UNIVERSITY OF ZAGREB
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

MASTER THESIS No. 529

A COLLABORATIVE APPLICATION FOR BIG DATA VISUALIZATION USING MIXED REALITY

Lea Brzica

Zagreb, June 2024

UNIVERSITY OF ZAGREB
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

MASTER THESIS No. 529

A COLLABORATIVE APPLICATION FOR BIG DATA VISUALIZATION USING MIXED REALITY

Lea Brzica

Zagreb, June 2024

UNIVERSITY OF ZAGREB
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

Zagreb, 04 March 2024

MASTER THESIS ASSIGNMENT No. 529

Student:	**Lea Brzica (0036517012)**
Study:	Computing
Profile:	Computer Science
Mentor:	prof. Lea Skorin-Kapov

Title: **A Collaborative Application for Big Data Visualization Using Mixed Reality**

Description:

Big data visualization techniques using mixed reality (MR) technologies leverage the immersive capabilities of MR to provide dynamic and interactive representations of vast datasets. These techniques blend virtual elements with the real world, offering users an intuitive way to explore and understand complex data. MR technologies enable users to manipulate and analyze data in real-time, facilitating insights that might be difficult to discern from traditional 2D visualizations. An example approach is to spatially map data points within a physical environment, allowing users to move around and interact with data in three dimensions. Your taks is to investigate and describe different big data visualization techniques using MR. Furthermore, your task is to implement a collaborative networked MR-based application designed to visualize one or more selected datasets, allowing multiple users to simultaneously interact with and analyze the data in an intuitive way. Finally, your task is to evaluate the developed application from a Quality of Experience perspective.

Submission date: 28 June 2024

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Zagreb, 4. ožujka 2024.

DIPLOMSKI ZADATAK br. 529

Pristupnica:	**Lea Brzica (0036517012)**
Studij:	Računarstvo
Profil:	Računarska znanost
Mentorica:	prof. dr. sc. Lea Skorin-Kapov

Zadatak: **Kolaborativna aplikacija za vizualizaciju velikih količina podataka pomoću tehnologije miješane stvarnosti**

Opis zadatka:

Tehnike vizualizacije velikih količina podataka pomoću tehnologija miješane stvarnosti (engl. Mixed Reality, skr. MR) koriste imerzivne sposobnosti MR-a kako bi pružile dinamične i interaktivne prikaze velikih skupova podataka. Ove tehnike kombiniraju virtualne elemente sa stvarnim svijetom, nudeći korisnicima intuitivan način istraživanja i razumijevanja kompleksnih podataka. MR tehnologije omogućuju korisnicima manipuliranje i analizu podataka u stvarnom vremenu, olakšavajući zaključke koje bi bilo teško razaznati iz tradicionalnih 2D vizualizacija. Primjer pristup je prostorno mapiranje podataka unutar fizičkog okruženja, omogućavajući korisnicima kretanje i interakciju s podacima u tri dimenzije. Vaš zadatak je istražiti i opisati različite tehnike vizualizacije velikih količina podataka pomoću tehnologije MR. Nadalje, Vaš zadatak je implementirati kolaborativnu umreženu MR aplikaciju za vizualizaciju jednog ili više odabranih skupova podataka, omogućujući višestrukim korisnicima istovremeno interakciju s podacima i analizu na intuitivan način. Konačno, vaš zadatak je evaluirati razvijenu aplikaciju iz perspektive iskustvene kvalitete krajnjih korisnika.

Rok za predaju rada: 28. lipnja 2024.

Thank you to my mentor prof. Lea Skorin-Kapov for all her help, guidance, and time. Also, thank you to mag.ing. Sara Vlahović and mag.ing. Mirta Moslavac for their support and assistance whenever it was needed. Finally, thank you to everyone else who supported me through my education and celebrated my successes.

CONTENTS

INTRODUCTION

In the realm of data visualization, the integration of immersive technologies marks a significant advancement, offering dynamic and interactive representations of extensive datasets. Technologies such as Mixed Reality (MR) blend virtual elements seamlessly into the real world [1], providing users with an intuitive platform to explore and comprehend complex data sets. Unlike traditional 2D visualizations, MR empowers users to manipulate and analyze data in real time, facilitating otherwise challenging insights. For instance, creating a geospatial visualization of data points allows users to navigate and interact with data in three dimensions, enhancing their understanding and engagement [2]. This thesis explores various big data visualization techniques using MR, focusing on the implementation of a collaborative networked MR-based application called *VisualizeAir* designed to visualize and analyze selected datasets. This application enables multiple users to interact simultaneously, fostering intuitive and collaborative data exploration. Furthermore, the thesis evaluates the developed application from a Quality of Experience (QoE) perspective, examining usability, immersion, and the effectiveness of collaboration in data visualization tasks.

The first chapter of the thesis gives a theoretical foundation covering topics such as big data visualization techniques, immersive technologies, immersive analytics, and the Air Quality Index (AQI). Chapter 2 covers related work reviewed in the fields of immersive data visualization, collaborative approaches in immersive environments, and specific studies on air quality data visualization. Furthermore, the third and fourth chapters detail *VisualizeAir*'s design, functionalities, and development process, including the implementation of technologies like the globe model and integration with OpenAQ data. Chapter 5 describes the hardware and software setup for the user study, the procedure, participant demographics, and the questionnaire. The sixth chapter discusses results from the user study highlighting findings on MR interaction, data visualization experiences, collaboration features, and identified limitations. Finally, the last chapter concludes with reflections on future enhancements and the potential of *VisualizeAir* to advance data visualization practices.

1. Theoretical Background

1.1. Big Data and Big Data Visualization

Big Data refers to massive data sets characterized by large, varied, and complex structures, which present challenges in storing, analyzing, and visualizing for further processes or results [3]. The term Big Data emerged from the necessity for large companies, such as Facebook, Yahoo, and Google, to analyze vast amounts of data [4]. Initially, Big Data was characterized by three V's: *volume*, *velocity*, and *variety* [5]. *Volume* refers to the large amount of data, while *velocity* indicates the speed at which data is generated. *Variety* concerns the diversity of data formats, sources, and structures. Data can originate from numerous sources and exist in various formats, such as numerical data, text documents, images, videos, sensor information, and social media. Despite the classification of data into structured, semi-structured, and unstructured categories, approximately 80% of the world's data is unstructured [6]. Subsequently, the three V's have been expanded to include a fourth V: *variability* or *value*, depending on the source [7]. There are also discussions about further expanding these characteristics, with *visualization* being one potential addition [8].

Visualization is the representation of data in a pictorial or graphical format, which facilitates the comprehension of complex concepts and the identification of new patterns [8]. Humans have an enhanced ability to process large volumes of data more efficiently when it is presented visually, as well as to identify patterns in visual representations [9]. Consequently, data visualization acts as a universal medium for conveying information, allowing the target audience to swiftly grasp the core message [10].

Using pictures, graphs, charts, and maps to understand data and information has been practiced for centuries [11]. Traditional visualization techniques include graphs such as line charts, bar charts, pie charts, area charts, and heat maps. However, due to the vast volume and high magnitude of big data, effective visualization is becoming increasingly challenging. Many visualization tools struggle with issues such as low performance in scalability and response time, and datasets are often too large to fit into memory [9] [12].

Modern data visualization techniques, such as word clouds, symbol maps, and connectivity charts, have been developed to handle and represent big data, particularly semi-structured and unstructured data [11].

1.2. Immersive Technologies

This section provides an overview of key theoretical concepts concerning immersive technologies, starting with a definition of *immersion* and its significance within this context. It then explores Virtual Reality (VR), Augmented Reality (AR), Mixed Reality (MR), and the broader concept of Extended Reality (XR). Due to potential inconsistencies, these concepts are defined and explained here to establish a shared understanding for the rest of the thesis.

Immersion is defined as a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences [13]. In digital media, immersion refers to the sensory and perceptual experience of being situated in a non-physical, mediated, or simulated virtual environment [14]. This state involves overriding one's sensory experiences of the real world, leading to a diminished awareness of physical surroundings and an enhanced engagement with virtual stimuli such as sight, sound, and touch.

Virtual Reality is an environment that is fully generated by digital means. To qualify as virtual reality, the virtual environment should differ from the local environment [1]. VR typically utilizes head-mounted displays (HMDs) to block out the physical world and replace it with simulated environments [15]. Its primary goal is to create a profound sense of presence and immersion, enabling users to interact with virtual objects and environments as if they were physically present.

Augmented Reality is defined as an environment containing both real and virtual sensory components. The augmented reality continuum runs from virtual content that is overlaid on a real environment (Assisted Reality) to virtual content that is seamlessly integrated and interacts with a real environment (mixed reality) [1]. Unlike VR, AR does not replace the physical world but supplements it through devices like smartphones, AR glasses, or other wearables [15].

Mixed Reality is an environment containing both real and virtual components that are seamlessly integrated and interact with each other in a natural way (one end of the augmented reality continuum) [1]. MR environments are spatially aware, enabling virtual objects to interact with physical surfaces and objects, thus creating immersive experiences where digital content integrates naturally with the user's physical surroundings.

Extended Reality is an environment containing real or virtual components or a combination thereof, where the variable X serves as a placeholder for any form of new environment (e.g., augmented, assisted, mixed, virtual reality) [1]. It is also often defined as an umbrella term as it encompasses the full spectrum of immersive technologies, including VR, AR, and MR [15].

All of these technologies can be used collaboratively and in two ways: co-located and remote. ***Co-located*** refers to physical presence of users in the same location, while ***remote*** involves interaction from different physical locations facilitated by technology [16].

1.3. Immersive Analytics

Immersive Analytics (IA) is a rapidly growing field that explores how new interactive and display technologies support analytical reasoning and decision-making [17]. IA leverages technologies such as XR and Large Interactive Displays (LID) to enable more engaging and immersive experiences for data analysis applications by utilizing their unique affordances [18]. Although immersive and visual analytics share some main goals, such as deriving insights from complex datasets, IA focuses on new immersive technologies and considers multi-sensory interaction in addition to visualization [17].

Academic literature has proposed six opportunities that set IA techniques apart from traditional techniques [19]. These opportunities include spatial immersion, situated analytics, embodied data exploration, collaboration, multi-sensory presentation, and engagement. *Spatial immersion* leverages an additional spatial dimension, allowing users to interact with data more intuitively and naturally. *Situated analytics* involves displaying visualizations in direct reference to their source, providing context that enhances understanding. *Embodied data exploration* uses spatial and natural interactions to manipulate and explore data visualizations, making the analysis process more dynamic and interactive. *Collaboration* in IA explores how users can share and collaborate within the same or different virtual spaces, facilitating teamwork and idea exchange. *Multi-sensory presentation* employs various custom devices to analyze data through senses such as smell or touch, adding a new layer of depth to data interpretation. Finally, *engagement* focuses on using creative methods, such as 3D visualizations or realism, to present data in a way that is more visceral and impactful. By harnessing these opportunities, IA aims to transform traditional data analysis, making it more immersive, interactive, and effective.

In contrast, Ens et al. [20] have conducted research and present 17 key challenges that address the rapidly evolving field of IA. These challenges cover the topics of spatially situated data visualization, interacting with immersive analytics systems, collaborative analytics, and user scenarios and evaluation. The exact challenges can be viewed in Table 1.1.

Table 1.1: Overview of Immersive Analytics Challenges [20].

Topics		**Challenges**
Spatially Situated Data Visualization	C1	Placing Visualizations Accurately in Space
	C2	Extracting and Representing Semantic Knowledge
	C3	Designing Guidelines for Spatially Situated Visualisation
	C4	Understanding Human Senses and Cognition in Situated Contexts
	C5	Applying Spatial Visualisation Ethically
Interacting with IA Systems	C6	Exploiting Human Senses for Interactive IA
	C7	Enabling Multi-Sensory Feedback for IA
	C8	Supporting Transitions around Immersive Environments
	C9	Coping with IA Interaction Complexity
Collaborative Analytics	C10	Supporting Behaviour with Collaborators
	C11	Overcoming Constraints of Reality
	C12	Supporting Cross Platform Collaboration
	C13	Integrating Current Collaboration Practice
	C14	Assessing Collaborative Work
User Scenarios and Evaluation	C15	Defining Application Scenarios for IA
	C16	Understanding Users and Contexts for Evaluation of IA
	C17	Establishing an Evaluation Framework for IA

1.4. Air Quality Index (AQI)

The Air Quality Index serves as a crucial tool in assessing and communicating the quality of ambient air in relation to human health [21]. It provides a standardized way to measure pollutants present in the air and their potential impacts on health.

The AQI is essential for public health as it translates complex air quality data into easily understandable metrics. By quantifying pollutants like particulate matter (PM), ozone (O3), nitrogen dioxide (NO2), sulfur dioxide (SO2), carbon monoxide (CO), and others, the AQI informs individuals and authorities about potential health risks associated with outdoor air quality. The AQI typically ranges from 0 to 500, where lower values indicate better air quality and higher values indicate poorer air quality [22]. Different numerical ranges correspond to different levels of health concern, from *Good* (0-50) to *Hazardous* (301-500), as shown in Table 1.2. Each category comes with associated health implications, helping the public and policymakers make informed decisions regarding outdoor activities and health precautions. The AQI scales can vary between countries due to differences in pollutant thresholds, health-based standards, and averaging times used to calculate the index. For instance, the United States Environmental Protection Agency (EPA) utilizes a scale specific to its air quality standards, which differs from scales used in Europe or Asia [23].

In this study, the American AQI scale is adopted due to its widespread recognition and comprehensive coverage of pollutants relevant to this research. The American scale includes PM2.5, PM10, ozone, carbon monoxide, sulfur dioxide, and nitrogen dioxide, aligning with the pollutants commonly monitored in the context of urban and environmental air quality studies.

Table 1.2: Main AQI Categories [23].

Index Value	Level of Concern	AQI Color	Description of Air Quality
0-50	Good	Green	Air quality is satisfactory, and air pollution poses little or no risk.
51-100	Moderate	Yellow	Air quality is acceptable. However, there may be a risk for some people, particularly those who are unusually sensitive to air pollution.
101-150	Unhealthy for Sensitive Groups	Orange	Members of sensitive groups may experience health effects. The general public is less likely to be affected.
151-200	Unhealthy	Red	Some members of the general public may experience health effects; members of sensitive groups may experience more serious health effects.
201-300	Very Unhealthy	Purple	Health alert: The risk of health effects is increased for everyone.
301-500	Hazardous	Maroon	Health warning of emergency conditions: everyone is more likely to be affected.

2. Related Work

This chapter provides an overview of existing solutions in fields pertinent to the subject of this thesis. The subchapters encompass an in-depth examination of data visualization in immersive environments, collaborative and immersive solutions, and air quality visualization solutions. Each section explores contemporary research and exemplary implementations within these domains.

2.1. Data Visualization in Immersive Experiences

Immersive technologies are widely utilized in education and scientific research to visualize complex scientific data, enabling researchers and students to interact with and comprehend intricate ideas. Brown et al. [24] compare traditional and VR visualizations of molecular structures, concluding that VR can significantly enhance the understanding of complex chemical concepts. Pietikäinen et al. [25] developed a VR application for organic chemistry education, while Xu et al. [26] created a web-based VR application for visualizing and studying macromolecular structures. Zhang et al. [27] showcase VR-assisted platform for integrated visual analysis of DNA/RNA sequences and protein structures. Fernandez et. al [28] created a virtual reality application for the visualization of large biological networks called *GeneNet VR*. Through structured interviews, they concluded that the use of VR for exploring this type of data may be beneficial for finding patterns.

Shifting focus from the themes of biology and chemistry, yet remaining within the broader realm of scientific and educational research, Huang [29] designed a VR-based application to improve the understanding and retention of complex astronomical concepts. Similarly, Pérez-Lisboa et al. [30] developed an AR application for children to support learning astronomical concepts without needing to observe the night sky. Janiszewski et al. [31] demonstrated two VR learning systems in the rock engineering field. One system was developed for training in structural mapping of discontinuities and rock mass characterization, and the other for teaching students how to identify rocks and minerals.

Beyond education and scientific research, visualization can also be used in healthcare through XR technologies. VR provides trainees or inexperienced surgeons with the opportunity to practice in VR surgical simulations. Ghaedania et al. [32] explore the use of AR and VR in spine surgery, noting both limitations and significant potential for the future. Ma et al. [33] propose an enhanced 3D AR system to provide surgeons with a high-quality, magnified 3D AR scene, particularly useful for precision surgery. Eom et al. [34] created *NeuroLens*, an AR-based system for trainee neurosurgeons, offering guidance for completing external ventricular drain catheter placements. Zhao et al. [35] propose an AR training solution that provides a complete visualization of the neonatal endotracheal intubation procedure, along with real-time guidance and assessment. Multiple AR applications also show promise in the field of anatomical visualization [36] [37] [38].

In the domain of music, several applications leverage XR technologies for musical education or to visualize music. Reddy et al. [39] created *Liquid Hands*, an AR music visualizer system, wherein 3D particles react to the flow of music. *Liquid Hands* was created to enrich the music listening experience and bridge the gap between physical and virtual concerts that happened post COVID-19 pandemic. Jin et al. [40] developed a tangible programming system called *Shape of Music*, designed for children. It acts as a bridge between programming and music, stimulating children's creativity.

In the realm of historical preservation and cultural heritage, XR technologies are transforming the presentation and conservation of heritage sites. Bozzelli et al. [41] utilize AR and VR to reconstruct ancient ruins, offering visitors an immersive experience of historical sites as they once existed. Their application allows users to interactively engage with the stories associated with artifacts and artworks. Similarly, Marto and Gonçalves [42] developed *DinofelisAR*, enabling users to view a reconstruction of a Roman Forum superimposed over its current ruins, blending past and present seamlessly.

Several studies have explored immersive analytics as a novel approach to presenting and understanding data. Iquiapaza et al. [43] developed *DeAR*, leveraging desktop and AR technologies to improve data analytics. This approach allows users to interact with both environments simultaneously, choosing the most suitable one depending on the specific data analysis needs.

The following works primarily utilize mobile technology rather than desktop or web platforms. Langner et al. [44] combine an AR headset and mobile device to enhance visual data analysis. Through six scenarios, they showcase a prototype framework that integrates one or multiple mobile displays with 2D or 3D augmentations around, above, or between them. Huang et al. [45] created *SPARVIS*, employing a similar approach that utilizes an AR headset and a mobile device. The headset displays various visualizations, while the mobile device is used to interact with and organize these visualizations in space. Their study focuses on the four main interactions required to use this cross-device framework.

Furthermore, creating immersive data visualizations can be time-consuming and often requires complex low-level programming. To address this, Sicat et al. [46] present *DXR*, a toolkit for building immersive data visualizations based on Unity. Similarly, Cordeil et al. [47] developed *IATK*, an immersive analytics toolkit that can be used as a Unity package. Both *DXR* and *IATK* enable rapid prototyping of immersive data visualizations by using a grammar of visualization primitives that can be configured through a graphical user interface (GUI).

2.2. Collaborative Data Visualization in Immersive Environments

The literature on collaborative immersive analytics compared to single-user immersive analytics is sparse, particularly regarding user studies that examine the impact of integrating both immersion and synchronous collaboration in the data analysis process [48]. As noted in Table 1.1, challenges C10-C14 illustrate the additional difficulties encountered when creating a collaborative immersive data analysis experience. The subsequent sections review the literature on both co-located and remote collaboration immersive solutions, in that order.

Lee et al. [49] focus on data analysis tasks conducted within a co-located, room-sized immersive environment. Participants in their study reported that the setup was effective for maintaining workspace awareness and sharing findings among team members. Garrido et al. [50] present a solution that displays data in the form of a 3D graph within two environments: desktop and VR. They conclude that co-located collaboration is beneficial for data interaction tasks, particularly when dealing with complex data. Additionally, Cavallo et al. [51] introduced *Immersive Insights*, a co-located, hybrid analytics environment for exploratory data analysis (EDA). This environment utilizes high-resolution movable screens to visualize statistical information, complemented by a central projection table that provides an overview of ongoing analyses. It also includes an AR view for interacting with high-dimensional data. The conclusion drawn from this study is that while XR technologies can reduce analysis time and enhance insights, they are not yet capable of completely replacing traditional EDA methodologies.

Legetth et al. [52] introduced *CellexalVR*, a VR system designed for visualizing and analyzing single-cell experiments. This system enables researchers to intuitively collaborate and gain insights into their data regardless of geographical location. Finally, Billinghurst et al. [53] review the integration of immersive technologies with visual analytics to facilitate both co-located and remote collaboration, finding that collaboration is especially advantageous when combined with immersive spaces for data analysis.

2.3. Air Quality Data Visualization

Data visualization has revolutionized the interpretation of complex environmental data, such as air quality metrics. Researchers employ advanced web-based platforms and diverse visualization techniques to thoroughly analyze air quality trends [54] [55] [56] [57] [58] [59]. These tools not only enhance data comprehension and support statistical analysis but also facilitate effective communication and informed decision-making in environmental sciences. This capability is especially crucial for policymakers and researchers, as it extracts meaningful insights from potentially overwhelming datasets, ensuring their relevance and applicability beyond specialist audiences. Given that this thesis focuses on visualizing air quality data, this section and the next provide a brief overview of related work in the field of air quality visualization.

Li et al. [54] introduce an efficient approach for visualizing spatio-temporal air quality data using various methods, including line charts, bar charts, heat maps, and geovisualization techniques. Their study demonstrated that these visualizations significantly enhance data comprehension and support further statistical analysis, focusing specifically on Chinese data. Similarly, Lu et al. [55] created interactive web map visualizations for urban air quality data in major Chinese cities, using custom tools like *timezoom.js* and *symadaptive.js*. Their research emphasized that visual representations provide deeper insights into data analysis compared to raw data tables.

Zeng et al. [56] developed a web-based big data visualization platform designed for air quality data in Taiwan. They found that their platform improves data communication and enables real-time monitoring of data changes over time. Additionally, Chen [57] utilized real-time air quality data from 23 observation points in Beijing, employing Google Earth and Keyhole Markup Language (KML) for visualization. This approach enabled interactive querying and used a gradient color scale to visually highlight air quality levels. The visualization was found to be more expressive and intuitive than raw data tables, enhancing the potential for further data analysis.

Carro et al. [58] present a solution to make vast amounts of air quality environmental data more understandable by exploring different visualization methods. Their study highlighted the advantage of using colors in plots based on AQIs. Teles et al. [59] developed *CityOnStats*, a 3D visualization tool that presents urban air quality data in a game-like environment. The application was designed as an alternative to traditional visualization methods, and their results demonstrated that *CityOnStats* is at least as efficient as conventional techniques. Additionally, it provides an experience similar to gaming, making the application not only efficient but also enjoyable.

2.4. Immersive Air Quality Data Visualization

Recent research has begun to explore the potential of immersive technologies in the field of air quality data visualization. These cutting-edge approaches aim to leverage XR to create more engaging and interactive visualizations of environmental data. The remainder of this subchapter delves into various studies that have implemented XR technologies to enhance the understanding and communication of air quality information. Initially, mobile solutions are discussed, followed by the subsequent consideration of the HMD solution towards the section's conclusion.

Prophet et al. [60] created an application for mobile devices aimed at cultivating environmental awareness. They leveraged AR technology to develop a playable and engaging solution where users plant a virtual tree and choose a location on the globe, such as Hong Kong or Wuhan. The tree's growth depends on real-life air quality data from the chosen location, and users must perform various actions to help their tree grow. Ramachandran et al. [61] present *USC AiR*, a mobile AR platform that translates air quality sensor feeds from the smart campus testbed into augmented reality visualizations for the USC community. The application includes a map with air quality data visualizations, aiming to engage the community by encouraging them to plant virtual trees in areas with suboptimal air quality. Similarly, Mathews et al. [62] showcase *AiR*, an AR application developed for Android mobile devices. Their goal was to make users aware of the air quality in their surroundings. The application retrieves the user's GPS coordinates and then obtains data from the nearest air quality monitoring station. The pollutants are then randomly distributed into the airspace around the user, providing an immersive and immediate understanding of the air quality in their vicinity.

Ultimately, Larsson et al. [63] developed a 3D VR prototype application for complex air data visualization as part of a transdisciplinary project on urban planning and stakeholder communication. The visualization featured both street-level and bird's-eye views and offered several visualization options: volumetric particles, a movable large cut plane with a heatmap, and small cut planes that could be fired from the user's hand.

3. *VisualizeAir*: Design and Functionalities

This chapter provides an insight into the design and key functionalities of the application. The first subchapter offers an overview of *VisualizeAir*, while the second details the data visualization and various user interactions.

3.1. *VisualizeAir* Overview

The application consists of a single scene shown in Figure 3.1. The scene contains four main components: *Main Informations Panel*, *Setup Panel with AQI Categories*, *Earth Renderer Options Panel*, and *the Earth* model. When starting the application, the user is positioned directly into the scene.

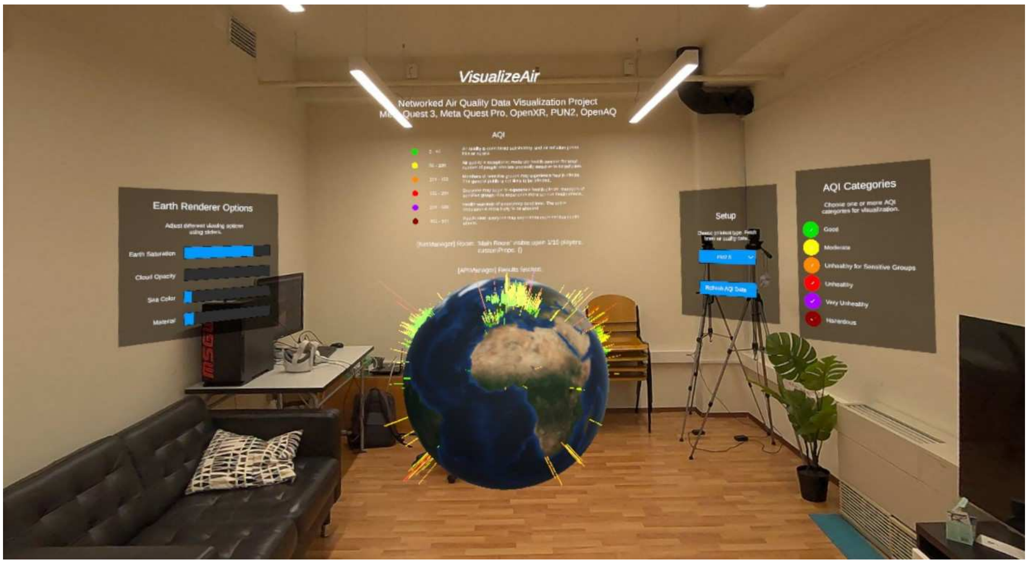

Figure 3.1: Main scene.

The *Main Informations Panel* is designed to introduce users to the application. Its primary components include the application name as the title, a subtitle containing the application description, a chart illustrating the AQI legend and descriptions, and a status text. The AQI legend, similar to Table 1.2, informs the user about how the data will be represented later on. The status text serves as a diagnostic tool in case issues arise with data fetching or the collaboration networking part. The previously mentioned elements are visually represented in Figure 3.2, showing the status text when everything is functioning correctly. This panel is non-interactive and serves solely for informational purposes, as suggested by its name. Users are encouraged to examine the panel at the start of their experience, and may also return to it later if needed.

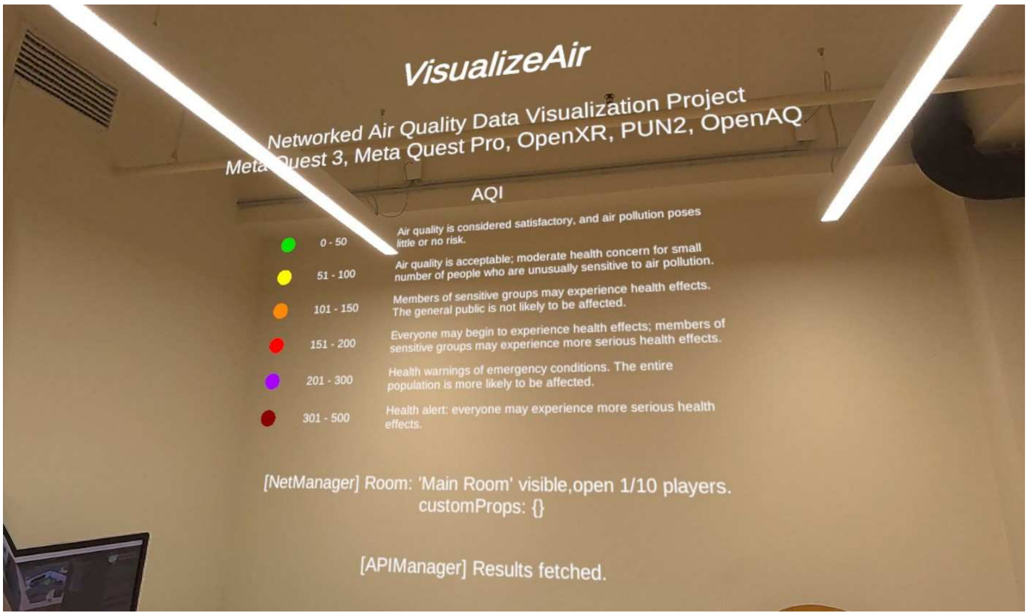

Figure 3.2: Main informations panel.

The first of two interactable panels is the *Setup Panel with AQI Categories*, which consists of two parts. As shown in Figure 3.3, the left side of the panel includes a dropdown menu with all supported pollutant types. By selecting the dropdown button, the user can change the default pollutant, *PM2.5*, to one of the other options: *PM10*, *CO*, *NO2*, *SO2*, and *O3*. Beneath the dropdown menu is another button that, when clicked, sends a new request to update all data points shown on the globe to the most recent values. The right side of the panel features interactable toggle buttons representing the same AQI categories as those seen in the Main Informations Panel. By selecting one of the toggle buttons, the user can decide which data points will be visible on the globe.

To continue, the second of the two interactable panels is the *Earth Renderer Options Panel*. This panel contains four sliders, each with a description written next to them as shown in Figure 3.4. The sliders are *Earth Saturation*, *Cloud Opacity*, *Sea Color*, and *Material*. The first slider, *Earth Saturation*, adjusts the saturation of the globe model.

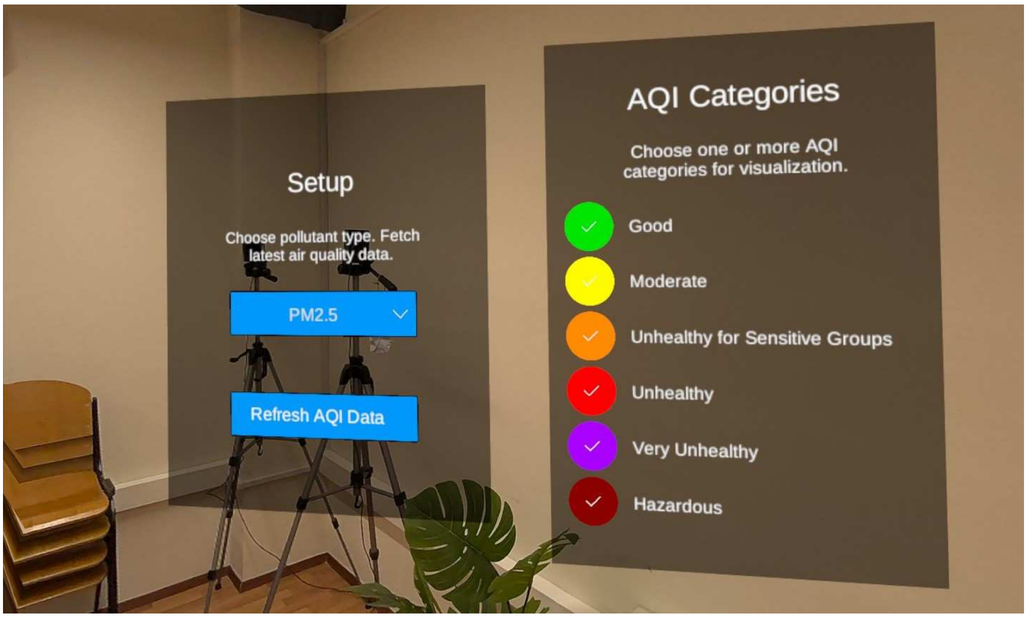

Figure 3.3: Setup panel with AQI categories.

The second slider, *Cloud Opacity*, controls the opacity of the clouds rendered on top of the model. The *Sea Color* slider modifies the blue color of the sea, transitioning from fully opaque blue to different hues depending on the ocean's depth at specific locations. The final slider, *Material*, adjusts the globe's material smoothness, ranging from matte to glossy. Any changes made to the slider values are immediately reflected on the globe model. While data is being fetched, all interactable components, both in this panel and the previously mentioned panel, are greyed out and disabled, as demonstrated in Figure 3.5.

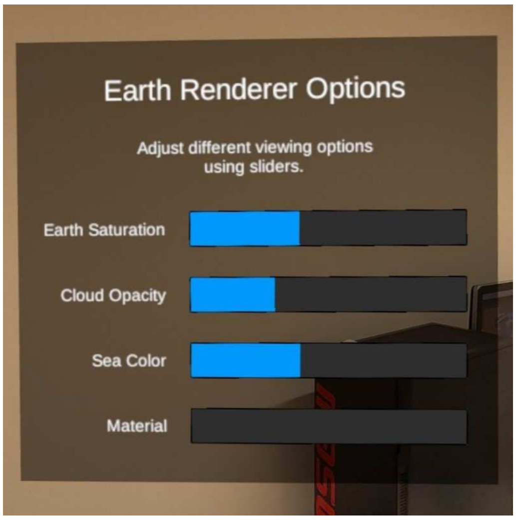

Figure 3.4: Earth renderer options.

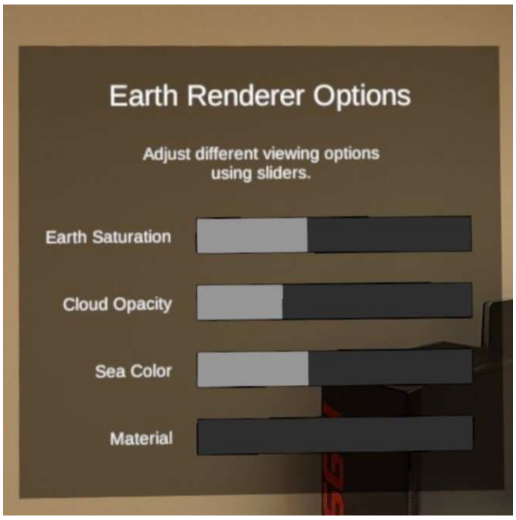

Figure 3.5: Earth renderer options - greyed out.

The fourth and final component is the globe which is the main visualization object in the scene with all others serving as informational and interactable panels only. As shown in Figures 3.6 and 3.7, each data point is represented as a ray emanating from the globe. These rays vary in color and height, with both attributes indicating the amount of pollution at their respective locations. Data visualization and all user interactions are explained further in the rest of the thesis.

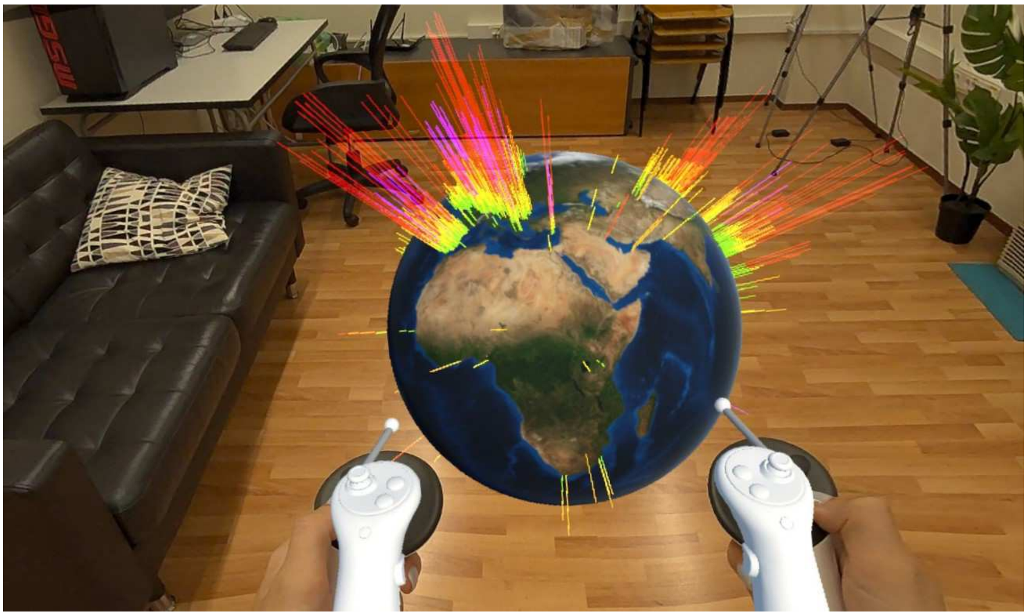

Figure 3.6: Earth with data points.

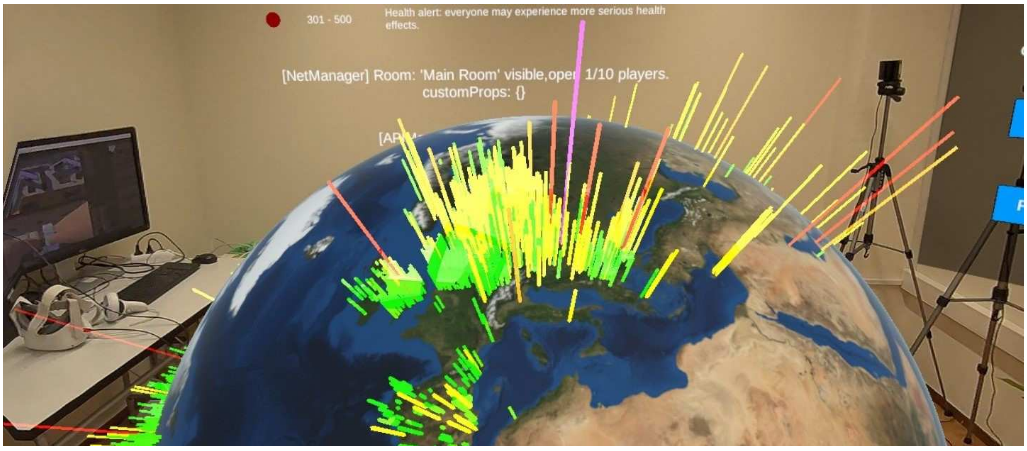

Figure 3.7: Earth with data points - close up.

The typical workflow for using the application involves several steps, visible in Figure 3.8. Initially, the user opens the application and allows a brief period for the data to load and for the application to establish a connection to the collaborative room if multiple users are present. Subsequently, the user reviews the main panel, which displays all pertinent information. Once familiar with this panel, the interactive phase commences. During this phase, the user can modify the appearance of the globe or determine which data is displayed by utilizing toggle buttons and a dropdown menu to select the pollution level for a specific pollutant type.

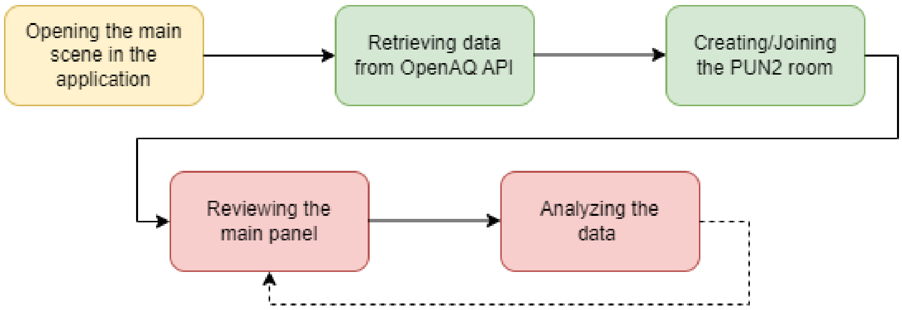

Figure 3.8: Typical workflow for using the application.

3.2. Key Functionalities

One of the key functionalities of the application is the air quality data visualization. As previously noted, the application visualizes data points as rays emanating from the globe. Each ray is uniquely colored and sized: the color indicates the pollution level, and the height reflects the intensity of the pollution. This visualization approach aims to provide an intuitive understanding of pollution levels and their intensities.

Users can interact with the data in real time, manipulating the globe to zoom in on specific regions or adjust viewing angles to explore data from different perspectives. These interactions are facilitated seamlessly using both controllers and hand gestures. Interactive actions include repositioning, rotating, and scaling the globe. UI interactions with buttons and panels utilize the same intuitive techniques.

The collaborative features of the *VisualizeAir* application are designed to foster seamless interaction and shared exploration of air quality data in MR environments. Users benefit from real-time data synchronization, allowing multiple participants to simultaneously view and manipulate data points and the globe as well. Session management features enable easy creation and joining, while object ownership maintains data integrity and control. Real-time updates ensure that all changes made by participants are instantly reflected for all users, while built-in conflict resolution mechanisms handle simultaneous interactions smoothly. These collaborative functionalities aim to enhance user engagement and facilitate collective insights into air quality patterns and trends.

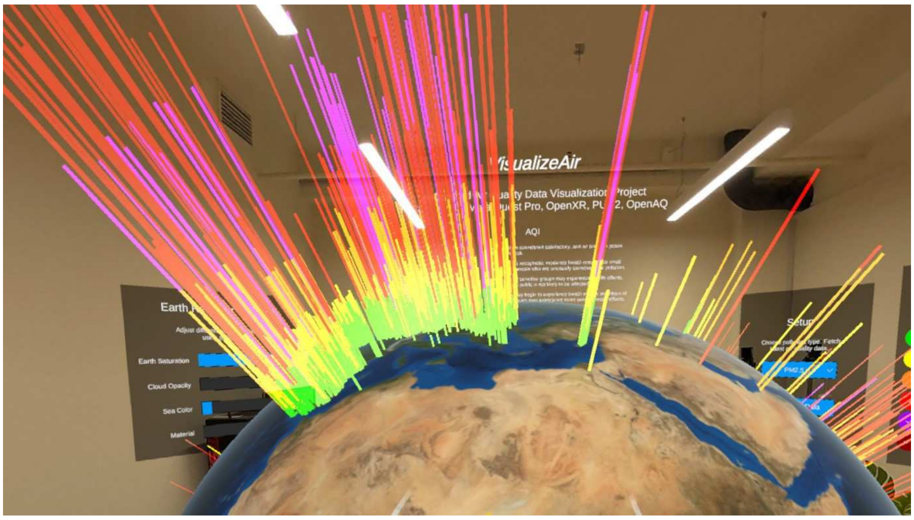

Figure 3.9: Earth scaled up.

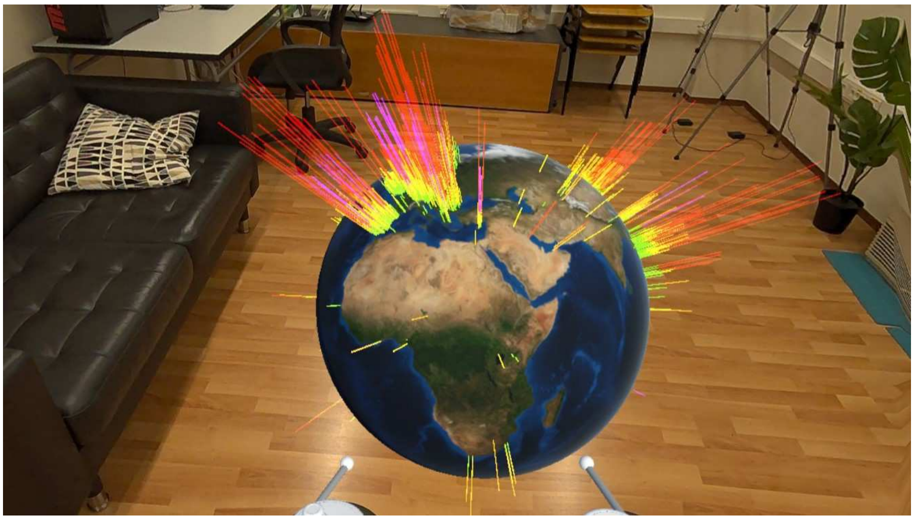

Figure 3.10: Earth scaled down.

4. Development of *VisualizeAir*

The developed collaborative application *VisualizeAir* is an immersive big data visualization experience created to enhance the understanding of large volumes of air quality data. The application was developed and tested for Meta Quest 3 and Meta Quest Pro devices using the Unity platform. Although it has a collaborative component, this application can also be used individually.

This application's primary goal was to make data more intuitive to the user by leveraging all the advantages that immersive technologies offer compared to traditional visualization techniques. Using spatial context should allow users to visualize air quality data in a three-dimensional space, providing a better understanding of the spatial distribution of pollutants globally. Interaction with data in real-time and manipulating the objects in the scene is also expected to help with the comprehension of data. It also tries to enhance comprehension and facilitates the identification of patterns and trends that may be difficult to discern from traditional 2D charts.

The next few subchapters contain an overview of used tools, softwares, and technologies, followed by implementation details and in the end the summary of the limitations of the application.

4.1. Used Technologies, Software and Tools

Unity[1], developed by Unity Technologies, is a cross-platform game engine that offers a comprehensive set of tools for rendering, physics, and input management. These tools enable developers to create both three-dimensional (3D) and two-dimensional (2D) virtual spaces without building everything from scratch [64]. Unity supports a wide variety of platforms, including desktop, mobile, console, and various XR devices. Unity's scripting is primarily done in C#, a general-purpose, multi-paradigm programming language known for its strong typing, object-oriented features, and seamless integration with the engine. For this application, version 2022.3.22f1 of Unity was used.

[1] `https://unity.com/`

Meta Quest 3[2] and **Meta Quest Pro**[3] are headsets developed by Meta, formerly known as Facebook. They both provide a wireless immersive experience, allowing users to move freely without being tethered to a computer. Both headsets feature high-resolution displays, advanced graphics, and improved processing power. They also offer built-in audio, hand tracking, and passthrough. However, Meta Quest Pro includes an additional feature, face and eye tracking.

Photon Unity Networking (PUN2)[4] is a cross-platform networking solution that simplifies the process of adding multiplayer functionality to Unity projects. Developed by Exit Games, PUN2 provides features such as built-in support for matchmaking, client-server architecture, and game object synchronization. Although it was announced that PUN2 would no longer get any new features, it remained a viable choice for this application because of prior experience with the package. Version 2.46 was used for this application.

While developing this application for the Quest 3 and Quest Pro, the open-standard plugin **OpenXR**[5] was utilized. OpenXR, developed by the Khronos Group, ensures cross-platform support and simplifies development by providing a unified Application Programming Interface (API) for various devices. Version 1.10.0 was used during development.

XR Interaction Toolkit[6] is a package created by Unity to help developers create interactive experiences in XR within Unity. It includes a set of tools, scripts, and prefabs that simplify common interactions such as grabbing objects, teleporting, and interacting with UI elements. It provides rapid prototyping while being easy to use and integrate into an existing Unity project. For this application, version 2.5.4 was used. Also an additional package, **XR Hands**[7] version 1.3.0, was imported to enable hand tracking support. By including this package, all the interactions done with controllers can also be done using only hands.

AR Foundation[8] is another Unity package designed to provide a framework for building AR applications. It acts as an abstraction layer over different AR SDKs, enabling developers to create cross-platform AR experiences with a single codebase. The version used in this application was 5.1.4.

[2] https://www.meta.com/quest/quest-3/
[3] https://www.meta.com/quest/quest-pro/
[4] https://www.photonengine.com/pun
[5] https://www.khronos.org/openxr/
[6] https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@3.0/
[7] https://docs.unity3d.com/Packages/com.unity.xr.hands@1.1/
[8] https://docs.unity3d.com/Packages/com.unity.xr.hands@1.1/

Unity OpenXR Meta[9] package relies on both the OpenXR Plug-in and AR Foundation, leveraging their functionalities to enable the creation of AR Foundation apps for the Meta Quest OpenXR runtime. By implementing support for Meta's OpenXR extensions, this package provides seamless integration with Meta devices in AR Foundation projects. With features ranging from session management to advanced tracking capabilities, this package streamlines the development process, ensuring optimal performance and compatibility for Meta device applications. Additionally, it defines AR Foundation subsystem implementations for various AR features, providing detailed documentation to guide developers through the integration process. Version used for *VisualizeAir* was 1.0.1.

All application development was carried out using the aforementioned C#. The development environment utilized was **Visual Studio Community 2022**[10], developed by Microsoft, specifically version 17.9.6.

Unity Asset Store[11] is an online marketplace that offers assets such as 3D models, textures, scripts, and tools. It offers a wide range of free and paid resources, enabling developers to save time and effort by integrating pre-made components into their applications. Some of the models in the application were acquired from Unity Asset Store.

Another model used in this application was obtained from **GitHub**[12]. GitHub is a web-based platform for version control and collaboration, allowing developers to store, manage, and share their code repositories. It also provides a repository where developers can access various 3D models and assets shared by the community, facilitating easy integration into their projects.

4.2. Globe Model Implementation

In the development of *VisualizeAir* application, the Unity project leveraged the *MoonAndEarth* repository available on GitHub, developed by Keijiro Takahashi [65]. This repository provides an implementation for rendering the Earth and Moon within Unity environments, offering high-quality and detailed models suitable for immersive experiences. Figure 4.1 displays the image of the Earth provided by NASA[13].

[9] `https://docs.unity3d.com/Packages/com.unity.xr.meta-openxr@1.0/`
[10] `https://visualstudio.microsoft.com/vs/`
[11] `https://assetstore.unity.com/`
[12] `https://github.com/`
[13] `https://visibleearth.nasa.gov/`

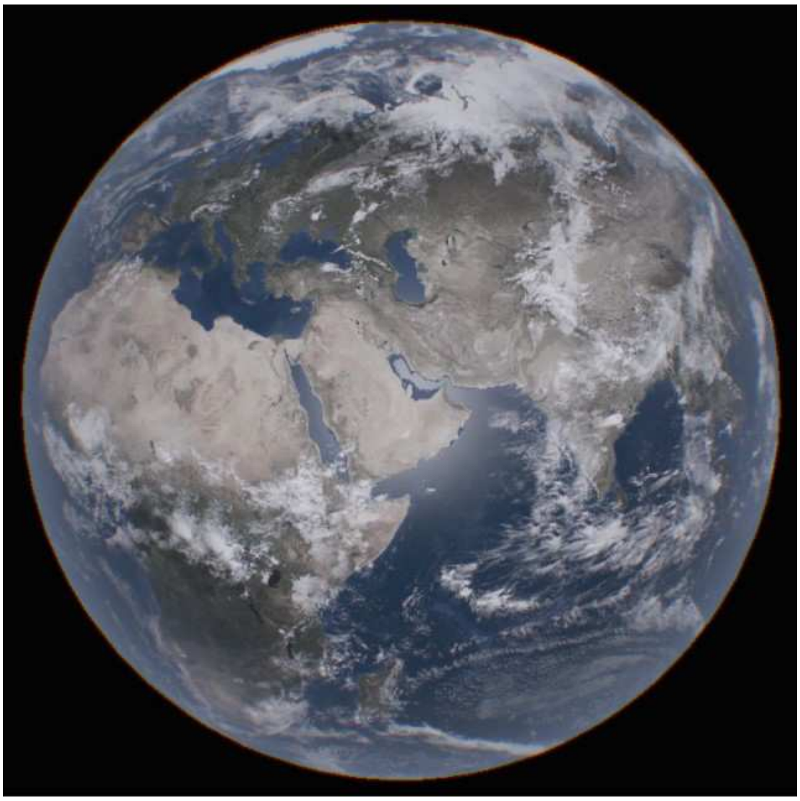

Figure 4.1: Earth material sample [65].

4.3. *OpenAQ* Implementation

Air quality data visualization was chosen as the focal point of this thesis. Following an assessment of a few different APIs (*AirVisual*[14], *WAQI*[15], *OpenWeather*[16], *Forecast*[17], *BreezoMeter*[18]), the *OpenAQ*[19] emerged as the preferred solution due to its easily accessible and large datasets. This API serves as an open-source air quality data platform that aggregates historical and real-time air quality data from diverse sources from various governmental and research sources. Figure 4.2 depicts the aggregation of data from all around the world and its different visualizations.

Data is acquired through the *OpenAQ* REST API. After *OpenAQ* pulls the data from the source, data is parsed into a standard file format, and saved into their database. *OpenAQ* provides programmatic and queryable access to the *OpenAQ* database by using different API calls. Each API call follows the format: `https://api.openaq.org/*`.

[14] `https://airvisual.com/api/documentation`
[15] `https://aqicn.org/api/`
[16] `https://openweathermap.org/api/air-pollution`
[17] `https://plumelabs.com/en/forecast-api/`
[18] `https://docs.breezometer.com/`
[19] `https://openaq.org/`

Figure 4.2: OpenAQ - Data Aggregation and Visualization [66].

Some of the possible API calls are presented in Table 4.1. Additionally, some of these results can be filtered using various parameters such as *limit, page, offset, date_from* and *date_to*.

Table 4.1: API Call Examples.

API Call	**Description**
`/v2/latest`	Provides a list of locations with the latest measurements.
`/v2/cities`	Provides a list of available cities, optionally filtered by country.
`/v2/countries`	Provides a list of all available countries.
`/v2/measurements`	Provides a list of measurements.
`/v2/parameters`	Provides a list of all available parameters (pollutants).
`/v2/sources`	Provides a list of data sources.
`/v2/averages`	Provides average values for specified parameters over a defined time or space.
`/v2/locations/{location_id}`	Provides detailed information about a specific location by its ID.

For the purpose of this thesis, the following GET request was used:

```
https://api.openaq.org/v2/latest?limit=5000
```

This request retrieves a list of locations along with the latest measurements in JavaScript Object Notation (JSON) format. The parameter *limit* was employed to restrict the number of locations returned in the response. These adjustments were made to enhance the application's performance while ensuring an adequate number of data points are displayed on the globe. By limiting the number of locations to 5,000, approximately 15,000 data points are generated, as some locations feature multiple sensors for different types of pollutant particles. This limitation is also discussed mentioned in Section 4.6.

After sending the request, it is necessary to wait a short while for a response. In case there is a network or protocol error, an error message will be shown on *MainInformationPanel* status text. If there is no error, the JSON response is deserialized into an `OpenAQReponse` data object and that data object is then processed further.

After a list of all locations from the response has been created, that list is then filtered to a new list as each location can contain multiple data sources. All unsupported measurements are discarded. Each data point in the new list is a new `DataPoint` object.

To map latitude and longitude values to Unity's 3D space, the following code was used:

```
Vector3 pointPosition;
pointPosition.x = 0.5f * Mathf.Cos((lng) * Mathf.Deg2Rad)
                         * Mathf.Cos(lat * Mathf.Deg2Rad);
pointPosition.y = 0.5f * Mathf.Sin(lat * Mathf.Deg2Rad);
pointPosition.z = 0.5f * Mathf.Sin((lng) * Mathf.Deg2Rad)
                         * Mathf.Cos(lat * Mathf.Deg2Rad);
```

4.4. Interactions Implementation

Repositioning the globe can be accomplished with a single hand or controller. Users initiate grabbing by pinching with their hand or pressing the trigger button on the controller. Rotating the globe is similarly achievable with one hand through a rotational motion while maintaining a grip on the object. With controllers, users can rotate and reposition the globe using the 3D thumbstick once they grabbed it. Scaling the globe involves using both hands or controllers. Users grab the globe in two separate places and adjust their hand positions to zoom in or out on specific areas of the Earth. To scale down the object, the user has to bring hands closer together and to scale it up, hands need to be moved further apart. Interactions with UI objects in the scene use the same intuitive gestures and actions. Selecting a button requires a pinch gesture with one hand or pressing the trigger on the controller. Adjusting slider values involves selecting the slider and dragging it while maintaining a grip. These interactions are specific to Meta Quest 3 and Meta Quest Pro and might not be the same on other HMD devices.

4.5. Multiuser Implementation

To facilitate collaboration between multiple users of this application, a reliable networking solution is essential. Various options are available, such as Unity's built-in networking solution *Netcode*[20], the *Mirror*[21] networking library, and several Photon services including *PUN2*[22] and *Fusion*[23]. Given Photon's reliability and prior experience with this package, *PUN2* was selected as the networking framework for this application. *PUN2* offers all the necessary components for this project and, despite multiple pricing tiers, the free version supports up to 20 clients simultaneously, which is sufficient for the scope of this application. Additionally, it provides features such as a matchmaking API, client-server architecture, Remote Procedure Call (RPC) communication, and synchronization of game objects across the network.

Setting up PUN is a straightforward process, as it can simply be added to Unity like any other Unity package. Once added, the *PUN Wizard* window appears and integrates the `PhotonServerSettings` file into the project. This file is primarily used by the `PhotonNetwork.ConnectUsingSettings()` method, which connects to the Photon server using predefined settings. The predefined settings include properties such as the *App ID* and *Dev Region*. The *App ID* is configured through the Photon Engine Dashboard and then copied into Unity, while the *Dev Region* can be optimally selected within Unity to ensure the application connects to the nearest Photon server. Beyond these configurations, the default settings proved sufficient for the scope of this project.

When implementing multiplayer functionality, several key methods and callbacks were utilized to establish and manage network connections and rooms. The first step in this process was adding the following namespaces to the script:

```
using Photon.Pun;
using Photon.Realtime;
```

Next, a new class called `NetworkManager` was created. This class inherits from `MonoBehaviourPunCallbacks`, which provides a `photonView`, various callbacks, and enables the overriding of all necessary events and methods. Several callback methods are implemented to handle various network events. Table 4.2 shows all the methods used in the scope of this project along with their descriptions.

[20] `https://unity.com/products/netcode`
[21] `https://mirror-networking.com/`
[22] `https://www.photonengine.com/pun`
[23] `https://www.photonengine.com/fusion`

Table 4.2: Used PUN methods and their descriptions.

Method	**Method Description**
`OnConnectedToMaster()`	Called when the client is connected to the Master Server and ready for matchmaking and other tasks.
`OnCreatedRoom()`	Called when this client created a room and entered it. `OnJoinedRoom()` will be called as well.
`OnJoinedRoom()`	Called when the `LoadBalancingClient` entered a room, no matter if this client created it or simply joined.
`OnLeftRoom()`	Called when the local user/client leaves a room, so the game's logic can clean up its internal state.
`OnDisconnected()`	Called to signal that the raw connection got established but before the client can call an operation on the server.

The method `PhotonNetwork.JoinOrCreateRoom()` is invoked to either join an existing room with the specified name or create a new room if one does not already exist. Only the Master client, meaning the first client to connect, can create a new room, while the other clients can only join. The `OnConnectedToMaster()` callback is triggered when the application successfully connects to the Photon server, prompting the aforementioned creation or joining of a room. `OnCreatedRoom()` and `OnJoinedRoom()` callbacks are executed when the room is successfully created or joined, respectively. After the player joins the room a new network player is instantiated using `PhotonNetwork.Instantiate()` method. The `OnLeftRoom()` callback is triggered when the local player leaves the room, enabling cleanup tasks, such as destroying the network player, to be executed. Lastly, the `OnDisconnected()` callback is implemented to handle disconnection events, providing feedback or performing cleanup actions based on the cause of the disconnection. Together, these methods and callbacks facilitate the establishment and management of network connections and rooms within the application, ensuring seamless multiplayer interactions for users. Users can know if connecting to the Photon server and joining the room was successful by looking at the status text on the *Main Informations Panel*.

Photon is employed for several key networking tasks, including creating Photon rooms, synchronizing objects across the network, and using RPCs to synchronize user inputs. The presence of a `PhotonView` component signifies that an object is intended for networked functionality, with each object possessing a unique View ID. To enact methods that affect multiple users across the network, these methods must be annotated with `[PUNRPC]`. For instance, a method responsible for tracking changes made to a slider would be paired with another method ensuring that any changes made by one client are promptly propagated to all other networked clients. All methods annotated with `[RPC]` are listed in the Table 4.3.

Table 4.3: RPC methods used in the thesis.

RPC Method	**Method Description**
`NetworkedPollutantType-DropdownValueChanged()`	Called when a user selects different pollutant type by clicking on a dropdown menu.
`NetworkedOnColorButton-ValueChanged()`	Called when a user toggles a color on the globe on or off.
`NetworkedOnEarthSatu-rationSliderValueChanged()`	Called when a user changes earth saturation on a globe model by dragging a slider.
`NetworkedOnCloudOpacity-SliderValueChanged()`	Called when a user changes cloud opacity on a globe model by dragging a slider.
`NetworkedOnSeaColor-SliderValueChanged()`	Called when a user changes sea color on a globe model by dragging a slider.
`NetworkedOnMaterialSmooth-nessSliderValueChanged()`	Called when a user changes material smoothness on a globe model by dragging a slider.

4.6. Limitations and Future Work

Several limitations were encountered during the development phase of *VisualizeAir*. This section acknowledges these challenges to enhance the reliability and resilience of the application in future iterations. Key limitations identified include issues with the anchoring system, performance challenges with large datasets, and concerns regarding networking reliability, which are elaborated upon in the following sections. Addressing these limitations will be essential for future work to enhance the functionality and user experience of *VisualizeAir*.

In the development of the collaborative aspect of our mixed reality application, a significant limitation has been identified due to the lack of an anchoring system. Time constraints during development prevented the implementation of an anchoring solution. Although *Azure Spatial Anchors*[24] was proposed as a reliable solution, issues with accessing the *Azure portal* and insufficient time to explore alternative options hindered its integration. Consequently, the application encounters substantial challenges in co-located environments. For example, if two or more users are using this application in the same physical space, they can see each other's representations both inside the virtual environment as well as in the real space, but those representations will not be aligned. All users should see the same virtual objects in the same locations relative to the physical space around them. Misalignment between their virtual worlds and objects can make users question their understanding of a unified environment. Incorporating an effective anchoring solution will be a priority for future improvements to ensure reliable and accurate MR interactions in shared physical spaces.

Due to the substantial size of the dataset, it was anticipated that certain limitations would arise. Handling large volumes of data imposes significant computational demands, which has been evident during the development of *VisualizeAir*. Particularly noticeable is the delay experienced during application startup, attributable to data retrieval from the OpenAQ API. Furthermore, the visualization suffers from jitteriness and lag when a high number of data points are loaded into the scene. Due to this issue, a compromise had to be struck between having too many data points, which caused the application to become jittery, and having too few data points, which resulted in poor visualization quality. It can be observed that disabling certain data categories using toggle buttons and thereby reducing the number of points leads to a significantly marked improvement in performance, resulting in a smoother scene.

Finally, there were intermittent connection problems when devices attempted to join the room. Each user is supposed to join the same room specified in the code. Typically, one user starts the application and joins the default room, followed by another user who should also join the same room upon starting the application. However, despite joining the room with identical specifications, users occasionally ended up in separate rooms. This inconsistency complicated the testing process and affected the assessment of the application's collaborative features. Since this issue occurred sporadically, further investigation is required to determine the root cause, which might be related to PUN2 inconsistencies. Additionally, it is worth noting that all devices in the study were connected to the same shared mobile hotspot which also has to be taken into consideration when investigating this problem.

[24]`https://azure.microsoft.com/en-us/products/spatial-anchors`

5. User Study Methodology

As part of this thesis, a a small scale pilot user study was conducted with the goal of investigating the impact of immersive technologies such as MR in collaborative big data visualizations. This study explores how immersive and interactive visualizations can enhance the understanding and analysis of large datasets. Table 5.1 contains the research questions (RQs) introduced to fulfill this goal.

Table 5.1: Research Questions.

RQ1	*How satisfied are users overall with MR-based big data visualization techniques compared to traditional 2D visualization methods?*
RQ2	*To what extent does the immersive nature of MR improve user engagement and understanding of complex datasets?*
RQ3	*What are the challenges and benefits of real-time collaboration in MR for big data analysis?*

5.1. Hardware and Software Setup

The study was conducted using Meta Quest 3 and Meta Quest Pro headsets. Participants used the Meta Quest 3 exclusively, while the facilitator always used the Meta Quest Pro. Both devices were connected to the Internet to access the air quality data and enable communication and collaboration. The Internet connection was secured through a shared Wi-Fi hotspot, facilitated by a mobile phone with an active mobile internet connection. Each participant had access to a computer for accessing the questionnaires, which were presented via an online form hosted on the Google Forms[1] platform.

[1] https://www.google.com/forms/about/

5.2. Procedure

Participants arrived individually at their designated time slots, with each session lasting approximately half an hour. Upon arrival, the participant was introduced to the theme and details of the user study they were participating in. Following that, the study member was provided with a consent form to sign and then guided to a computer where they began with the first part of the research study. This initial phase included a demographic sub-questionnaire and inquiries about their familiarity with immersive technologies, data visualization, and collaborative tools. The questionnaire is further detailed in Section 5.4. Once the first part of the questionnaire was completed, the participant received instructions on interacting with objects in MR and using the *VisualizeAir* application. A study facilitator was present throughout to address any questions or assist with potential issues. Each participant had five minutes to familiarize themselves with the application and explore the virtual environment independently. Following this, the study facilitator joined using another device to test the collaborative features. After spending an additional 5-10 minutes discussing and analyzing the data, the participant proceeded to complete the second part of the questionnaire, which focused on their experience using the application. Figure 5.1 displays the outline of the study procedure.

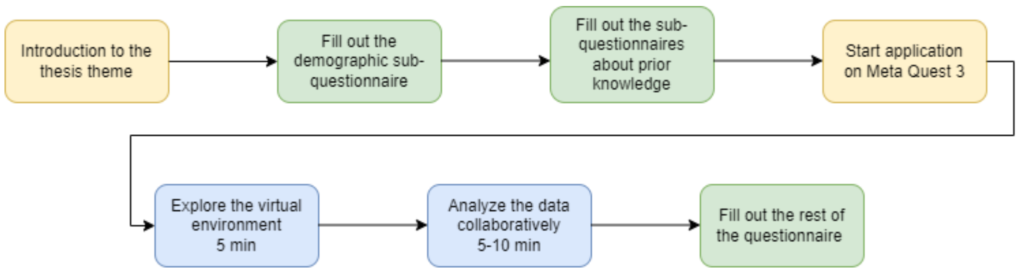

Figure 5.1: User study procedure outline.

5.3. Participants

This user study involved six participants, equally divided between *male* and *female*. Participants' ages ranged from 21 to 24, with a mean age of 23.17 and a median age of 24. Each participant rated their general familiarity with technology on a 5-point Likert scale. Most participants rated their familiarity as 4, resulting in a mean rating of 4.16 and indicating a high level of comfort with technology.

Participants were queried about their prior experience with immersive technologies such as VR, AR, or MR and all participants reported having used such technologies before. Regarding the utilization frequency, half had only tried it once, while the other half had used it a few times. None of the participants had extensive experience using these technologies on a yearly, monthly, or weekly basis. When asked about their proficiency with XR technologies, both mean and median values were 2 on a 5-point Likert scale, which correlates with their usage rate. Table 5.2 contains responses regarding the contexts in which participants used XR technologies. Similarly, Table 5.3 lists the types of devices participants had previously used. When prompted about their comfort with navigating and interacting within XR environments the mean comfort rating was 3.17.

Table 5.2: Contexts in which participants have already encountered XR technologies.

Context	User Count
Gaming	2
Work	0
School	2
Research	3
Social Interaction & Communication	1

Table 5.3: Types of XR devices participants used before.

Device Type	User Count
HMD	6
Mobile phone/tablet	2
Smart Glasses	2
HUD	2

Participants were also asked to self-report their proficiency and prior experience with data visualization. They rated their experience with traditional visualization tools such as *Excel*, *Tableau*[2], and *PowerBI*[3]. The mean rating was 3.17, with a median of 3. Most participants indicated that they rarely use visualization tools in their daily lives and are not familiar with analyzing large datasets. Table 5.4 lists the types of visualizations with which the participants in the study are most comfortable.

Finally, prior experience with collaborative tools was assessed. Five out of six participants reported frequent use of collaborative tools such as *Slack*[4] and *Microsoft Teams*[5].

[2] `https://www.tableau.com/`
[3] `https://www.microsoft.com/en-us/power-platform/products/power-bi`
[4] `https://slack.com/`
[5] `https://www.microsoft.com/en-us/microsoft-teams/group-chat-software`

Table 5.4: Visualizations participants are familiar with.

Visualization Type	User Count
Line Graphs	6
Bar Charts	6
Scatter Plots	4
Heat Maps	0

When asked about the frequency of collaborating with others while analyzing data, the mean response was 2.5, with a median of 2. None of the participants had reported ever using a collaborative MR application.

5.4. Questionnaire Description

A questionnaire consisting of 34 distinct questions was developed for this user study. Table 5.5 illustrates the various types of questions, including examples from the questionnaire. The questionnaire, which is available in its entirety in Appendix A, is organized into several sub-sections: demographic information is collected first, followed by brief sections on prior experience with XR, data visualization, and collaborative tools. Each of these sub-sections contains between 3 to 5 questions. In the end, participants complete a sub-questionnaire on their general experience after testing the application. The last section includes 17 questions, two of which are optional and used for providing additional details on previous answers.

Table 5.5: Questionnaire question types and examples.

Question Type	Question Example
closed with yes/no answers	*Have you ever used collaborative MR applications?*
open-ended	*Were there any interaction techniques that you found particularly helpful or difficult to use?*
multiple choice	*Which features of the MR-based visualization did you find most useful?*
5-point Likert scale	*How satisfied are you with the overall experience using the MR-based big data visualization application?*

6. User Study Results and Discussion

This chapter delves into the findings and analysis derived from the user study conducted to evaluate the *VisualizeAir* application. This study aims to gather insights into user satisfaction and data comprehension when comparing MR-based visualizations to traditional 2D visualizations. The feedback from participants highlights both the strengths and areas for improvement of the application, providing a comprehensive understanding of the user experience of the application.

6.1. Mixed Reality and Interaction Experience

83.3% of participants reported being *very satisfied* with the overall experience of using the application, as shown in Figure 6.1. While their satisfaction was high, it was not rated as *extremely satisfied*, indicating that there is still room for improvement. However, the overall experience remains positive. Figure 6.2 illustrates that 66.7% of users *strongly agrees* with the statement, *I felt immersed in the MR environment while analyzing data*.

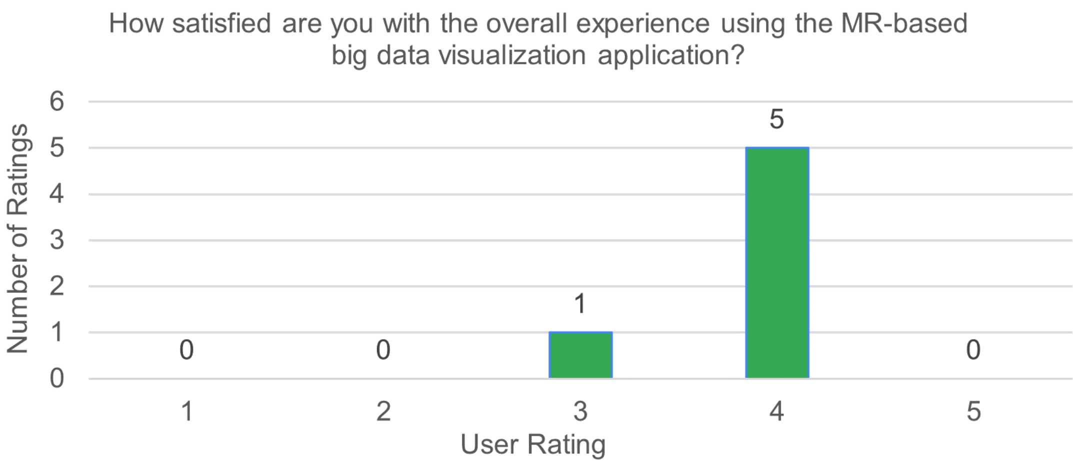

Figure 6.1: Satisfaction with the overall experience.

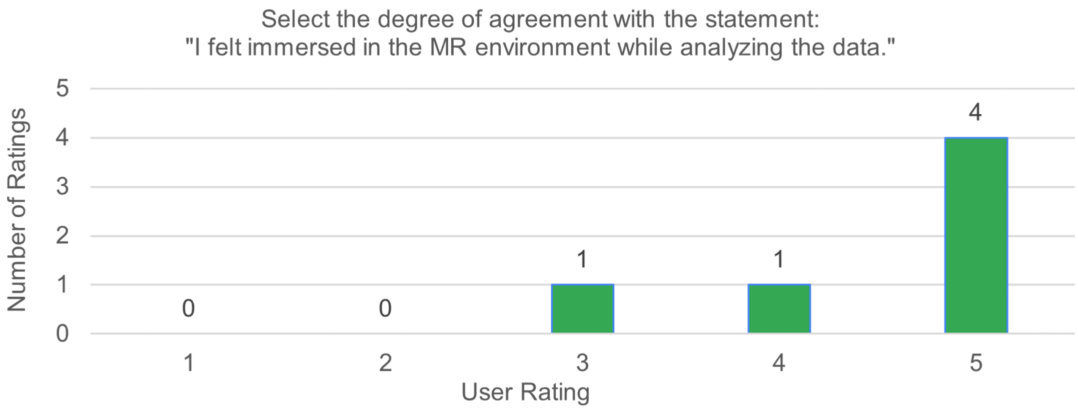

Figure 6.2: Feeling of immersion in the MR environment.

Rating the experience of learning to use an MR-based application and navigating within it, it was expected that there might be some difficulties since most of the users do not have extensive prior knowledge of XR technologies. However, Figures 6.3 and 6.4 show that for most users it was *very easy* to learn how to use this application, compared to navigating which is rated as just a bit harder.

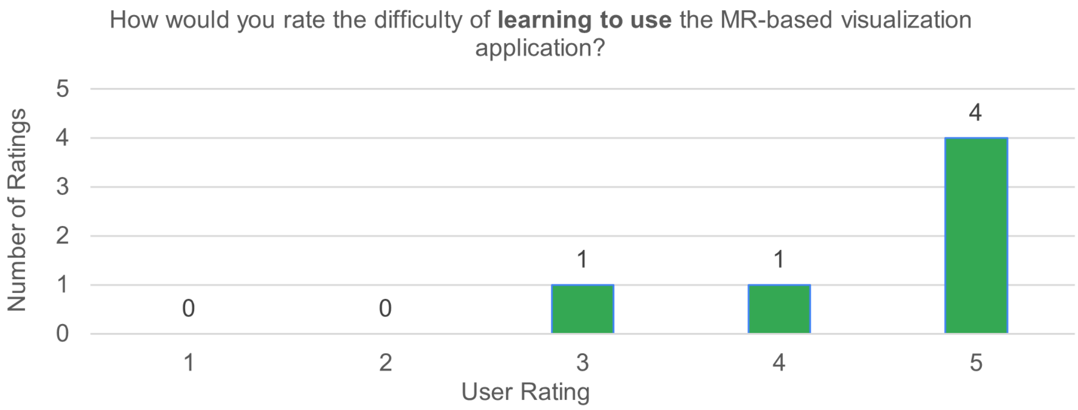

Figure 6.3: Getting familiar with the application.

The intuition was also assessed through several questions. All participants responded affirmatively when asked if they found the geospatial visualization of data points intuitive. Further elaboration revealed that they appreciated the Earth model as a globe as it was clear where they had to look to find the desired information. Figure 6.5 demonstrates that participants rated interaction techniques in the MR environment as *extremely intuitive* (50%) or *very intuitive* (50%). Upon further discussion, half of the participants identified the pinching action used for object selection with their hands as the most challenging interaction technique.

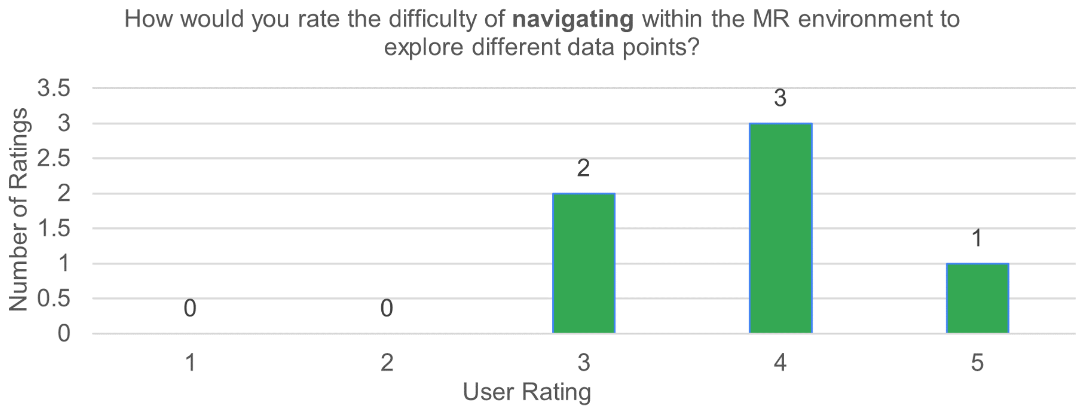

Figure 6.4: Navigating through the application.

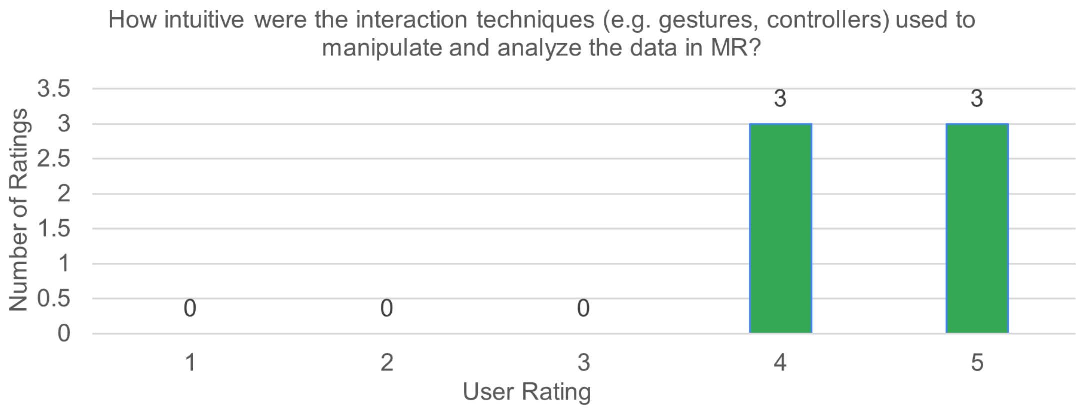

Figure 6.5: Intuitiveness of interaction techniques.

6.2. Data Visualization Experience

As Figure 6.6 shows, half of the participants expressed that MR visualization was *very effective* in understanding the structure and patterns within the data, and the other half responded that it was *extremely effective*. Additionally, the participants were asked to express how they would rate the difficulty of drawing insights from the MR visualization compared to traditional 2D visualizations. Figure 6.7 displays that 66.6% of them rated it as *very easy*. In addition, geospatial visualization was identified as the most useful feature by all six participants, followed closely by data manipulation, which four out of six participants found particularly beneficial. Finally, when asked to choose between MR-based visualizations and traditional visualization methods for data analysis tasks, 50% chose the first and 50% the second. The first half reported that traditional methods are more accessible and that MR is not yet completely natural to use, while the second half commented on MR being more fun and more natural to look at and interact with.

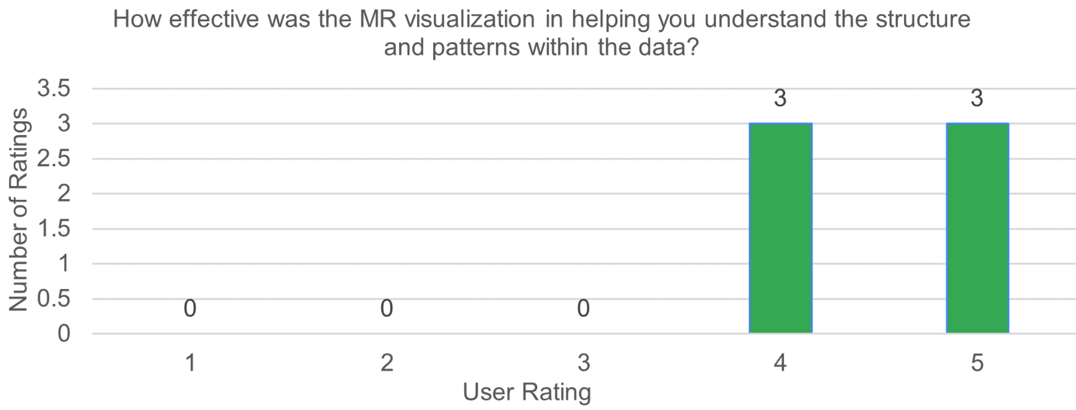

Figure 6.6: Effectiveness of data visualization.

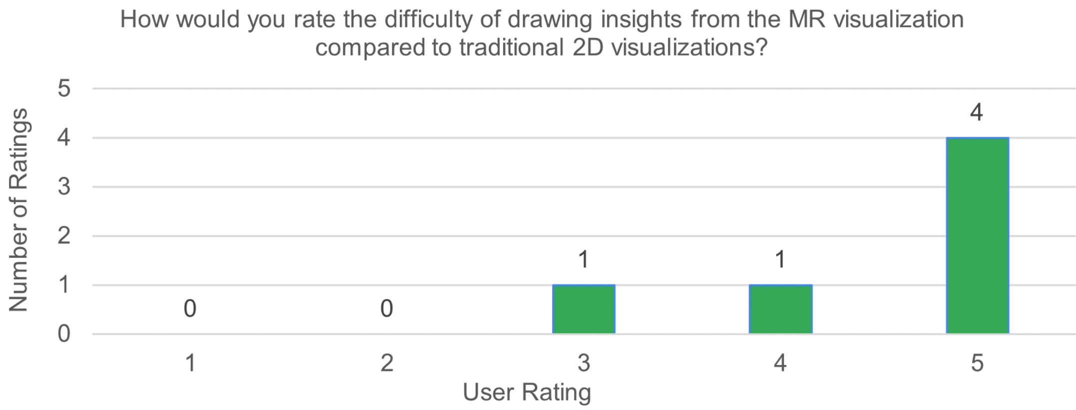

Figure 6.7: Difficulty of drawing insights from this MR visualization.

6.3. Collaboration Experience

In terms of the collaborative aspect of *VisualizeAir*, participants expressed varying degrees of insight and feedback, as represented in Figure 6.8. Half of the participants rated collaboration as *very easy*, while the remaining half was distributed between *neutral* (33.3%) and *easy* (16.7%). This distribution shows a more pronounced variation in user responses, which can be attributed to the networking issues already addressed in Section 4.6. This correlation was further substantiated by participants' comments, which highlighted instances where collaboration initially failed and required restarting the application once or multiple times. It is also important to note that for this part of the study, each participant was paired with a study facilitator they already knew. This should be taken into account when interpreting the results.

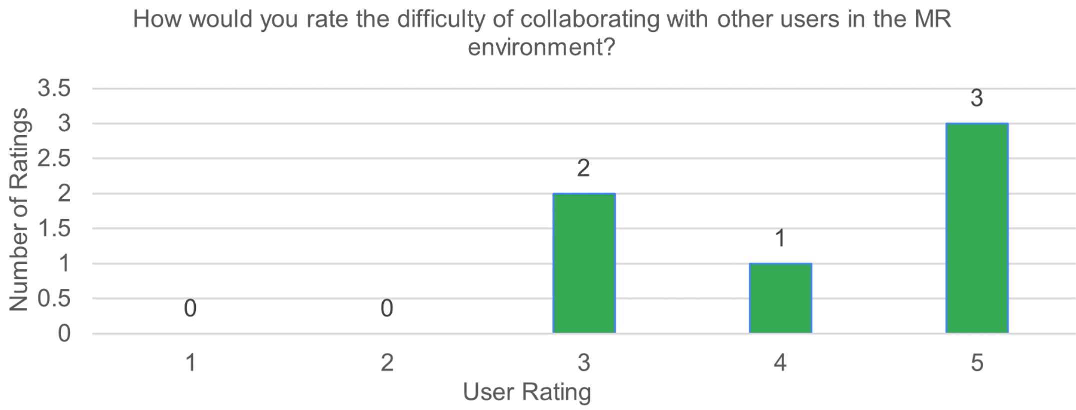

Figure 6.8: Difficulty of collaborating with other users.

6.4. Limitations

The study encountered several limitations that impacted its outcomes. Primarily, the number of participants was limited, meaning there was no sufficient generalization and variation in sample size within the group. The absence of standardization in the questionnaire may have introduced variability in how participants interpreted and responded to the questionnaire items, potentially affecting the reliability of the collected data. Hence, when interpreting the study results, caution must be exercised. Furthermore, issues with the collaborative features were frequently reported. Participants identified problems with collaboration and encountered issues with jitteriness and lag during usage. These technical challenges likely influenced their feedback and satisfaction levels. Nevertheless, the study demonstrates that the design of the *VisualizeAir* application effectively facilitated user comprehension and interaction ease. Addressing these limitations in future iterations of the application will be crucial to enhance the overall user experience and improve reliability.

CONCLUSION

This thesis includes research, development, and evaluation of *VisualizeAir*, a collaborative big data visualization application that uses MR. Through a thorough exploration of theoretical foundations including big data visualization techniques, immersive technologies, and immersive analytics, as well as an overview of related work in immersive data visualization and collaborative environments, this study has contributed to understanding the potential of MR in data visualization. *VisualizeAir* was designed to use a geospatial visual representation of air quality data on a model of Earth, enabling intuitive exploration and manipulation in three dimensions.

The development process highlighted key functionalities and the integration of technologies such as the globe model and air quality data. Methodologically, the user study provided insights into MR interaction, data visualization experiences, and collaborative functionalities.

Despite limitations such as a small participant sample and technical network challenges, *VisualizeAir* demonstrated promising outcomes in enhancing user engagement and comprehension of complex datasets. Looking ahead, future enhancements could address identified limitations and further refine *VisualizeAir*'s capabilities, thereby advancing the application of MR in diverse fields of data visualization.

BIBLIOGRAPHY

[1] International Telecommunication Union. Itu-t recommendation p.1320 - qoe assessment of extended reality (xr) meetings, 2022.

[2] Maxim Spur. *Immersive Visualization of Multilayered Geospatial Urban Data for Visual Analytics and Exploration*. PhD thesis, École centrale de Nantes, 2021.

[3] Seref Sagiroglu and Duygu Sinanc. Big data: A review. In *2013 international conference on collaboration technologies and systems (CTS)*, pages 42–47. IEEE, 2013.

[4] Dan Garlasu, Virginia Sandulescu, Ionela Halcu, Giorgian Neculoiu, Oana Grigoriu, Mariana Marinescu, and Viorel Marinescu. A big data implementation based on grid computing. In *2013 11th RoEduNet International Conference*, pages 1–4. IEEE, 2013.

[5] Doug Laney et al. 3d data management: Controlling data volume, velocity and variety. *META group research note*, 6(70):1, 2001.

[6] bigdatacharacteristics, 2023. URL `https://www.artera.net/en/data-science-en/big-data-characteristics-3v-5v-10v-14v/#:~:text=The%20first%203%20Vs%20of,%3A%20Volume%2C%20Variety%2C%20Velocity`. Last Accessed: 2024-May-30.

[7] Ibrar Yaqoob, Ibrahim Abaker Targio Hashem, Abdullah Gani, Salimah Mokhtar, Ejaz Ahmed, Nor Badrul Anuar, and Athanasios V Vasilakos. Big data: From beginning to future. *International Journal of Information Management*, 36(6):1231–1247, 2016.

[8] ANM Bazlur Rashid. Access methods for big data: current status and future directions. *EAI Endorsed Transactions on Scalable Information Systems*, 4(15), 2018.

[9] Syed Mohd Ali, Noopur Gupta, Gopal Krishna Nayak, and Rakesh Kumar Lenka. Big data visualization: Tools and challenges. In *2016 2nd International conference on contemporary computing and informatics (IC3I)*, pages 656–660. IEEE, 2016.

[10] James D Miller. *Big data visualization*. Packt Publishing Ltd, 2017.

[11] Parul Gandhi and Jyoti Pruthi. Data visualization techniques: traditional data to big data. *Data Visualization: Trends and Challenges Toward Multidisciplinary Perception*, pages 53–74, 2020.

[12] Rajeev Agrawal, Anirudh Kadadi, Xiangfeng Dai, and Frederic Andres. Challenges and opportunities with big data visualization. In *Proceedings of the 7th International Conference on Management of computational and collective intElligence in Digital EcoSystems*, pages 169–173, 2015.

[13] International Telecommunication Union. Itu-t recommendation g.1035 - influencing factors on quality of experience for virtual reality services, 2021.

[14] Chenyan Zhang. The why, what, and how of immersive experience. *Ieee Access*, 8: 90878–90888, 2020.

[15] Philipp A Rauschnabel, Reto Felix, Chris Hinsch, Hamza Shahab, and Florian Alt. What is xr? towards a framework for augmented and virtual reality. *Computers in human behavior*, 133:107289, 2022.

[16] Moinak Ghoshal, Juan Ong, Hearan Won, Dimitrios Koutsonikolas, and Caglar Yildirim. Co-located immersive gaming: A comparison between augmented and virtual reality. In *2022 IEEE Conference on Games (CoG)*, pages 594–597. IEEE, 2022.

[17] Tom Chandler, Maxime Cordeil, Tobias Czauderna, Tim Dwyer, Jaroslaw Glowacki, Cagatay Goncu, Matthias Klapperstueck, et al. Immersive analytics. In *2015 Big Data Visual Analytics (BDVA)*, pages 1–8. IEEE Computer Society, 2015.

[18] David Saffo, Sara Di Bartolomeo, Tarik Crnovrsanin, Laura South, Justin Raynor, Caglar Yildirim, and Cody Dunne. Unraveling the design space of immersive analytics: A systematic review. *IEEE Transactions on Visualization and Computer Graphics*, 2023.

[19] Kim Marriott, Jian Chen, Marcel Hlawatsch, Takayuki Itoh, Miguel A Nacenta, Guido Reina, and Wolfgang Stuerzlinger. Immersive analytics: Time to reconsider

the value of 3d for information visualisation. *Immersive analytics*, pages 25–55, 2018.

[20] Barrett Ens, Benjamin Bach, Maxime Cordeil, Ulrich Engelke, Marcos Serrano, Wesley Willett, Arnaud Prouzeau, Christoph Anthes, Wolfgang Büschel, Cody Dunne, et al. Grand challenges in immersive analytics. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2021.

[21] AQIStandard, 2024. URL `https://www.who.int/tools/air-quality-standards/`. Last Accessed: 2024-June-05.

[22] AQIBasics, 2024. URL `https://www.airnow.gov/aqi/aqi-basics/`. Last Accessed: 2024-June-05.

[23] AQI, 2024. URL `https://www.airnow.gov/aqi/aqi-basics/`. Last Accessed: 2024-May-30.

[24] Corina E Brown, Dalal Alrmuny, Mia Kim Williams, Ben Whaley, and Richard M Hyslop. Visualizing molecular structures and shapes: A comparison of virtual reality, computer simulation, and traditional modeling. *Chemistry Teacher International*, 3(1):69–80, 2021.

[25] Otso Pietikäinen, Perttu Hämäläinen, Jaakko Lehtinen, and Antti J Karttunen. Vrchem: a virtual reality molecular builder. *Applied Sciences*, 11(22):10767, 2021.

[26] Kui Xu, Nan Liu, Jingle Xu, Chunlong Guo, Lingyun Zhao, Hong-Wei Wang, and Qiangfeng Cliff Zhang. Vrmol: an integrative web-based virtual reality system to explore macromolecular structure. *Bioinformatics*, 37(7):1029–1031, 2021.

[27] Jimmy F Zhang, Alex R Paciorkowski, Paul A Craig, and Feng Cui. Biovr: a platform for virtual reality assisted biological data integration and visualization. *BMC bioinformatics*, 20:1–10, 2019.

[28] Álvaro Martínez Fernández, Lars Ailo Bongo, and Edvard Pedersen. Genenet vr: Interactive visualization of large-scale biological networks using a standalone headset. *arXiv preprint arXiv:2109.02937*, 2021.

[29] Huizhong Huang. *Exploring the Cosmos: Prototype of a Gesture-Controlled VR System for Interactive Astronomical Education*. PhD thesis, CALIFORNIA STATE UNIVERSITY SAN MARCOS, 2024.

[30] Sandra Pérez-Lisboa, Carmen Gloria Ríos-Binimelis, and J Castillo Allaria. Augmented reality and stellarium: Astronomy for children of five years. *Alteridad*, 15: 25–35, 2020.

[31] Mateusz Janiszewski, Lauri Uotinen, Jeremiasz Merkel, Jussi Leveinen, and Mikael Rinne. Virtual reality learning environments for rock engineering, geology and mining education. In *ARMA US Rock Mechanics/Geomechanics Symposium*, pages ARMA–2020. ARMA, 2020.

[32] Hamid Ghaednia, Mitchell S Fourman, Amanda Lans, Kelsey Detels, Hidde Dijkstra, Sophie Lloyd, Allison Sweeney, Jacobien HF Oosterhoff, and Joseph H Schwab. Augmented and virtual reality in spine surgery, current applications and future potentials. *The Spine Journal*, 21(10):1617–1625, 2021.

[33] Longfei Ma, Zhencheng Fan, Guochen Ning, Xinran Zhang, and Hongen Liao. 3d visualization and augmented reality for orthopedics. *Intelligent Orthopaedics: Artificial Intelligence and Smart Image-guided Technology for Orthopaedics*, pages 193–205, 2018.

[34] Sangjun Eom, David Sykes, Shervin Rahimpour, and Maria Gorlatova. Neurolens: Augmented reality-based contextual guidance through surgical tool tracking in neurosurgery. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 355–364. IEEE, 2022.

[35] Shang Zhao, Xiao Xiao, Qiyue Wang, Xiaoke Zhang, Wei Li, Lamia Soghier, and James Hahn. An intelligent augmented reality training framework for neonatal endotracheal intubation. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 672–681. IEEE, 2020.

[36] Ma Meng, Pascal Fallavollita, Tobias Blum, Ulrich Eck, Christian Sandor, Simon Weidert, Jens Waschke, and Nassir Navab. Kinect for interactive ar anatomy learning. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 277–278. IEEE, 2013.

[37] Chung Le Van, Trinh Hiep Hoa, Nguyen Minh Duc, Vikram Puri, Tung Sanh Nguyen, and Dac-Nhuong Le. Design and development of collaborative ar system for anatomy training. *Intelligent Automation and Soft Computing*, 27(3):853–871, 2021.

[38] Cita Nørgård, L O'Neill, Kurt Gammelgaard Nielsen, SH Juul, and John Chemnitz. Learning anatomy with augmented reality. In *EDULEARN18 Proceedings*, pages 1413–1422. IATED, 2018.

[39] GS Rajshekar Reddy and Damien Rompapas. Liquid hands: Evoking emotional states via augmented reality music visualizations. In *Proceedings of the 2021 ACM International Conference on Interactive Media Experiences*, pages 305–310, 2021.

[40] Qiao Jin, Danli Wang, Haoran Yun, and Svetlana Yarosh. Shape of music: Ar-based tangible programming tool for music visualization. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference*, pages 647–651, 2023.

[41] Guido Bozzelli, Antonio Raia, Stefano Ricciardi, Maurizio De Nino, Nicola Barile, Marco Perrella, Marco Tramontano, Alfonsina Pagano, and Augusto Palombini. An integrated vr/ar framework for user-centric interactive experience of cultural heritage: The arkaevision project. *Digital Applications in Archaeology and Cultural Heritage*, 15:e00124, 2019.

[42] Anabela Marto and Alexandrino Gonçalves. Mobile ar: User evaluation in a cultural heritage context. *Applied Sciences*, 9(24):5454, 2019.

[43] Yhonatan Iquiapaza, Jorge Wagner, and Luciana Nedel. Dear: Combining desktop and augmented reality for visual data analysis. In *Proceedings of the 25th Symposium on Virtual and Augmented Reality*, pages 233–237, 2023.

[44] Ricardo Langner, Marc Satkowski, Wolfgang Büschel, and Raimund Dachselt. Marvis: Combining mobile devices and augmented reality for visual data analysis. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2021.

[45] Jinbin Huang, Shuang Liang, Qi Xiong, Yu Gao, Chao Mei, Yi Xu, and Chris Bryan. Sparvis: Combining smartphone and augmented reality for visual data analytics. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 111–117. IEEE, 2022.

[46] Ronell Sicat, Jiabao Li, Junyoung Choi, Maxime Cordeil, Won-Ki Jeong, Benjamin Bach, and Hanspeter Pfister. Dxr: A toolkit for building immersive data visualizations. *IEEE transactions on visualization and computer graphics*, 25(1):715–725, 2018.

[47] Maxime Cordeil, Andrew Cunningham, Benjamin Bach, Christophe Hurter, Bruce H Thomas, Kim Marriott, and Tim Dwyer. Iatk: An immersive analytics toolkit. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 200–209. IEEE, 2019.

[48] Bruno Fanini and Giorgio Gosti. A new generation of collaborative immersive analytics on the web: Open-source services to capture, process and inspect users' sessions in 3d environments. *Future Internet*, 16(5):147, 2024.

[49] Benjamin Lee, Xiaoyun Hu, Maxime Cordeil, Arnaud Prouzeau, Bernhard Jenny, and Tim Dwyer. Shared surfaces and spaces: Collaborative data visualisation in a co-located immersive environment. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1171–1181, 2020.

[50] Daniel Garrido, João Jacob, and Daniel Castro Silva. Performance impact of immersion and collaboration in visual data analysis. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 780–789. IEEE, 2023.

[51] Marco Cavallo, Mishal Dolakia, Matous Havlena, Kenneth Ocheltree, and Mark Podlaseck. Immersive insights: A hybrid analytics system forcollaborative exploratory data analysis. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology*, pages 1–12, 2019.

[52] Oscar Legetth, Johan Rodhe, Stefan Lang, Parashar Dhapola, Mattias Wallergård, and Shamit Soneji. Cellexalvr: A virtual reality platform to visualize and analyze single-cell omics data. *IScience*, 24(11), 2021.

[53] Mark Billinghurst, Maxime Cordeil, Anastasia Bezerianos, and Todd Margolis. Collaborative immersive analytics. *Immersive Analytics*, pages 221–257, 2018.

[54] Huan Li, Hong Fan, and Feiyue Mao. A visualization approach to air pollution data exploration—a case study of air quality index (pm2. 5) in beijing, china. *Atmosphere*, 7(3):35, 2016.

[55] Wei Lu, Tinghua Ai, Xiang Zhang, and Yakun He. An interactive web mapping visualization of urban air quality monitoring data of china. *Atmosphere*, 8(8):148, 2017.

[56] Yu-Ren Zeng, Yue Shan Chang, and You Hao Fang. Data visualization for air quality analysis on bigdata platform. In *2019 international conference on system science and engineering (ICSSE)*, pages 313–317. IEEE, 2019.

[57] Pengyu Chen. Visualization of real-time monitoring datagraphic of urban environmental quality. *Eurasip Journal on Image and Video Processing*, 2019(1):42, 2019.

[58] Gustavo Carro, Olivier Schalm, Werner Jacobs, and Serge Demeyer. Exploring actionable visualizations for environmental data: Air quality assessment of two belgian locations. *Environmental Modelling & Software*, 147:105230, 2022.

[59] Bruno Teles, Pedro Mariano, and Pedro Santana. Game-like 3d visualisation of air quality data. *Multimodal Technologies and Interaction*, 4(3):54, 2020.

[60] Jane Prophet, Yong Ming Kow, and Mark Hurry. Cultivating environmental awareness: Modeling air quality data via augmented reality miniature trees. In *Augmented Cognition: Intelligent Technologies: 12th International Conference, AC 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15-20, 2018, Proceedings, Part I*, pages 406–424. Springer, 2018.

[61] Gowri Sankar Ramachandran, Biayna Bogosian, Kunal Vasudeva, Sushanth Ikshwaku Sriramaraju, Jay Patel, Shubhesh Amidwar, Lavanya Malladi, Rohan Doddaiah Shylaja, Nishant Revur Bharath Kumar, and Bhaskar Krishnamachari. An immersive visualization of micro-climatic data using usc air. In *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*, pages 675–676, 2019.

[62] Noble Saji Mathews, Sridhar Chimalakonda, and Suresh Jain. Air: An augmented reality application for visualizing air pollution. In *2021 IEEE Visualization Conference (VIS)*, pages 146–150. IEEE, 2021.

[63] Clara Larsson, Beata Stahre Wästberg, Daniel Sjölie, Thommy Eriksson, and Håkan Pleijel. Visualizing invisible environmental data in vr: Development and implementation of design concepts for communicating urban air quality in a virtual city model. In *International Conference on Computer-Aided Architectural Design Futures*, pages 253–267. Springer, 2023.

[64] Maxwell Foxman. United we stand: Platforms, tools and innovation with the unity game engine. *Social Media+ Society*, 5(4):2056305119880177, 2019.

[65] EarthAndMoon, 2024. URL `https://github.com/keijiro/MoonAndEarth`. Last Accessed: 2024-May-12.

[66] openAQ, 2024. URL `https://openaq.org/`. Last Accessed: 2024-May-30.

LIST OF FIGURES

List of Tables

ABBREVIATIONS

2D	*Two-Dimensional*
3D	*Three-Dimensional*
API	*Application Programming Interface*
AR	*Augmented Reality*
AQI	*Air Quality Index*
CO	*Carbon Monoxide*
EDA	*Exploratory Data Analysis*
GUI	*Graphical User Interface*
HMD	*Head-Mounted Display*
IA	*Immersive Analytics*
JSON	*JavaScript Object Notation*
LID	*Large Interactive Display*
MR	*Mixed Reality*
NO2	*Nitrogen Dioxide*
O3	*Ozone*
PM	*Particulate Matter*
RPC	*Remote Procedure Call*
RQ	*Research Question*
QoE	*Quality of Experience*
SO2	*Sulfur Dioxide*
UI	*User Interface*
VR	*Virtual Reality*
VRE	*Virtual Reality Environment*
XR	*Extended Reality*

Appendix A. User Study Questionnaire

A Collaborative Application for Big Data Visualization Using Mixed Reality - User Study for Master's Thesis

Thank you for participating in this user study. The purpose of this study is to evaluate a collaborative application designed for big data visualization using mixed reality (MR) technologies. This research is part of a Master's Thesis focused on exploring how MR can enhance the understanding and analysis of large datasets through immersive and interactive visualizations.

During this study, you will be asked to interact with the MR application and provide feedback on your experience. The information and responses gathered will be used solely for research purposes and possible further improvements. Your personal information will remain anonymous, as all data will be observed and analyzed collectively.

Thank you for your time and input:)

* Indicates required question

Demographic Questionnaire

1. Name and Surname *

____

2. Gender *

Mark only one oval.

- ◯ Female
- ◯ Male
- ◯ Non-Binary
- ◯ Transgender Female
- ◯ Transgender Male
- ◯ Prefer Not To Say
- ◯ Other: _______________________________

3. Age *

____

4. How would you rate your general familiarity with technology? *

Mark only one oval.

	1	2	3	4	5	
Not	◯	◯	◯	◯	◯	Very Familiar

5. Have you used Virtual Reality (VR), Augmented Reality (AR) or Mixed Reality (MR) before? *

Mark only one oval.

◯ Yes *Skip to question 6*

◯ No *Skip to question 18*

Prior Experience With Virtual/Augmented/Mixed Reality

6. How frequently do you use VR, AR or MR? *

Mark only one oval.

◯ Only tried it once

◯ Used it a few times

◯ Few times per year

◯ Montly

◯ One or more times per week

7. In which contexts have you used VR, AR, or MR? *

Tick all that apply.

☐ Gaming
☐ Work
☐ School
☐ Social Interaction and Communication
☐ Other: ____________________

8. Select the types of devices you have used. *

Tick all that apply.

- [ ] Mobile phone or tablet
- [ ] Head-Mounted Display - HMD e.g. Meta Quest, HoloLens, HTC Vive, ...
- [ ] Smart Glasses e.g. Google Glasses, ...
- [ ] Head-Up Display - HUD e.g. in a car, in a plane, ...
- [ ] Other: ____________________

9. How would you rate your proficiency with VR/AR/MR technologies? *

Mark only one oval.

	1	2	3	4	5	
Begi	◯	◯	◯	◯	◯	Expert

10. How comfortable are you with navigating and interacting within VR/AR/MR environments? *

Mark only one oval.

	1	2	3	4	5	
Not	◯	◯	◯	◯	◯	Very Comfortable

Prior Experience With Data Visualization

11. How would you rate your experience with traditional data visualization tools (e.g., Excel, Tableau, Power BI)? *

Mark only one oval.

	1	2	3	4	5	
No E	◯	◯	◯	◯	◯	Expert

12. How often do you work with data visualization tools in your professional or academic life? *

Mark only one oval.

	1	2	3	4	5	
Nev	◯	◯	◯	◯	◯	Very Often

13. How familiar are you with analyzing large datasets? *

Mark only one oval.

	1	2	3	4	5	
Not	◯	◯	◯	◯	◯	Very Familiar

14. Which types of data visualizations are you most comfortable with? *

Tick all that apply.

- [ ] Bar Charts
- [ ] Line Graphs
- [ ] Scatter Plots
- [ ] Heat Maps
- [ ] Other: ______________________

Prior Experience With Collaborative Tools

15. How often do you use collaborative tools (e.g. Google Docs, Microsoft Teams, Slack) in your work or studies? *

Mark only one oval.

	1	2	3	4	5	
Nev	◯	◯	◯	◯	◯	Very Often

16. Have you ever used collaborative MR applications? *

Mark only one oval.

- ⬭ Yes
- ⬭ No

17. How often do you collaborate with others when analyzing data? *

Mark only one oval.

	1	2	3	4	5	
Nev	⬭	⬭	⬭	⬭	⬭	Very Often

Post Application Use Questionnaire

Please proceed to the next section only when instructed to do so by the study facilitator.

General Experience

18. How satisfied are you with the overall experience using the MR-based big data visualization application? *

Mark only one oval.

	1	2	3	4	5	
Not	⬭	⬭	⬭	⬭	⬭	Very Satisfied

19. Select the degree of agreement with the statement: *
"*I felt immersed in the MR environment while analyzing the data.*"

Mark only one oval.

	1	2	3	4	5	
Stro	⬭	⬭	⬭	⬭	⬭	Strongly Agree

20. How would you rate the difficulty of **learning to use** the MR-based visualization application? *

Mark only one oval.

	1	2	3	4	5	
Very	⬭	⬭	⬭	⬭	⬭	Very Easy

21. How would you rate the difficulty of **navigating** within the MR environment to explore different data points? *

Mark only one oval.

	1	2	3	4	5	
Very	⬭	⬭	⬭	⬭	⬭	Very Easy

Geospatial Visualization

Geospatial visualization refers to the visual representation of data that has a geographical or spatial component. It involves mapping and displaying data points on a spatial plane, often using maps, globes, or other spatially-relevant contexts to show the relationship between data and geographic locations.
For instance, plotting data points corresponding to various geographical locations directly onto a 3D model of the Earth.

22. Did you find the **geospatial visualization** of data points intuitive? *

Mark only one oval.

⬭ Yes

⬭ No

23. Please elaborate on your previous answer.

____

____

____

____

____

24. How intuitive were the interaction techniques (e.g. gestures, controllers) used to manipulate and analyze the data in MR? *

Mark only one oval.

	1	2	3	4	5	
Not	◯	◯	◯	◯	◯	Very Intuitive

25. Were there any interaction techniques that you found particularly helpful or difficult to use? *

____
____
____
____
____

26. How effective was the MR visualization in helping you understand the structure and patterns within the data? *

Mark only one oval.

	1	2	3	4	5	
Not	◯	◯	◯	◯	◯	Very Effective

27. How would you rate the difficulty of **drawing insights** from the MR visualization compared to traditional 2D visualizations? *

Mark only one oval.

	1	2	3	4	5	
Very	◯	◯	◯	◯	◯	Very Easy

28. How would you rate the difficulty of **collaborating with other users** in the MR environment? *

Mark only one oval.

	1	2	3	4	5	
Very	◯	◯	◯	◯	◯	Very Easy

29. Which features of the MR-based visualization did you find most useful? *

Tick all that apply.

- [ ] Geospatial Visualization
- [ ] Real-Time Interaction
- [ ] Data Manipulation
- [ ] Other: ____________

30. Are there any features you would like to see improved or added? *

____
____
____
____
____

31. Would you prefer using MR-based visualizations over traditional methods for future data analysis tasks? *

Mark only one oval.

◯ Yes

◯ No

32. Please elaborate on your previous answer.

____
____
____
____
____

33. Do you have any suggestions for improving the MR-based big data visualization application? *

34. Do you have any additional comments or suggestions regarding your experience with the MR application? *

A COLLABORATIVE APPLICATION FOR BIG DATA VISUALIZATION USING MIXED REALITY

Abstract

This Master's Thesis presents *VisualizeAir*, a collaborative big data visualization application leveraging Mixed Reality (MR) to provide dynamic and interactive representations of air quality data from around the world. The application employs geospatial visualization, allowing users to explore and manipulate data on a three-dimensional model of the Earth. The collaboration feature allows all users to interact with and analyze data together in real time. The thesis includes an overview of related work in data visualization for immersive experiences, with a particular focus on air quality data visualization. A comprehensive evaluation of the developed application from a Quality of Experience (QoE) perspective assesses usability, immersion, and the efficacy of collaboration in data visualization tasks. Findings indicate favorable user satisfaction with the application's immersive MR environment, addressing the engagement and comprehension of complex datasets, alongside highlighting benefits and challenges. Acknowledging limitations such as a small participant sample, non-standardized questionnaires, and developmental networking issues, the study identifies opportunities for future enhancements. Overall, *VisualizeAir* demonstrates promising potential for advancing data visualization practices in diverse fields with continued refinement.

Keywords: AR, MR, XR, IA, Mixed Reality, Immersive Technologies, Big Data, Data Visualization

KOLABORATIVNA APLIKACIJA ZA VIZUALIZACIJU VELIKIH KOLIČINA PODATAKA POMOĆU TEHNOLOGIJE MIJEŠANE STVARNOSTI

Sažetak

Diplomski rad predstavlja razvoj kolaborativne aplikacije za vizualizaciju velikih skupova podataka, *VisualizeAir*, koja koristi tehnologiju miješane stvarnosti (MR) kako bi omogućila dinamične i interaktivne prikaze podataka o kvaliteti zraka iz cijelog svijeta. Aplikacija koristi geoprostornu vizualizaciju koja korisnicima omogućava istraživanje i manipulaciju podacima na trodimenzionalnom modelu Zemlje. Mogućnost kolaboracije omogućuje svim korisnicima zajedničku i istovremenu interakciju i analizu podataka u stvarnom vremenu. U sklopu rada napravljeno je pregled literature u područjima imerzivne vizualizacije podataka, s fokusom na aplikacije koje prikazuju podatke o kvaliteti zraka. Diplomski rad također uključuje ispitivanje kvalitete korisničkog iskustva. Ispitivana je upotrebljivost, imerzija korisnika te učinkovitost kolaborativnog aspekta prilikom vizualizacije podataka. Rezultati studije ukazuju na zadovoljstvo korisnika imerzivnim MR okruženjem aplikacije, koje potiče uključenost i bolje razumijevanje kompleksnih skupova podataka. Istovremeno su istaknute prednosti i izazovi te prepoznata ograničenja kao što su mali uzorak sudionika, nestandardizirani upitnici i poteškoće u razvoju u mrežnom aspektu. Finalno, pokazano je da postoji potencijal za korištenje imerzivnih tehnologija vizualizacije velikih skupova podataka u različitim područjima, ali uz kontinuirano unaprjeđivanje.

Ključne riječi: AR, MR, XR, IA, Miješana stvarnost, Imerzivne tehnologije, Veliki skupovi podataka, Vizualizacija podataka