposes thanks to its characteristics of immersivity, intended not only as spatial but also sensorial, defined as the capacity of virtual reality experiences to submerge the user's perceptual system in such a way that they can actually feel "inside" the digital environment. This plays well with the large and complex inner world of people suffering from ASD and allows technology to communicate with them in more effective ways.

6.6 AR

The experience will take place inside the *Museo Civico di Storia Naturale di Milano* and will be a tour guided by an augmented reality smartphone application from the outside, in the private courtyard of the museum inside the *Giardini Indro Montanelli* park, to the museum exhibit rooms.

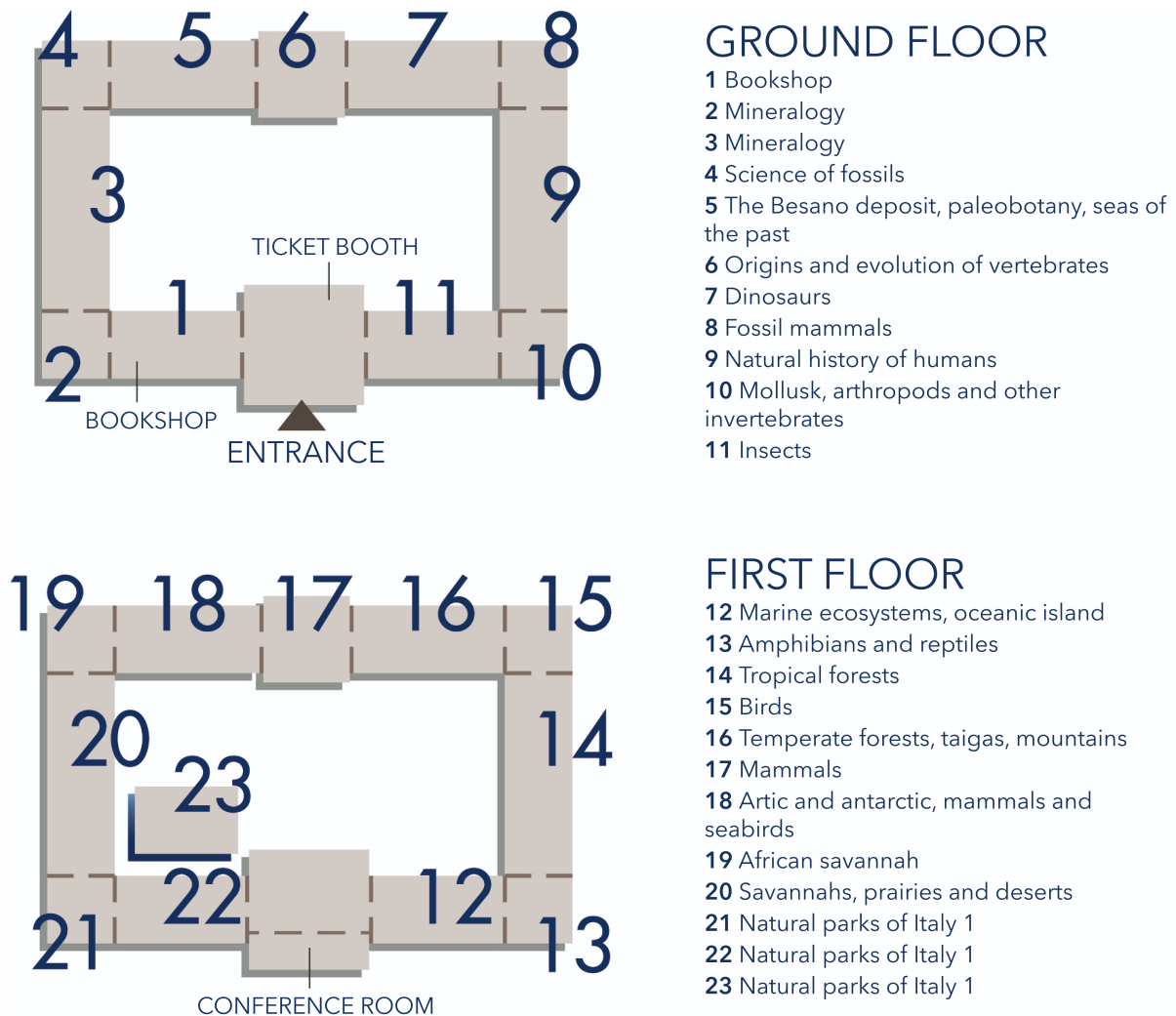


Figure 6.1. Map of the Museo Civico di Storia Naturale di Milano

6.7 Research Design

This study requires a collaboration between the two parts, the researchers and test subjects. The study wants to test if the overall system works and proves the theory behind it, but also is paying attention to how the single features are received. Participants needs and feedbacks will be listened carefully as may contribute to future developments of the application. Finally, this is an adaptable study that can work in the context of different museums and may be repeated for future studies.

The study is a *Within subjects* study: the same person tests all the conditions (i.e., all the user interfaces). By the intrinsic design of the 5A application, the test should maximize the learning and transfer across the two conditions (VR and AR). To avoid order effects, the test subjects will not complete the experience in AR all together, but will be divided in small group manageable by one person, to test each participant feedback and use of the application without being influenced by the group.

6.8 Participants

The experimental group is constituted of 5 subjects, and it is aimed exclusively at high-functioning autistic subjects, that means that are all affected by autism with mild symptoms. The selected subjects are from 11 to 20 years old.

6.9 Execution

The whole testing phase will take place in a single day. The VR application will be tested inside the therapy center structure, while the AR application testing will take place directly at the *Museo Civico di Storia Naturale di Milano*.

6.9.1 VR session

First of all, all the subjects will be introduced to the whole experience. Then the VR session will begin. The VR testing session will be composed of the following steps:

- Introduction to the Oculus device: users will be briefly instructed about the Oculus correct use and a short explanation on input device will be given (5 min).
- Application setup: application will be launched and all the difficulty settings will be tweaked by us.
- Testing the game: the experience will begin from the main menu and will comprehend all the museum rooms. The user will be said to start the experience by simply selecting the Start button. All the testing process will be supervised by a team member and a therapist.
- Filling the questionnaire: After the user completes the experience, they will have to fill a google form with some questions on usability and likeability of the application.

All the steps are repeated for every user and will take about 25/30 minutes.

6.9.2 AR session

In order to grant to the participants the best and most accurate experience, the AR session will be preceded by a preliminary visit to the museum composed of the following steps:

- Entrance to the museum and introduction to the supervisors about the purposes of our visit, that will in any case already have been informed by email.
- Measurement of the hall of the museum together with the bookshop, represented by the room 1 and the one with the entrance in the image 6.1.
- Measurement of the rooms 2, 3 and 4 that will be considered as a single room during the experience.

- Measurement of the rooms 5, 6 and 7 that will be the main rooms in which the experience will take place, and will be considered as three separate rooms.
- Writing and uploading of the JSON files to the web server.
- Posting of the QR codes in the various rooms of the museum.
- Preliminary testing of the whole experience in order to make sure that each feature works as it should. This step includes also the fixing of any application bugs that might come up during the testing.

The AR session will be composed of the following steps:

- **Introduction:**
 - Meeting with participants and therapists nearby the museum entrance. Explanation and short tutorial of the application.
 - Subdivision of the participants in 3 groups of 3/4 people. Each group is then assigned to a member of the development team and at least a therapist; group 2 will join 10 minutes after group 1, and group 3 will join 10 minutes after group 2.
- **Experience repeated for each group:**
 - Each group member is given a smartphone device with already installed the AR application of 5A-Museums and GPS location and mobile data set active.
 - Participants are taken to a chosen position of the gardens outside the museum, in order to test the efficiency of the GPS pathfinding.
 - Participant are instructed to follow the instructions of the application and, if they find themselves in difficulty, ask the VA for help. In case of inability to continue or problems related to the application itself, then the developers or the therapists will be available to help them, but they will try as much as possible to leave job to the VA.
 - An example of QR code is shown to the participants and they are instructed on how to scan them and on what is their purpose.
 - Participants launch the application and are instructed to select Italian language and *guided* mode.
 - Participants are instructed to follow the tutorial and any questions from them are answered.
 - The group is instructed to use manual functions or speech to text according to their preferences and needs.
 - At the time all participants end the tutorial phase, the experience starts, and so they are given more freedom in order to follow the application's instructions. Even if they are given more freedom, they are never left by themselves, in order to gather data during the experience and to support them in case they need urgent support.

- Once the participants reach the hall of the museum and the VA asks them if they already bought the ticket, they are instructed to select the negative answer, in order to be guided to the ticket office to recover their ticket.
- Once the participants reach the museum rooms where the last part of the experience will take place, they are instructed to interact with the VG and to try the treasure hunt game.
- After the participants complete the treasure hunt at least one time, they are instructed to try the encyclopedia feature to consult the information about the museum artifacts discovered through the interactive game.
- When the group ends the experience, each participant is given a questionnaire on google form with questions about usability and likeability of the application.

The duration of the testing for each person is estimated about 1 hour and 30 minutes. Beyond the 7th room, the participant will be asked to give back the smartphones as the application is implemented only up until that point.

6.10 Questionnaire

At the end of both the AR and the VR sessions, the users will undergo a questionnaire phase in which they will be asked to fill a Google form in order to collect their feedback about the experiences.

Both the surveys follow the same structure and are composed by the same number of questions that can be answered with a number between 1 and 5, where the first means strong disagreement and the latter means strong agreement. The first 10 question are extracted from the Affinity for Technology Interaction Scale, or ATI scale [46], and are designed to assess a person's tendency to actively engage in intensive technology interaction. In other words, how much a person is capable and fond of dealing with technology.

Next, 10 question on the usability of the system are presented. These are from the System Usability Scale (SUS), a reliable and widely used tool to asses the usability of an application. The rest of the questions are a selection from the a special questionnaire (VRUse [1]) that has been designed to measure the usability of a VR system according to the attitude and perception of its users. A few questions were changed and adapted to work also for AR infrastructures.

The scales of evaluation of these three systems will be used to assess the results and interpreting the answers.

The full questionnaires can be found in the appendix of the document (A).

Chapter 7

Conclusions

7.1 Contribution

The main contributions of this thesis work are mainly two. From a technological point of view, we developed a new extended reality project, defining a new connection between the Virtual and Augmented Reality technologies. These two, in fact, are generally treated separately and used for different scopes and purposes. 5A-Museums, instead, comprehend a design and implementation process that takes into account an experience that fully integrates both the potentialities of VR and AR. From a methodological point of view, the study defines a new possible approach in the treatment of ASD with the support of modern technological devices that are becoming more and more popular for therapy treatments.

7.2 Limitations

First of all it's important to remark that, due to Covid-19 related problems, the testing phase is still in progress, so, for now, the results are not available.

Concerning limitations on technological aspect, the main problems were related to the hardware processing skills and performance limits. In particular:

- For the VR application, running on Oculus Quest, we had to limit the scene size and the quantity of objects present in every scene. Moreover all the used graphical assets had to be simplified, in terms of used polygons and vertices to compose the assets. Another problem related to the Oculus hardware limit was the poor computing ability to calculate light effects and reflections that forced us to use only pre-computed light effects.
- The AR application was tested only on smartphones with high-end processors and Graphics processing units (GPUs). Moreover, the used plugins ARCore and ARKit have limits on the supported operating systems, working only with latest versions of Android and iOS (Android 7.0 and later releases, iOS 11 and later releases).

Concerning the used tools, other limitations that we had to take in consideration were:

- The Azure free plan limit: even if Azure Speech is a good solution, it is not an open source product. It has a monthly limit on the number of characters that can be read by the vocal synthesizer, that is extremely permissive for a testing phase and a small number of users, but it can be a problem for an extension of the target user. Since all the 5A project is no-profit, this could be a possible problem in future.
- Unity frequent changes: Unity engine is constantly updated and expanded. Importing the project in a new Unity version is generally a safe procedure. However, it can happen that, if the two Unity versions are really different and distant in time, the re-imported project has to be reworked to fit the new update.

7.3 Future works

5A-Museums has been designed as a modular experience to permit an high scalability of the system. The solution that we implemented is a minimum part of all the modules that can be added. Here report a list of possible future works:

- Adding more rooms to VR museum: all the VR rooms are contained by different scenes, so adding rooms is a relative easy procedure.
- New games to engage users inside exhibit rooms both in AR and VR;
- More sophisticated Conversational Agent (CA): the current CA is a basic prototype and can recognize only few words and sentences. A possible extension is to improve the CA quality.
- Finding or implementing an open source solution for STT/TTS for Oculus devices.
- Adding indicative silhouettes to image recognition: image recognition in some cases works only if the object is at a certain angle to the camera. Adding a semi-transparent silhouette of the object can help the user to stay at the correct angle to recognize the item faster.
- Developing an online map editor: the AR application needs to have JSON files containing data on the museum structure. Creating this files is not always fast, and it requires a site inspection of the target place. Having an online map editor can speed up this process.
- Adding new checkpoints and making some of them optional: inside the already existent rooms (in VR) some tasks, like a queue task, can be added and other ones, like those regarding the ticket, can be made optional.
- Porting on other devices (Oculus devices or smartphones for VR, smartglasses for AR)
- Multiplayer experience for VR: multiplayer could improve social interactions between users.

- Increase the expressiveness of the Assistant and the Guide: improving the range of expressions of the two characters could train the user to recognize expressions.

Appendix A

First Appendix

Here we report a list of all the questions inserted in the post testing modules. In all the questions users can answer with a value from 1 (completely disagree) to 5 (completely agree). Both the modules for VR and AR are divided in several sections.

A.1 ATI Scale Questions

- 1) I like to use technological systems more thoroughly.
- 2) I like to test the functions of new technological systems.
- 3) I deal with technological systems mainly because I am obliged to do so.
- 4) When I have a new technological system in front of me, I immediately try it with great interest.
- 5) I like to spend my time discovering and learning about new technological systems.
- 6) For me it is enough that a technological system works, I don't care how or why.
- 7) I try to understand how exactly a technological system works.
- 8) For me it is enough to know the basic functions of a technological system.
- 9) For me it is enough to know the basic functions of a technological system.

A.2 Usability

- 10) I think that I would like to use this system frequently.
- 11) I found the system unnecessarily complex.
- 12) I thought the system was easy to use.
- 13) I think that I would need the support of a technical person to be able to use this system.
- 14) I found the various functions in this system were well integrated.
- 15) I thought there was too much inconsistency in this system.
- 16) I would imagine that most people would learn to use this system very quickly.
- 17) I found the system very cumbersome to use.
- 18) I felt very confident using the system.
- 19) I needed to learn a lot of things before I could get going with this system.

A.3 Interface

- 20) It was easy to access all the features on the home menu.
- 21) It was easy to access all the features during the user experience
- 22) It was difficult to remember all the functions available.
- 23) I understood the meaning of the various buttons on the control interface
- 24) I didn't need to use all the features provided.
- 25) Overall I would rate the system in terms of features:
(1-Very unsatisfactory; 5-Very satisfactory).

A.4 Inputs commands

- 26) I kept making mistakes while interacting with the system.
- 27 a) It was easy to select the objects in the virtual environment (VR).
- 27 b) It was easy to click the icons on the screen (AR).
- 28) I found it easy to move or reposition myself in the virtual environment. (only VR)
- 29) Overall, I would evaluate the user input as:
(1-Very unsatisfactory; 5-Very satisfactory).

A.5 Display and Audio

- 30) The amount of glitches and delays in the image affected my performance.
- 31) The image quality affected my performance.
- 32) The information displayed was too complicated.
- 33) Audio feedback helped my performance.

A.6 Engagement

- 34) I would feel comfortable using this system for long periods.

A.7 reliability

- 35) Overall, I would evaluate the reliability of the system as:
(1-Very unsatisfactory; 5-Very satisfactory).

A.8 Virtual Assistant

- 36) The Virtual Assistant is easily visible and accessible on the screen and in the scenario.
- 37) The virtual assistant was useful for me to find the right path.
- 38) The virtual assistant was clear as he explained the instructions for the next activity
- 39) The virtual assistant was annoying
- 40) The virtual assistant was too intrusive
- 41) I would have liked more help from the Virtual Assistant

A.9 Virtual Guide

- 42) The Virtual Guide was easily visible and accessible on the screen and in the scenario
- 43) The Virtual Guide was clear while giving instructions for the treasure hunt game

A.10 Facilitator Graphics

- 44) The images of the Virtual Assistant dialogues were clearly understandable.
- 45) The images of the Virtual Assistant dialogues were helpful
- 46) I would have preferred more help pictures

A.11 Text-to-Speech

- 47) The Virtual Assistant's voice was clear
- 48) The voice of the Virtual Guide (Ciro) was clear
- 49) The voice prompts were more helpful than the written dialogues
- 50) The voice prompts made me feel more confident using the game

A.12 Speech-to-text

- 51) I used the voice input feature
- 52) Speech recognition was responsive
- 53) Speech recognition has always been accurate
- 54) I would have felt lost without being able to ask for help through my voice

A.13 Facilitators

- 55) The directional arrow was clearly visible on the screen.
- 56) The directional arrow was useful for me to find the right path

A.14 Exclusive questions on VR questionnaire

A.14.1 Facilitators

- 57) The areas illuminated in orange were clearly visible and accessible.
- 58) The orange-lit areas were helpful in finding the right path.
- 59) The road signs were clearly visible
- 60) Road signs were helpful in finding the right path

A.14.2 Distractors

- 61) Ambient sounds within the game were too distracting
- 62) The crowds of people and the cars were too annoying

A.14.3 Treasure Hunt

- 63) I had a hard time finding the animals / artifacts required by the treasure hunt
- 64) I had trouble selecting the animals / artifacts required by the treasure hunt

A.15 Exclusive questions on AR questionnaire

A.15.1 Image recognition

- 65) The exhibits were easily recognizable by the system
- 66) The acknowledgment feedback was clear

A.15.2 QR codes

- 67) QR codes were easy to find. 68) The QR codes were blocked or unusable.
- 69) There were enough QR codes in each area of the museum.

Acronyms

ASD	Autism Spectrum Disorder
AR	Augmented Reality
CA	Conversational Agent
DLS	Daily Living Skills
DSM-5	Diagnostic and Statistical Manual of Mental Disorders
FG	Facilitator Graphics
GPUs	Graphics processing units
HUD	Head-up display
NDD	Neuro-Developmental Disorders
NPCs	Non-Player Characters
PCS	Picture Communication Symbol
UI	User Interface
VA	Virtual Assistant
VG	Virtual Guide
VR	Virtual Reality
TTS	Text-To-Speech
STT	Speech-To-Text

Bibliography

- [1] Roy S Kalawsky. “VRUSE—a computerised diagnostic tool: for usability evaluation of virtual/synthetic environment systems”. In: *Applied ergonomics* 30.1 (1999), pp. 11–25.
- [2] Roxana Moreno and Richard E Mayer. “Cognitive principles of multimedia learning: The role of modality and contiguity.” In: *Journal of educational psychology* 91.2 (1999), p. 358.
- [3] Sarah Parsons et al. “Development of social skills amongst adults with Asperger’s Syndrome using virtual environments: the ‘AS Interactive’ project”. In: *Proc. The 3rd International Conference on Disability, Virtual Reality and Associated Technologies, ICDVRAT*. 2000, pp. 23–25.
- [4] Tina R Goldsmith and Linda A LeBlanc. “Use of technology in interventions for children with autism.” In: *Journal of Early and Intensive Behavior Intervention* 1.2 (2004), p. 166.
- [5] Justin HG Williams et al. “Visual–auditory integration during speech imitation in autism”. In: *Research in developmental disabilities* 25.6 (2004), pp. 559–575.
- [6] Janet Preis. “The effect of picture communication symbols on the verbal comprehension of commands by young children with autism”. In: *Focus on Autism and Other Developmental Disabilities* 21.4 (2006), pp. 194–208.
- [7] Ingrid Maria Hopkins et al. “Avatar assistant: improving social skills in students with an ASD through a computer-based intervention”. In: *Journal of autism and developmental disorders* 41.11 (2011), pp. 1543–1555.
- [8] Lizbeth Escobedo et al. “MOSOCO: a mobile assistive tool to support children with autism practicing social skills in real-life situations”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2012, pp. 2589–2598.
- [9] American Psychiatric Association, American Psychiatric Association, et al. “Diagnostic and statistical manual of mental disorders: DSM-5”. In: *Arlington, VA* (2013).
- [10] Frank Biocca and Mark R Levy. *Communication in the age of virtual reality*. Routledge, 2013.
- [11] Hsin-Kai Wu et al. “Current status, opportunities and challenges of augmented reality in education”. In: *Computers & Education* 62 (2013), pp. 41–49. ISSN: 0360-1315. DOI: <https://doi.org/10.1016/j.compedu.2012.10.024>. URL: <https://www.sciencedirect.com/science/article/pii/S0360131512002527>.

-
- [12] Sara Bernardini, Kaśka Porayska-Pomsta, and Tim J Smith. “ECHOES: An intelligent serious game for fostering social communication in children with autism”. In: *Information Sciences* 264 (2014), pp. 41–60.
- [13] Lizbeth Escobedo et al. “Using augmented reality to help children with autism stay focused”. In: *IEEE Pervasive Computing* 13.1 (2014), pp. 38–46.
- [14] Vanessa Hus Bal et al. “Daily living skills in individuals with autism spectrum disorder from 2 to 21 years of age”. In: *Autism* 19.7 (2015), pp. 774–784.
- [15] Anita Thapar and Michael Rutter. “Neurodevelopmental disorders”. In: *Rutter’s child and adolescent psychiatry* (2015), pp. 31–40.
- [16] Asmaa Bakroon and Vasudevan Lakshminarayanan. “Visual function in autism spectrum disorders: a critical review”. In: *Clinical and Experimental Optometry* 99.4 (2016), pp. 297–308.
- [17] David F Cihak et al. “Evaluating augmented reality to complete a chain task for elementary students with autism”. In: *Journal of Special Education Technology* 31.2 (2016), pp. 99–108.
- [18] Haylie L Miller and Nicoleta L Bugnariu. “Level of immersion in virtual environments impacts the ability to assess and teach social skills in autism spectrum disorder”. In: *Cyberpsychology, Behavior, and Social Networking* 19.4 (2016), pp. 246–256.
- [19] Ali Adjorlu et al. “Daily living skills training in virtual reality to help children with autism spectrum disorder in a real shopping scenario”. In: *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*. IEEE. 2017, pp. 294–302.
- [20] Edward P Alexander, Mary Alexander, and Juilee Decker. *Museums in motion: An introduction to the history and functions of museums*. Rowman & Littlefield, 2017.
- [21] US Department of Health, Human Services, et al. “Autism spectrum disorder: communication problems in children”. In: *National Institute on Deafness and Other Communication Disorders [NIDCD]* (2017).
- [22] Christian Moro et al. “The effectiveness of virtual and augmented reality in health sciences and medical anatomy”. In: *Anatomical sciences education* 10.6 (2017), pp. 549–559.
- [23] Patrick Schueffel. *The Concise Fintech Compendium*. Sept. 2017. ISBN: 9782940384440.
- [24] Salatiel Dantas Silva et al. “Knowledgemon hunter: a serious game with geolocation to support learning of children with autism and learning difficulties”. In: *2017 19th symposium on virtual and augmented reality (SVR)*. IEEE. 2017, pp. 293–296.
- [25] Jenny A Vlachou and Athanasios S Drigas. “Mobile Technology for Students & Adults with Autistic Spectrum Disorders (ASD).” In: *International Journal of Interactive Mobile Technologies* 11.1 (2017).
- [26] Elaine C Zachi et al. “Color vision losses in autism spectrum disorders”. In: *Frontiers in psychology* 8 (2017), p. 1127.

-
- [27] Shahira Sharaf Eldin and Hadeel Fathy El Naggar. “Potential of Virtual reality in special needs higher education”. In: *Eurasia International Conference*. 2018.
- [28] Liliana Laranjo et al. “Conversational agents in healthcare: a systematic review”. In: *Journal of the American Medical Informatics Association* 25.9 (2018), pp. 1248–1258.
- [29] A Rega et al. “New technologies and autism: can augmented reality (ar) increase the motivation in children with autism”. In: *INTED2018 Proceedings* (2018), pp. 4904–4910.
- [30] Natalia Stewart Rosenfield et al. “A virtual reality system for practicing conversation skills for children with autism”. In: *Multimodal Technologies and Interaction* 3.2 (2019), p. 28.
- [31] Katherine Valencia et al. “The impact of technology on people with autism spectrum disorder: A systematic literature review”. In: *Sensors* 19.20 (2019), p. 4485.
- [32] Mohammad Rafayet Ali et al. “A virtual conversational agent for teens with autism spectrum disorder: Experimental results and design lessons”. In: *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*. 2020, pp. 1–8.
- [33] Carmen Berenguer et al. “Exploring the impact of augmented reality in children and adolescents with autism spectrum disorder: A systematic review”. In: *International Journal of Environmental Research and Public Health* 17.17 (2020), p. 6143.
- [34] Ian T Miller et al. “Virtual Reality Air Travel Training Using Apple iPhone X and Google Cardboard: A Feasibility Report with Autistic Adolescents and Adults”. In: *Autism in Adulthood* 2.4 (2020), pp. 325–333.
- [35] Alessandra Pasini. 2020. URL: <http://hdl.handle.net/10589/177776>.
- [36] Elio Salvini. 2020. URL: <http://hdl.handle.net/10589/177805>.
- [37] Gray Atherton and Liam Cross. “The Use of Analog and Digital Games for Autism Interventions”. In: *Frontiers in Psychology* (2021), p. 3049.
- [38] Salvatore Vita et al. “ARtis: how AR supports the guided experience in museums for people with autism”. In: *teleXbe*. 2021.
- [39] In: (). URL: <http://angsa.it/autismo/numeri/>.
- [40] URL: <https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@1.0/manual/index.html>.
- [41] URL: <https://docs.unity3d.com/Manual/com.unity.xr.management.html>.
- [42] URL: <https://docs.unity3d.com/Packages/com.unity.xr.arfoundation@4.1/manual/index.html>.
- [43] URL: <https://docs.unity3d.com/Packages/com.unity.xr.arcore@4.1/manual/index.html>.
- [44] URL: <https://docs.unity3d.com/Packages/com.unity.xr.arkit@4.1/manual/index.html>.

- [45] URL: <https://azure.microsoft.com/en-us/services/cognitive-services/speech-to-text/#overview>.
- [46] *ATI Scale*. URL: <https://ati-scale.org/>.
- [47] Nigel Newbutt. "WE THE CURIOUS VIRTUAL REALITY TOUR". In: ().
- [48] Hannah Ritchie. "Neurodevelopmental disorders". In: (). URL: <https://ourworldindata.org/neurodevelopmental-disorders#prevalence-of-autistic-spectrum-disorders>.
- [49] Sol Rogers. "How Virtual Reality Can Help Those With Autism". In: (). URL: <https://www.forbes.com/sites/solrogers/2019/04/03/how-virtual-reality-can-help-those-with-autism/?sh=3610cc01198e>.