Boston College

Morrissey College of Arts and Sciences: Computer Science

Honors Thesis

DataGarden:
Exploring Our Community in a VR Data Visualization

Joy Kondo

May 2023

Advisor: Nam Wook Kim

With special thanks to Justin Park

1 Abstract

With the ever-increasing need to enhance data literacy, methods and approaches to effectively represent and analyze data are being developed. Traditional representations of data such as charts and abstract visuals detract from the humanity represented by datasets.

This thesis presents DataGarden, a system that supports embodied interactions with humane representations of data in an immersive VR environment for users to think about the people behind the data. In this interactive system, users are able to interact with their personal visualizations in our virtual garden, as well as with the visualizations created using the data of the Boston College community.

2 Acknowledgements

I would like to express my gratitude to my supervisor, Nam Wook Kim, who has guided my research and supported me over the past year and a half. I also want to thank Justin Park, who has worked with me for the past year to implement the DataGarden system and without whom the amazing VR environment would not have been made. Thank you also to Will Morrison who helped me get the system up and running when I started my research. I also want to thank Josiah Kondo, my brother, who supported my research, implemented many of my ideas, and helped conduct user studies.

I am thankful for my parents who always provided for me and gave me the best so that I could come to college and have the opportunity to do this research in the first place. I also thank my church family who have supported me throughout my college years. I would not have been able to finish college without your love and prayers. Lastly, I thank God who has faithfully kept me with His steadfast love that endures forever. I owe my life to Him and I would not be here without Him.

Contents

3 Introduction

3.1 Background

People have traditionally used 2D visualizations to represent data. A focus on minimalistic, objective, and accurate visualizations has led to the use of abstract representations of data, such as bars, lines, and pies. These methods of visualizing data have proven perceptually effective for representing statistical data. In our increasingly data-driven society, such visualizations are used across disciplines to communicate and analyze complex, multi-faceted data. As a society we have learned to utilize visual perception and analytical skills to effectively garner salient information and draw relevant insights into these abstract forms of data visualization. Various data visualization programs such as Tableau, PowrBI, and R, have enabled analysts to explore data efficiently and discover meaningful insights.

3.2 Problem

However, traditional methods of aggregate, abstract representations prove to be ineffective in trying to represent non-abstract data, because they undervalue the humanity behind the data [6]. Design principles such as avoidance of visual clutter have been infused into abstract visualizations, consequently reducing the humanity represented by data visualizations to mere bars and dots[5]. The personal stories, emotions, and meaning in the datasets that ought to be interwoven into their respective data visualizations are instead perceived by viewers as numbers, statistics, and patterns. The data that is drawn from people's lived experiences and that captures a snapshot of our world ought to foster an emotional connection with those people represented by the data. Though 3D visualizations have typically been used in scientific domains where the datasets have physical properties that can be translated into 3D models, we propose that 3D visualizations could be used powerfully beyond this scope to humanize data and effectively cause viewers to consider the people behind the data.

3.3 Research Question & Methodology

In this study, we are investigating how to leverage immersiveness and humane representations of data to help people think about the people behind the data. We utilize immersiveness to enable embodied interactions with data that could set people in a frame of mind similar to real-world experiences [9] in which they can relate the 3D visualizations to real-world objects and their own life experiences. Virtual reality (VR) technology can mimic our natural 3D viewing environment to advance immersive, interactive data analysis and support better understanding of visual information[3]. Effective visualization of more dimensions beyond 2D can support recognition of interesting patterns, correlations, or outliers in data. Figurative, humanistic representations of data could heighten people's interest by helping them relate to the data. Studies have found that rather than simplifying visualizations and eliminating all visual clutter, aesthetic and enjoyable visualizations could increase user interactivity and foster engagement with the data visualizations[7]. Unit visualizations that use visual groupings made of figurative individual marks could prompt affective responses and encourage storytelling through anthropomorphism to convey the stories behind the data. We want to evoke visceral feelings, or a "subjective sensation of being there in a scene depicted by a medium, usually virtual in nature" in order to convey the semantics of the data and promote a better understanding of our community through the 3D visualizations. The VR environment can offer an innovative data experience to make data visualizations more enjoyable and encourage deeper engagement with the data[11], as well as span multiple perspectives to give the viewer different experiences as they are immersed in the environment from multiple viewpoints[8].

Taking inspiration from DataSelfie, we created a questionnaire to gather data for our virtual garden. We conducted interviews with seven faculty members from six different departments at Boston College to optimize the structure of the questionnaire in order to obtain insights about the community. The feedback we received from the interviews was diverse and helpful, with each interviewee providing valuable ideas. For example, an Environmental Science professor recommended including a question about people's outlook on life or their optimism for the future. Plastic usage was also suggested as a generic yet indicative data point to include. A Sociology professor recommended questions about social media usage and encouraged analyzing social behaviors across different groupings. This led us to consider comparing data across various categories. A Psychology professor suggested examining people's stress levels, which could be a good indicator of their mental state. Finally, a Philosophy professor proposed analyzing people's relationships and how connected or disconnected they felt from their community. With these suggestions in mind, we created a question bank containing over 200 questions, from which we could create a well-structured questionnaire for our virtual garden that would capture diverse data points.

4 Related Work

4.1 Visualization Aesthetics

While traditional information visualizations have been created with basic design goals of relieving distracting elements or extraneous information, studies have found that there is a tradeoff between efficient designs and comprehension/recall on the part of users[5]. Cognitive efficiency could in fact limit the depth of understanding and fail to capture the complex nature of datasets that are not purely abstract and quantitative.

4.1.1 Humane Representations

Visual metaphors are often used to represent data in a way that is recognizable by users. For example, UbiFit Garden tracks users' physical activities with the metaphor of blooming plants in a garden. Such intuitive metaphors allow users to easily understand what the visualization is trying to represent[8][2][7]. Similarly, figurative representations of data as determined by users of DataSelfie enabled users to reflect more deeply on themselves, encouraging perhaps slower but deeper engagements with data to humanize users' data experience. Another study shows that figurative symbols prove to increase memorability and engagement with data.

4.1.2 Enactive Interfaces

A study on Anthropographics, using design strategies that help an audience relate to the people represented by the data and have an emotional connection, shows that the distinctiveness of each visualized data point is effective in fostering an emotional connection and generating a response from an audience, rather than aggregating data points[8]. Another study shows that unit visualizations where individual marks correspond to individual data points foster an emotional connection with the people behind the data[6]. These ideas tie in with enactive interfaces, which are interfaces specifically designed to augment a user's sense-making of the data by means of creating new modes of perceptual interactions through their own actions. The emotional connection drawn out from anthropomorphic representations of data can make up an enactive interface that allows users to make better sense of humane data[10].

In this project, our goal is to foster an emotional connection between users and the data. We utilize humane representations of data rather than abstract visualizations to engage users and encourage a deeper understanding of the data. In addition, each individuals' data is represented by a single visualization so that users can have an emotional connection with the data that represents individual people. The aesthetics and design of our system combines humane representations of data with an enactive interface to achieve integration between the user and the environment. The enactive interactions promoted by our interface, together with the humane representations of data, serve to support deeper understanding of the data and engagements with the data.

4.2 Immersive visualization

Immersive environments have been used in research varying from a Wall Street Journal 3D VR tour of the Nasdaq to bring financial reporting to life, to Dear Pictograph, a project that investigated enjoyment of visualizations using immersion. Dear Pictograph utilizes an immersive environment to give users the ability to relate their immersive virtual experiences with real-world objects and their own life experiences[1][11]. The large workspace that VR affords allows users to analyze large amounts of data[9]. Additional research shows that immersiveness allows users to have an intuitive understanding of the environment and the visualizations that it contains[3]. The multi-sensory feedback and output that immersive experiences provide also allow people to utilize their most familiar and appropriate human senses and instincts in terms of movement. [4] In addition, the situational awareness that is provided by immersive VR environments supports a better understanding of visual information[4][11][9]. Although immersive environments are arguably not the best means of attaining quantitative accuracy in data visualizations, users have more dramatic sensual perceptions of the scale of the data.[9]

Dear Pictograph also shows that engaging with data visualizations in immersive environments increases the enjoyability of the visualizations. Such enjoyable data visualizations can in turn draw a wider audience to view and interact with the data, as well as take more time to understand the data[1].

Our research builds on the benefits of immersiveness that enable users to have an intuitive data experience and to increase their enjoyability. We utilize the perceptual effectiveness of immersion to set participants in a frame of mind that allows them to use their intuitive senses to understand the data. We further explore how the ability for people to relate their real-world experiences with immersive environments can further draw out the humanity behind the datasets and allow them to relate to the people represented by the visualizations to encourage people to reflect deeply on themselves.

4.3 Community Data

Giorgia Lupi's Dear Data project introduced data humanism to facilitate deeper engagement with data through personalization of data[11]. Subsequent works like Dear Pictograph have expanded her work to explore making data visualization experiences more enjoyable through personalization and immersion, and DataSelfie equips people to design their own visual vocabulary for qualitative and nuanced aspects of personal data[7]. These works show that personal data is effective in captivating users by helping them relate to the data.

While previous works have incorporated personal data to engage users and foster deeper interactions with and insights into the data, we add community data into the picture to enable users to interact with their personal visualizations in the context of a larger community. When each member in a community shares their personal data, we can collect data on the community as a whole, and expand research from an individual personalized data visualization to a unique representation of a community made up of individuals. We explore how the ability to analyze personal data as well as community data can promote more engagement, deeper insights, and develop a connection between the user and the data represented by the people in their community.

5 Design Goals

We wanted to build a system through which individuals can engage with their personal data and community data through an aggregation of personalized visualizations from members in the community. Our main goal through our design is to enable people to see the humanity within the data while having deeper engagements with the visualizations.

5.1 Collecting community data with a questionnaire

In order to engage audiences more and humanize our data, our surveys are designed to collect personal data that can only be answered by each individual. The collection of personal data amounts to community data that is represented in our system. The survey questions cover a range of categories, from categorical data such as class and average daily social media usage, to personal data including one's personality and outlook on life. The collective representations that visualize our community support data humanism, since information about individuals is reflected in the data that is represented.

5.2 Creating humane representations of data

To foster memorability of the data and garner an emotional response from the data, we intend to create humane data visualizations where each person's data is visualized by an individual flower or tree. Humane representations help audiences relate to the people represented by the data and effectively support deeper engagement with the visualizations.

5.3 Utilizing immersive embodied interactions in a VR environment

To foster deeper engagement with the data visualizations, we also utilize immersiveness in a VR environment. The immersive quality of the visualization fosters memorability and deeper interactive analysis of the data. The added dimension of the visualizations also helps to support better understanding of the visual information, because people can view the visualizations intuitively, in a manner similar to their interactions with the natural world. Embodied interactions allow users to achieve deeper understanding through the knowledge they acquire from their own actions and movements.

6 System Description

DataGarden consists of five main components: a questionnaire that collects data, a spreadsheet that maps the data to a personalized visualization, a user interface in the form of a virtual environment, a legend to support analysis, and interactions within the environment for users to engage with the data visualizations.

6.1 Questionnaire and Spreadsheet: Collecting Community Data

To engage users by visualizing their personal data, data points in the visualization are initially collected by means of a questionnaire (Figure 1A). As a user responds to a series of personal questions in the questionnaire, the answers are collected in a spreadsheet that maps the data to an element in the virtual environment (Figure 1B). The collection of personal data amounts to community data, with each element in the virtual environment representing something about the community.

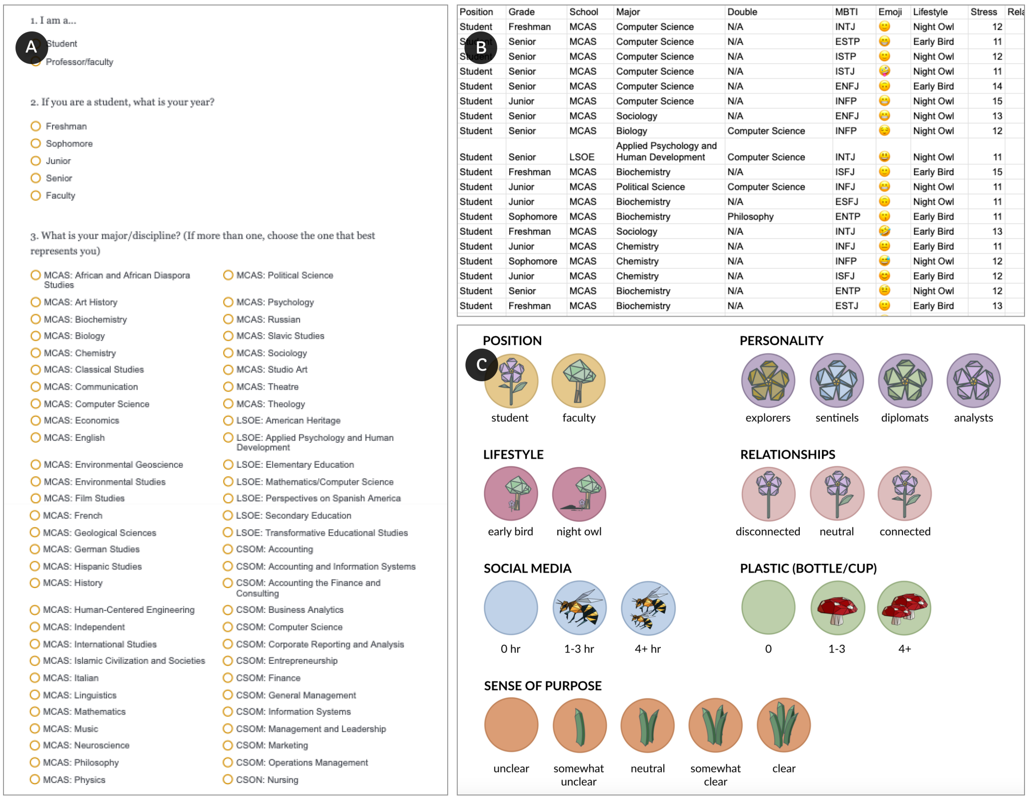

Figure 1: DataGarden's iterative process begins with a survey containing both quantitative and qualitative questions. The first three questions are shown (A). Responses to the survey are collected in a spreadsheet (B). Each answer to the questions in the survey is mapped to a variable in DataGarden, as shown in this legend (C). The legend is also available in the virtual environment for participants to reference during the interactive experience.

The rest of this section (6.2) pertains to the variable mappings in the first iteration of our system. Initially, users were asked to choose between two options: student or professor/faculty. The choice they make determined whether their data would be visualized as a flower or a tree, respectively. An individual's Myers-Briggs personality type determined the color of the corresponding flower in the virtual garden. The colors represented the different personality types present in the community. The 16 personality types were categorized into four categories, Diplomats, Sentinels, Explorers, and Analysts. These categories, taken from the Myers-Briggs personality test website, correlate to the following:

Diplomats: INFJ, INFP, ENFJ, ENFP
Sentinels: ISTJ, ISFJ, ESTJ, ESFJ
Explorers: ISTP, ISFP, ESTP, ESFP
Analysts: INTJ, INTP, ENTJ, ENTP

Each category was assigned a random color, with Diplomats represented by red, Sentinels by blue, Explorers by pink, and Analysts by purple.

The garden was split into four quadrants according to the four schools at Boston College: Morrissey College of Arts and Sciences (MCAS), Carroll School of Management (CSOM), Lynch School of Education (LSOE), and Connell School of Nursing (CSON). Users' responses to questions about feeling connected or disconnected in their community determined the number of leaves on flowers or branches on trees. The number of clouds corresponded to users' average daily plastic bottle usage (Figure 2C).

Most of the variables were not able to be included in this iteration and were set aside for the time being. The questions we were not able to include in the virtual environment were questions about users' major/department, the emoji that best represented how users felt most of the time, their lifestyles (early bird/night owl), average stress levels, average social media usage, how much they felt their life had a sense of purpose, and their outlook on life where they indicated whether they thought the future was bright or grim.

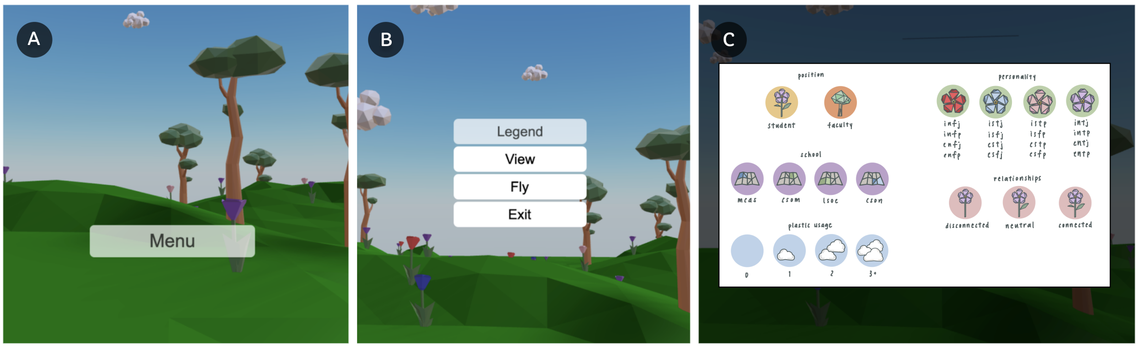

Figure 2: Initial design of DataGarden. A menu button is always visible in DataGarden (A). Upon clicking the menu button, users are presented with 4 choices: Legend, View, Fly, Exit (B). Users can click the legend button to view the legend that shows the mappings from the data to the visualizations in the virtual environment (C).

6.2 Environment Components: Humane Data Visualizations

Users' unique responses generate a personalized data visualization. Individuals are able to see how their responses correlate to a characteristic of their own visualization, and are also able to compare their own visualization to the other visualizations in the virtual environment. The uniqueness of each visualization enables users to reflect on the individuality of each person behind the data, and the surrounding visualizations work together to provide unique information about the community, such as the most common personality type or the average stress levels of people in the Boston College community.

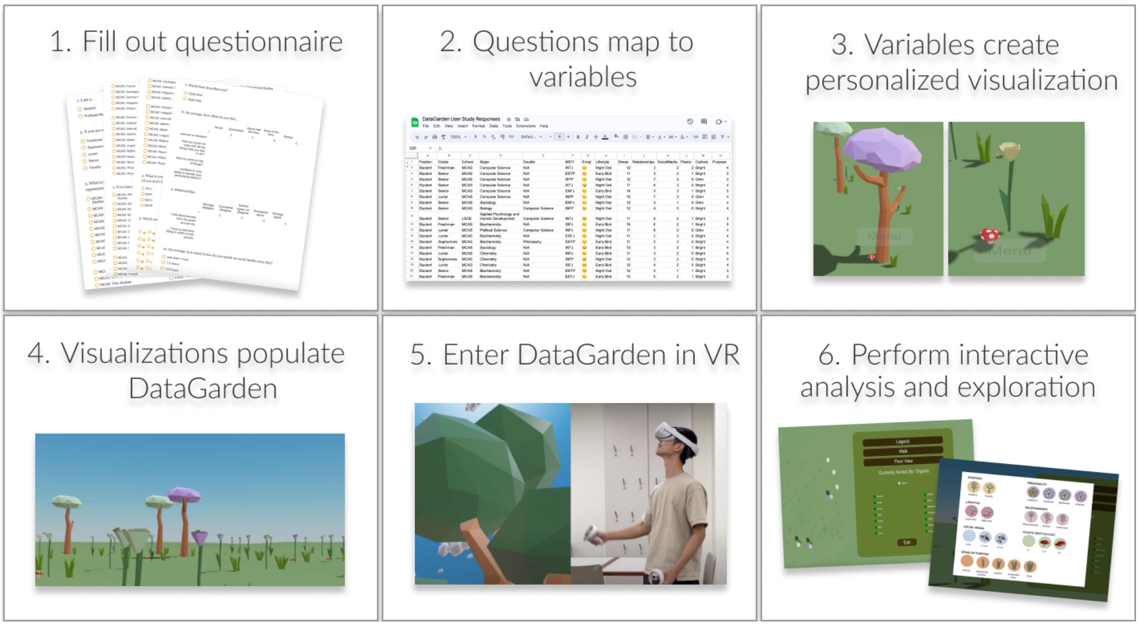

Figure 3: Schematic diagram of the DataGarden pipeline.

6.3 Immersive and Interactive VR Environment

Our system enables users to interact with the data visualizations in an immersive and interactive VR environment, and is implemented based on the schematic diagram above (Figure 3). A legend that maps individuals' survey responses to the data visualizations aids users in their analysis of the data and fosters a better understanding of the visual representations (Figure 2C). Users have the ability to navigate through the virtual environment by means of a number of controls listed in the given menu (Figure 2B). They can change their perspective from a ground level view (Figure 4E) to an aerial birds'-eye-view (Figure 4D) and to a top-down view (Figure 4C), giving them the ability to see individual visualizations as well as all the community visualizations as a whole. Users can use the "walk" mode to navigate through the environment on ground level, or use the "fly" mode and have the freedom to move around in the environment both horizontally and vertically without the pull of gravity. These interactions support engagement with the visualizations and an understanding of the humanity behind each element in the environment.

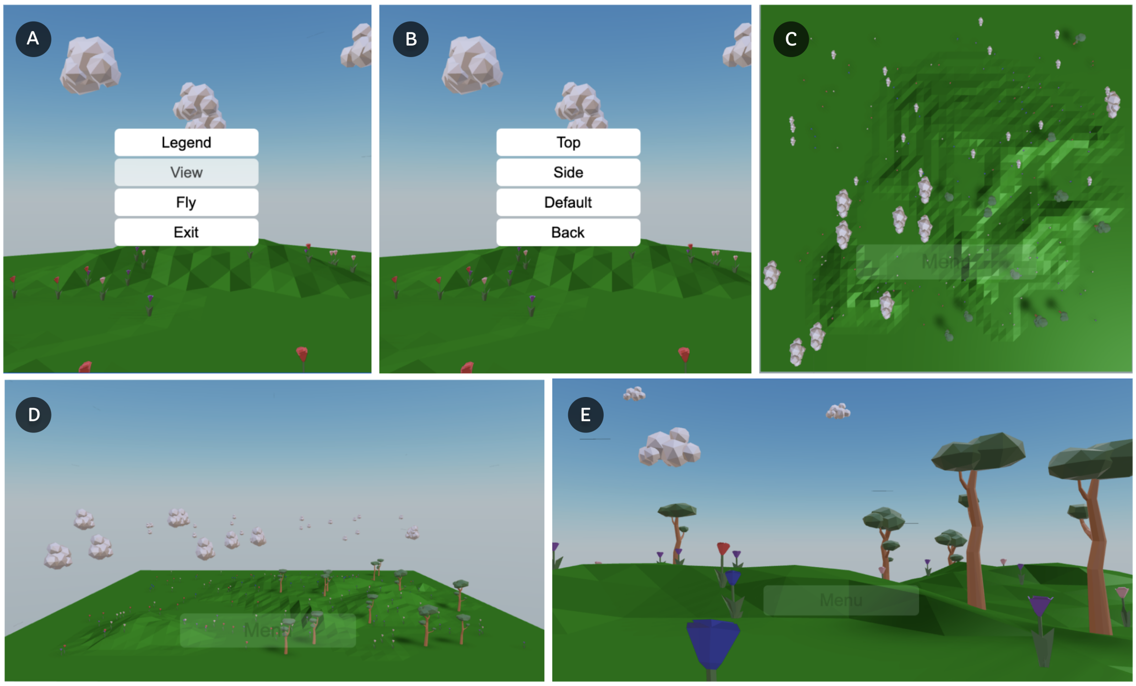

Figure 4: Initial design of DataGarden. Users can select "View" from the menu to change the angle at which they view the environment (A). DataGarden supports 3 different perspectives: Top, Side, and Default (B). If the top button is selected, the user's perspective is automatically adjusted to be directly above the garden. Users must physically tilt their heads down to view the garden from a top-down viewpoint (C). If the side button is selected, the user's perspective is shifted to a bird's-eye view slightly off to the side of the garden, such that the user is able to view the entire garden from an aerial perspective (D). If the default button is selected, the user's position is shifted down to ground level (E).

7 User Study

7.1 Data Collection

In order to populate our virtual garden with real data so that the garden could be used for a user study, we created a data collection survey (Figure 1A). The data collection survey was sent to departmental mailing lists in order to reach as many people in the BC community as possible. The data collection survey as described in the system description included questions about class, major, personality, habits, lifestyle, and outlook on life. The answers to these questions created a personalized visualization for each participant, and the visualizations were added into our virtual environment.

7.2 Usability Study

We conducted a usability study to evaluate the effectiveness of DataGarden in fostering deeper engagement with data and in helping people see the humanity behind the data. The purpose of this initial study was to determine the usability of the system and to receive initial feedback from users who had no experience with our system.

7.3 Participants

A select group of 5 participants who completed the data collection survey were included in the usability study. The participants were carefully selected among volunteers, with the goal of diversifying the group in terms of gender, race, class, and major as much as possible. We compensated these participants with $25 Amazon gift cards for a 45-60 minute session.

7.4 Procedures & Tasks

Each 45-60 minute session began with a pre-study background survey, which asked questions regarding race and gender, as well as the participant's previous experience with data visualization and with virtual reality devices. A research moderator then performed an Oculus demonstration in order to introduce participants to basic functions of the Oculus device they would be using for the study. Then participants were given some time to try using the device to become acquainted with the basic controls before completing 5 tasks.

In the first task, we asked participants to count the number of red flowers in order to see how well they could identify elements. In the second task, they were asked to compare two similar elements and explain the differences. In the third task, we tested how well users could analyze data by asking them an indirect question that would require analysis rather than simple identification. In the fourth task, we asked specific questions about elements in the environment to see how well participants could identify relationships between different variables. In the fifth task, participants were asked about an overall pattern in the environment to see if they could summarize trends in the data.

Following the completion of the tasks, participants were asked to complete a post-study usability survey to assess the usability of the system. The survey was followed by a semi-structured interview during which participants could share about their experiences to provide the researchers with feedback.

7.5 Results

We utilized the System Usability Score to determine the usability and intuitiveness of our system. We noticed that users' responses and familiarity with the system largely depended on their previous experience playing video games or using VR devices. We had a rather small participant pool of 5 people and each response varied greatly from the others. Our system usability score came out to be 66.5, right below the average usability score of 68. All participants were able to complete all the tasks, though some took significantly longer than others. The tasks and average completion times for each task were as follows:

1. Which color is the most dominant? (21 sec)
2. Compare any 2 student visualizations and explain how they differ (163 sec)
3. What is your impression of how much plastic is used on average among CSOM students? (77 sec)
4. Choose one flower. What is this individual's personality type and school at Boston College? (57 sec)
5. Which personality type is dominant among MCAS students? (46 sec)

Personalization of visualizations. Participants all expressed positive feelings about the personalization of visualizations. "The personalization made the data more human, especially because the individual visualizations had human characteristics" -P1 and "[Personalization of visualizations] made it more meaningful, since there was less generalization of the data points" -P3. Other participants noted that the individual personalized visualizations made it easier for them to understand a large amount of data. They were impressed by the amount of information that could be visualized at once while the data still remained personal. However, one participant also noted that it was a lot of information to take in and keep track of, and suggested making the legend easier to reference.

Enjoyability & engagement of immersion. All the participants made positive comments about their experience, describing it as "enjoyable" and said that they were engaged during the entire experience. Multiple participants noted that the colors especially made the experience enjoyable, and one participant mentioned that the colors helped to make the patterns easier to notice. One participant was particularly engaged by the immersion factor and said, "it's the feeling of being in a game" -P2. P2, P3, and P5 pointed out that they enjoyed flying, and the physical movements within the environment made the experience fun. They also commented on the aesthetics, "The choice of visualizations is aesthetically pleasing" -P4. P3 mentioned that the interactions that were available made the experience intriguing, and expressed an interest in further developments in terms of interactions within the virtual environment.

Learnability & usability of immersion. The learnability of the system varied depending on the participant. Those with prior experience in gaming or using a VR device were able to figure out the functionalities a lot quicker than those with no prior gaming/VR experience. "It took some time to get used to- especially with the pointer and clicking the menu buttons. Once I got the hang of it it was ok but maybe it could be changed so that the clicking is more intuitive" -P1. However, one participant had a different opinion. "VR works the way it does because you can move around and touch things in a way that's not necessarily intuitive. One perk of VR is moving around physically but it's unfamiliar so it's hard for people... but I don't think that's a bad thing" -P3. The participants who found the controls to be hard to use had a difficult user experience with the system. They attributed the difficulty to the unfamiliarity with the virtual reality environment. Improving the user experience by making the controls more intuitive and user-friendly could help mitigate these issues.

Information overload & navigation. Some participants felt overwhelmed by the amount of information presented in the DataGarden system. "I feel like there's a trade-off between the amount of data that's included here. On one hand it's really cool to be able to contain so much information about individuals in a way that other charts and visualizations wouldn't be able to. But at the same time it's a lot to take in" -P5. They found it difficult to keep track of all the data presented and found the legend hard to access, requiring multiple taps. Accessing the menu was also considered unintuitive, which made navigating the system more challenging. "I wish the menu and legend didn't require as many steps because I found that I needed to access it a lot. It would help if each menu element could be one click away" -P4. Additionally, some participants did not understand the purpose of the top-down view and found it to be unnecessary. We concluded that improving the navigation of the system and making the information more accessible could help alleviate these issues.

Use cases. The participants in the user study of the DataGarden system highlighted several interesting use cases for the technology. They noted that the system could be used for social impact by providing information about populations and humanizing data to encourage empathy and understanding. Additionally, the system could be useful in a company management setting by providing a novel way of visualizing data that encourages viewing data from different angles. The ability to zoom into the data and see individual answers, as well as patterns and overviews, could facilitate decision-making and problem-solving. Overall, the participants found the DataGarden system to be an innovative and engaging approach to data visualization that has the potential to be applied in a variety of contexts.

8 Discussion

8.1 Suggestions

Participants suggested having a legend that would be always visible, but smaller in size and located on the side of the platform. Additionally, the legend could be viewed with a single click for better accessibility. To make the legend more user-friendly, another suggestion was to replace the labels of the personality types with category names that summarized those personalities. For example rather than labeling the yellow flower under "Personality" as "ISTP, ISFP, ESTP, ESFP," the label could read, "Explorers," so that users could have an idea of the type of people who were represented by each color.

Another recommendation was to make the menu easier to access, since navigating the platform could be difficult for users who were unfamiliar with the VR technology. Although participants had varied preferences based on their prior experiences with VR technology and video games, it was observed that the VR controls were not intuitive. However, as one participant noted, one of the advantages of using VR is the ability to physically move around and interact with virtual objects, which may not come naturally to new users.

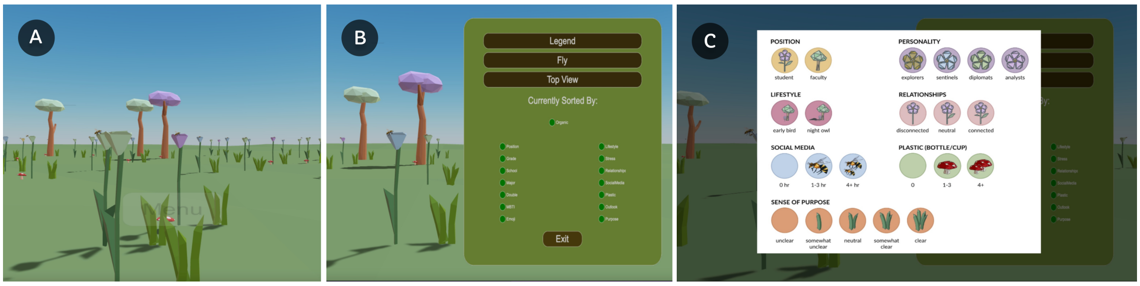

Figure 5: Updated design of DataGarden. A menu button is always visible in DataGarden (A). Upon clicking the menu button, users are presented with the choices: Legend, Fly, Top View, Exit, as well as radio buttons for sorting (B). Users can click the legend button to view the legend that shows the mappings from the data to the visualizations in the virtual environment (C). The legend shows information for variables that are not included in the tooltip.

A third suggestion was to incorporate more interactions into the platform. Users could pick up objects so that the hand component of VR technology could be better utilized. Additionally, the legend could be designed to be held as a map to make it more interactive.

Finally, participants suggested that the design of the platform could be improved by showing rain going down from clouds to individual flowers to make the connection more clear. Furthermore, it was observed that certain flowers on the platform were overlapping, which could be addressed to enhance the overall aesthetic appeal.

Overall, the suggestions aimed to improve the usability and interactivity of DataGarden and to make it more engaging and user-friendly for both new and experienced users.

8.2 Upgraded System

After conducting our initial user study and gathering feedback from participants, we took the valuable insights we gained and utilized them to improve the DataGarden system. Based on the results and feedback from the study, we focused on enhancing the existing features and functionalities to better meet the needs and preferences of our users. Through this process, we were able to refine the user interface and make it more intuitive and user-friendly, while also improving the visualizations and data representation. Our aim was to create a more engaging and interactive experience for users, while also providing them with more robust and insightful data analysis capabilities. As a result, we were able to create a more comprehensive and dynamic system that provides users with a more personalized and meaningful data experience.

Figure 6: Updated design of DataGarden. Users can select "Fly" from the menu to change the angle at which they view the environment in an experience similar to flying (A). Users can also switch back to "Walk" to return to a ground-level view (B).

Variables The upgraded version of DataGarden now includes most of the variables and direct mappings (Figure 5C). The choice between one's position as a student or professor/faculty still determines whether the data will be represented as a flower or a tree, respectively. The color of an individual flower or tree is determined by the person's Myers-Briggs personality type and represents the different personality types present in the community. We were able to match the four colors we use for the flowers and trees in the garden with the colors specified by the Myers-Briggs' website to represent each personality type group. Green represents Diplomats, blue represents Sentinels, yellow represents Explorers, and purple represents Analysts.

Users' responses to lifestyle questions map to the shadow on the flower or tree, with early birds having no shadow and night owls having a shadow. Users' responses to questions about feeling connected or disconnected in their community determine the number of leaves on flowers or branches on trees. The number of bees and mushrooms around individual flowers and trees correspond to users' average social media usage and average plastic bottle usage, respectively. The quantity of shrubs around flowers and trees corresponds to users' answers to the question of how much they feel their life has a sense of purpose. Users' major/discipline and the emoji that best represents how they feel most of the time are included in the tooltip as additional information (Figure 8). Variables that have yet to be incorporated into the system are users' average stress levels and their outlook on life, where they indicate whether they think the future is bright or grim.

Interactions In the upgraded system, we made a significant change that allows users to change the groupings of the flowers and trees in the garden (Figure 7). The new feature enables users to analyze patterns and trends among different groups of people at Boston College, by sorting the visualizations according to similar answers to questions such as "relationships," "social media," and "emoji." Users can now engage with the data more and explore trends in a way that was not possible before. This change was made with the goal of providing a more embodied and interactive experience for the users.

Another new feature added to the system was the tooltip (Figure 8). When a user hovers the VR control over a flower or tree, the tooltip displays information about the user's major/discipline and the emoji that best represents how they feel most of the time. While we discussed the possibility of removing the legend and including more information in the tooltip, we kept the legend in an updated form to help users understand the data better. These enhancements were made to bring us closer to our goal of humanizing data and allowing users to see the data as representative of individual people rather than mere abstract representations.

Figure 7: One of DataGarden's interactive capabilities includes moving elements around in the garden to group them differently. This functionality allows for interactive exploration and analysis of the data. The natural or "organic" view is shown, when the user enters the garden and sees an overview of all the visualizations present (A). Users can view a menu with different grouping options, such as by "position," which splits the visualizations into the groups "student" and "faculty," represented by flowers and trees, respectively (B). Upon selecting a different variable that has more than two groupings, users can see the visualizations transition to reflect the new grouping (C).

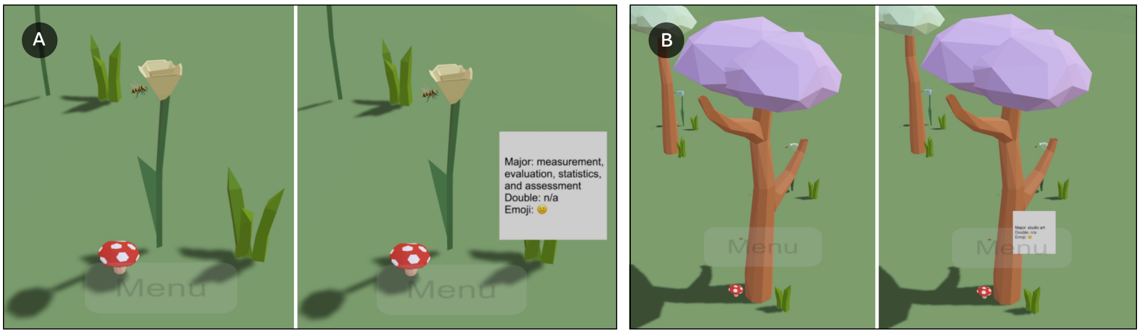

Figure 8: Another one of DataGarden's interactive capabilities includes a tooltip for each visualization that appears when a user hovers their VR controller over an individual flower or tree. The resting state shows no tooltip as users move around (left) but when the user uses the controllers to hover over the flower as pictured (right), the tooltip appears containing relevant information (A). The same happens when a user hovers their controller over an individual tree (B).

8.3 Future Work

In the future, we plan to expand our research and improve the DataGarden system by adding new features and conducting comparative studies. One of the proposed additions is to incorporate sound effects in the virtual garden to make it more immersive and realistic. We also intend to introduce more interactions, such as the ability to pick up objects in the environment and interact more with the various elements.

Additionally, we plan to create a 2D equivalent of the virtual garden so that we can conduct a comparative study between the 2D and 3D visualization methods. These future developments will help us to evaluate the effectiveness of our system in engaging users and fostering emotional connections with data, as well as to identify areas for further improvement. By continuously refining our approach, we aim to create a powerful and innovative tool for representing and humanizing data in a meaningful way.

9 Conclusion

In conclusion, we have developed DataGarden, a unique and innovative data visualization tool that humanizes data by using natural elements such as flowers and trees to represent individual data points. Through the use of our system, users can explore patterns and trends among different groupings of people at Boston College based on various variables.

Overall, DataGarden has the potential to be an effective tool for data exploration, analysis, and communication. It allows users to engage with data in a more intuitive and personalized way, making data more accessible and meaningful, and, most importantly, allowing people to think about the people behind the data. We believe that our tool could have a wide range of applications, from academic research to market analysis, and we look forward to seeing how it will be used and further developed in the future.

References

[1] Aleszu Bajak. Why the wsj used a virtual roller coaster ride to bring stock prices to life, Apr 2015.

[2] Sunny Consolvo, David W. McDonald, Tammy Toscos, Mike Y. Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, Louis LeGrand, Ryan Libby, Ian Smith, and James A. Landay. Activity sensing in the wild: A field trial of ubifit garden. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, page 1797–1806, New York, NY, USA, 2008. Association for Computing Machinery.

[3] Ciro Donalek, S. G. Djorgovski, Alex Cioc, Anwell Wang, Jerry Zhang, Elizabeth Lawler, Stacy Yeh, Ashish Mahabal, Matthew Graham, Andrew Drake, Scott Davidoff, Jeffrey S. Norris, and Giuseppe Longo. Immersive and collaborative data visualization using virtual reality platforms. In *2014 IEEE International Conference on Big Data (Big Data)*, pages 609–614, 2014.

[4] Barrett Ens, Benjamin Bach, Maxime Cordeil, Ulrich Engelke, Marcos Serrano, Wesley Willett, Arnaud Prouzeau, Christoph Anthes, Wolfgang Büschel, Cody Dunne, Tim Dwyer, Jens Grubert, Jason H. Haga, Nurit Kirshenbaum, Dylan Kobayashi, Tica Lin, Monsurat Olaosebikan, Fabian Pointecker, David Saffo, Nazmus Saquib, Dieter Schmalstieg, Danielle Albers Szafir, Matt Whitlock, and Yalong Yang. Grand challenges in immersive analytics. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery.

[5] Jessica Hullman, Eytan Adar, and Priti Shah. Benefitting infovis with visual difficulties. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2213–2222, 2011.

[6] Alexander Ivanov, Kurtis Thorvald Danyluk, and Wesley Willett. Exploration & anthropomorphism in immersive unit visualizations. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, page 1–6, New York, NY, USA, 2018. Association for Computing Machinery.

[7] Nam Wook Kim, Hyejin Im, Nathalie Henry Riche, Alicia Wang, Krzysztof Gajos, and Hanspeter Pfister. Dataselfie: Empowering people to design personalized visuals to represent their data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–12, New York, NY, USA, 2019. Association for Computing Machinery.

[8] Luiz Morais, Yvonne Jansen, Nazareno Andrade, and Pierre Dragicevic. Showing data about people: A design space of anthropographics. *IEEE Transactions on Visualization and Computer Graphics*, 28(3):1661–1679, 2022.

[9] Dheva Raja, Doug Bowman, John Lucas, Chris North, and Virginia Tech. Exploring the benefits of immersion in abstract information visualization. *8th Int'l Immersive Projection Technology Workshop (IPT '04)*, 01 2004.

[10] Andrés Rodríguez, Pascual González López, and Gustavo Rossi. Sketching for designing enactive interactions. In *Proceedings of the XV International Conference on Human Computer Interaction*, Interacción '14, New York, NY, USA, 2014. Association for Computing Machinery.

[11] Hugo Romat, Nathalie Henry Riche, Christophe Hurter, Steven Drucker, Fereshteh Amini, and Ken Hinckley. Dear pictograph: Investigating the role of personalization and immersion for consuming and enjoying visualizations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–13, New York, NY, USA, 2020. Association for Computing Machinery.