A Real Time Three Dimensional Representation of the Social Network 'Reddit' for use in a Virtual Reality Environment.

Adam Edward D'Arcy
12523233

Final Year Project – 2017
B.Sc. Computer Science and Software Engineering

Department of Computer Science
Maynooth University
Maynooth, Co. Kildare
Ireland

A thesis submitted in partial fulfilment of the requirements for the B.Sc. in Computer Science and Software Engineering

Supervisor: Dr. Kevin Casey

Table of Contents

Timestamp
2017/03/15 1:25:45 PM GMT
2017/03/15 1:29:38 PM GMT
2017/03/15 1:39:09 PM GMT
2017/03/15 1:41:19 PM GMT
2017/03/15 3:06:12 PM GMT
2017/03/15 7:15:56 PM GMT

Have you ever experienced Virtual Reality
Yes
Yes
Yes
No
Yes
Yes

If so, how would you rate the experience?
7
10
8
9
8

At the moment, do you believe good VR experiences are readily accessible?
No
Yes but the price needs to decrease
Yes, but through mobile
No, too expensive
Yes but the price needs to decrease
No, too expensive

Do you prefer data presented in an infographic or a plain table, please give a short reason why?
Yes, easy to read
Infographic
infographical as I feel it displayes the information clearly
Infographic is much easier to follow.
Infographic - Easier to look at and understand
Infographic - it's easier to understand the presented info

How would you rate the ease of use of the application?
9
9
8
9
7
8

Declaration

I hereby certify that this material, which I now submit for assessment on the program of study as part of B.Sc. Honours Computer Science and Software Engineering qualification, is *entirely* my own work and has not been taken from the work of others - save and to the extent that such work has been cited and acknowledged within the text of my work.

I hereby acknowledge and accept that this thesis may be distributed to future final year students, as an example of the standard expected of final year projects.

Signed: Date: 3rd/March/2017

Acknowledgements

It is with great pleasure that I thank, everyone from the Final Year Lab, my friends, for their help and support throughout the past year, my family and girlfriend for providing constant feedback and support throughout the development of this project, the faculty of the Department of Computer Science, Maynooth University, for sharing their collective knowledge and expertise over the past 4 years, and finally, a special thanks to my supervisor Dr. Kevin Casey, without whom this project would not have been possible.

Abstract

The social network Reddit is an ever expanding universe, containing Subreddit's for every possible interest. The challenge of this project was to develop a real-time, three-dimensional Virtual Reality ready representation of Reddit. The results of which are discussed throughout the following thesis.

The above challenge can be better described as two problems;

1. Developing a method by which real-time Reddit data can be represented in a three-dimensional environment.
2. Using this method to develop a Virtual Reality ready representation of Reddit.

Virtual Reality is a new frontier in data visualization and this thesis details a method by which a complex data set, Reddit, can be represented in a vivid, exciting manner, and transformed into, not only a data visualization in Virtual Reality, but an experience. Data representations are moving on from flat visualizations of information, to immersive experiences that can better represent the complex world we live in and the sheer amount of data we produce. With Virtual Reality we can better exploit the power of our visual sense, to experience information as opposed to simply viewing it.

List of Figures

List of Tables

Chapter One: Introduction

Summary

In this chapter the main topics addressed in the project will be discussed, the motivations behind the project, the high level abstract and technical problem the project solves and finally how solving these problems was approached, and what was achieved in doing so.

1.1 Topics addressed in this project

The main topic addressed in this project was to create an immersive, real time representation of the social network Reddit. In order to achieve a modern, useful and truly immersive experience the development of this representation was geared toward use in a Virtual Reality environment.
There is no greater way to provide an immersive experience at the moment than through that provided by Virtual Reality. Headsets such as the Oculus Rift and HTC Vive and mobile alternatives like Google's Cardboard and Samsung's Gear VR are providing new avenues to train professional sports stars [1] and treat posttraumatic stress disorder [2] among many other applications of the technology.

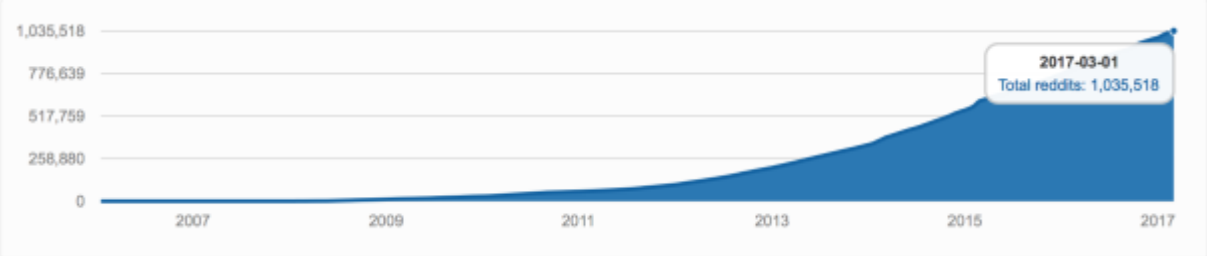


Figure 1 Growth of Number of Subreddits on Reddit, 2006/01/17 – 2017/03/01 [3]

Reddit is often described as the "front page of the internet", but the front page is just the tip of the iceberg. As of 1 March 2017, 1,035,518 Subreddits existed on Reddit according to redditmetrics.com [3], the same day approximately 16.7 Subreddits a minute were created on the site. The topic of this project is to develop a way of representing as much of this information as possible into a real time, interactive experience to provide a glimpse of a larger portion of Reddit than isn't obvious from interacting through Reddit's text based UI and in doing so give some insight into how these Subreddits are connected and finally to show something about Reddit not visible in the usual boring two-dimensional screen based world we live in.

1.2 Motivation

There is something for everyone on Reddit. There are Subreddits for everything from ‘r/FUNNY’ to ‘r/PERSONALFINANCE’ and communities aimed at programmers (r/programming), cat people (r/CatsStandingUp), freaks (r/Freaks), geeks (r/Nerd_Corner) and those who just want to watch the world burn (r/ooerintensifies) among more than one million others. But what’s the difference between r/programming and r/coding, and how are they connected? Or are they even connected? And how would it be possible to represent this information in an intuitive manner without overwhelming the end User? These are the questions I attempt to address with this project.
With this project I attempted to create a representation of Reddit that encapsulates more than individual Subreddits, individual User preferences or even a subsection of Reddit as a whole. I strived to provide the end User with an experience that immerses them deep within Reddit’s molecular build up, to develop a representation of Reddit, as much of Reddit as possible, that encapsulates how Subreddits are connected, which Subreddits are connected, which Subreddits aren’t connected and to provide a way of instantly delving deep into the soul of Reddit.

1.3 Problem Statement

Reddit is a social network compromised of more than one million Subreddits, how would it be possible to develop an intuitive and functional yet enticing representation of this complex network without overwhelming the end User? And would it be possible to show how Subreddits are connected to one another? The main technical problem that has to be solved during the development of this project is to develop an algorithm capable of interpolating Reddit data in real-time and producing useful results that can be used to represent Reddit in a three-dimensional Virtual Reality environment.

1.4 Approach

To solve the problem at hand I designed a system to graph Reddit in real time. In the graph Subreddits would be represented as Nodes and the system would update Node positions based on how they are connected too other Nodes. The system will find connections between Nodes by investigating User connections between Subreddits.

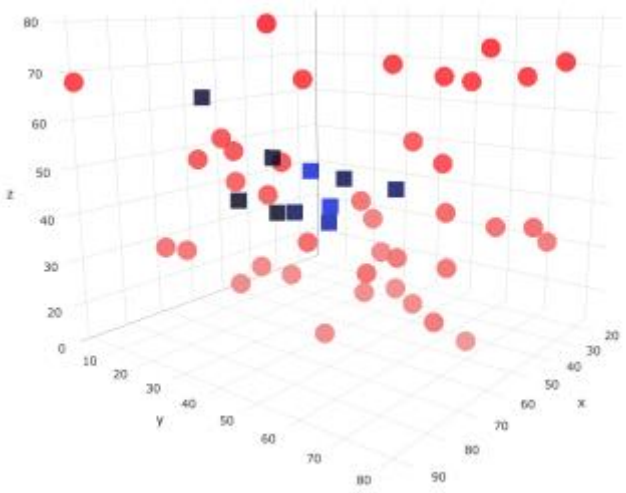


Figure 2 An early model of the Reddit Graph – Subreddits are represented as red spheres, Users are represented as blue cubes.

To decide how closely connected Subreddits are and to position them accordingly the system will investigate User connections between Subreddits. If User A comments on Subreddit0 and Subreddit1, then Subreddit0 and Subreddit1 are connected and should be updated accordingly.
In evaluating the solution the graph should provide a visual representation directly comparable to the raw data.
Initially the system acted on a large of set of Reddit data to provide a base for future real time updates. Using this base the system will search for new Subreddits and add them to the graph, updating older Nodes positions as it does so.

1.5 Project

Throughout this project there were three main achievements;

1. Developing a system to position and update nodes to graph Reddit: Developed an online algorithm to interpolate Reddit data in real time and return three-dimensional coordinates associated with Subreddits and Users in order to position them as nodes in a graph. The algorithm updates the positions of these nodes in real time.
2. Designing and implementing a three-dimensional visual representation of Reddit: Using the position data from the above algorithm a three-dimensional representation of Reddit was designed that updates in real time and allows for User interaction with nodes.

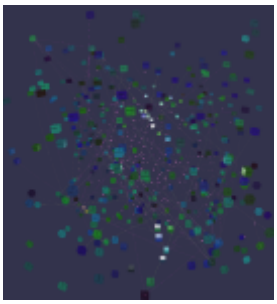


Figure 3 Screenshot of three-dimensional representation of Reddit – Subreddits are represented as coloured cubes, Users are represented as pink spheres.

3. Finally, implementing the design in Virtual Reality.

Chapter Two: Technical Background

Summary

In the following chapter the material addressed in the project will be discussed, both technical and non-technical. For example, what work has already been completed in this area and tools used throughout the development of the project.

2.1 Topic Material

There have been a variety of applications developed using the Reddit API, both real time and interactive.

The site redditviz [4] used Reddit data to create an interactive map of Reddit using the Gephi OpenOrd Layout [5] algorithm to decide the layout of the map. In the words of Fastcodedesign.com "Reddit is the "front page of the internet" but this handy tool can take you straight to page 52" [6]. Secondly, the site redditmetrics.com [7] tracks Reddit and Subreddit metrics such as "New Subreddits by date" and "Fastest growing".

Figure 4 Screen shot from redditmetrics.com of a "New Subreddits by date" graph.

Across the web there are a wide variety of three dimensional representations of data that demonstrate how data can be deployed not only as a source of information, but also as art. Some of the best examples of this "information art" are provided below.

Metrogram3D [8] is a real time simulation of the Tokyo metro built with WebGL and is a premium example of how real time data three dimensional representations can be used to convey complex network information in interesting, effective and enticing ways that a two-dimensional static simulation could never hope to achieve.

Finally, armsglobe [9] a three-dimensional "interactive visualization of government-authorized small arms and ammunition transfers from 1992 to 2010" produced by Google as part of the Google Ideas INFO Summit with the Peace Research Institute Oslo small arms database is a stunning example of how databases can be represented to convey the information at their heart.

2.2 Technical Material

Throughout the development of this project I used a variety of different technical tools and leant on many different resources in order to help implement them.

<u>The tools used throughout the project are as follows:</u>

- HTML5
- CSS5
- JQuery
- JavaScript
- TypeScript, BabylonJS
- Python, PRAW
- Reddit API
- TSlint, PYlint
- Firebase

At the beginning of the process I undertook the task of learning TypeScript and Baylonjs to develop the visualization for the project.

I also began to learn PRAW in order to obtain the data needed from Reddit with Python. PRAW was chosen due to its ease of use compared to the raw Reddit API, and my previous experience with manipulating data in Python.

This process was completed in the early stages of the project, however as the project progressed I continued to refer to a variety of resources and material such as Stack Overflow and tutorial sites.

Other technical material used included:

- Reddit API [10] this was the starting point for interacting with Reddit.
- PRAW "Getting Started". [11]
- The Python.org python-firebase [12] webpage to learn from information regarding Firebase setup in Python setup and for interacting with Firebase in Python.
- A "Getting started" tutorial on the Firebase website [12] regarding using Firebase in a web app.
- Information on how to resolve JQuery with TypeScript via Stack Overflow [14]
- Tutorials on Babylonjs [15]
- Information regarding the Leap Motion SDK [16] and getting started with the Leap Motion in Virtual Reality [17].

All of this material proved useful at certain stages throughout the project. In section 4.1 I will demonstrate in more detail how some of above technology was used in key areas throughout the development stage.

Chapter Three: The Problem

Summary

In the following chapter the project documentation and analysis of the Project Problem will be discussed. The section includes documentation regarding the development of the project and discussion surrounding both the functional and non-functional requirements of the Project.

3.1 Project UML Documentation

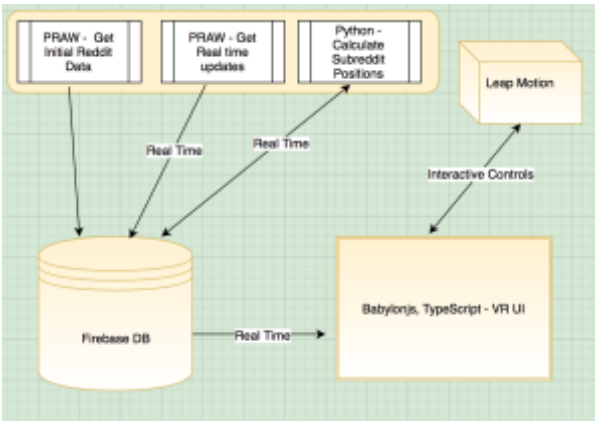


Figure 5 High Level Description of Project Architecture

3.2 Project Analysis

The main aim of this project is to develop an interactive three-dimensional visualization that in some way accurately represents Reddit, ideally a three-dimensional map. To create an accurate representation of Reddit I need some information about the site. To do this I need to gather data about Reddit as it is in real time. Using this data I need to create a program that will output some information about Reddit, such as three-dimensional coordinates corresponding to Subreddits, for use in the visualization. These outputs will need to be an accurate representation of Reddit in some way and updateable in real time. I also need some way of creating a representation of the data for viewing by the User.

Ideally the outputs, f(x), will represent how strongly connected two Subreddits are to one another. For example, if two Subreddits are connected they should move toward one another and the more strongly connected the Subreddits the greater the step toward one another should be. We want to maximize the step size as the distance between two connected nodes decreases, see below;

$$As\ x\ approaches\ 0, f(x)\ increases.\ Where\ x\ is\ the\ Distance\ between\ two\ nodes.$$

The problem can be separated into several parts:

1) Obtaining real time Reddit data.
2) Create a system using the data to output and update coordinates for Subreddit Nodes to be used in a three-dimensional map of Reddit.
3) Using the system outputs mentioned above develop, a visualization of Reddit in three-dimensions for use in a Virtual Reality environment.

The functional requirements of the project are as follows:

- Output coordinate information for Subreddits.
- Develop a system that moves nodes to represent how they are connected to other nodes.
- Update coordinate information as nodes are moved.
- Develop a functional and useful Virtual Reality ready visualization of the outputs.

The non-functional requirements are:

- Develop a visually enticing representation of Reddit.
- Develop a system that can be used across a variety of platforms and devices.

Chapter Four: The Solution

Summary

Throughout the following section, the solution to the problem at hand is discussed, in so far as to explain the tools used throughout the process and why these tools where chosen, and secondly the process by which the solution was developed.

4.1 Choice of Tools

A number of tools and languages were needed throughout the development of this project. The main languages used are Python and TypeScript.

I decided to use Python as it was the language I was most proficient in and I had previous experience using the language to manipulate data structures and interact with NoSQL databases. PRAW (discussed in section 4.1.1) is an excellent Python Wrapper for the Reddit API which was used throughout the project and made Python the ideal environment for dealing with the Reddit data.

TypeScript was used during the project as it is the language of choice for the framework BabylonJS (section 4.1.3). TypeScript is a typed superset of JavaScript that compiles to plain JavaScript [18]. The TypeScript type system can reveal many common errors as they will be flagged by the transpiler revealing many obvious mistakes in your code which can make debugging much easier, lending itself as an ideal language for web development.

In the following sections the technologies used throughout the project will be discussed in further detail.

```
tsc app/js/view.ts app/babylon.2.4.d.ts
```

Figure 6 An example of compiling a BabylonJS .ts file

4.1.1 PRAW

PRAW (Python Reddit API Wrapper) is a Python package that allows for simple access to Reddit's API. It allows you to access Reddit through calling PRAW methods on a Reddit instance. This allows for quick and easy access to Reddit data without needing to build Reddit URLs or worry about endpoints. PRAW also provides a built in delay functionality so the application will not elapse Reddit's API call limit of one request every two seconds.

```
11
12      r = praw.Reddit(user_agent=user_agent)
13      subreddits = r.get_popular_subreddits()
14
15      for subreddit in subreddits:
16
17          subreddit_ = str(subreddit)
18
19          comments_data = r.get_subreddit(subreddit_)
20          comments_data = comments_data.get_comments(limit=10, comment_sort="top")
21          comments_data_list = [i for i in comments_data]
22
```

Figure 7 Example of PRAW code from project

4.1.2 Firebase

Firebase is a web application platform that provides a real time database and backend as a service. Firebase also provides hosting for the application complete with HTTPS and SSL encryption (see Figure 8).

Firebase was chosen for the project due to its ease of interfacing with both Python and TypeScript making it an ideal choice for every aspect of the design.

Figure 8 SSL report for the project site (https://adamedwarddarcy.tech/) [19]

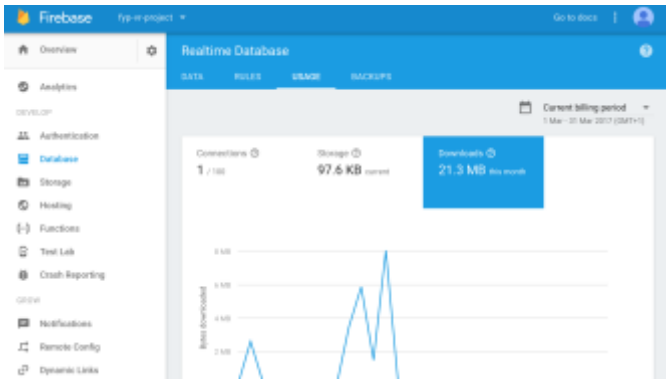


Figure 9 The Firebase console

4.1.3 BabylonJS

BabylonJS is a TypeScript framework for building three-dimensional applications with HTML5, WebGL and Web Audio. BabylonJS allows Users to develop three-dimensional environments using TypeScript and basic HTML5 elements such as the canvas element.

The built in Virtual Reality cameras and device orientation tracking sets BabylonJS apart as the perfect framework for building a Virtual Reality application.

4.2 Implementation

In the following section the implementation of the project is discussed alongside the associated files which can be found in the appendix. After deciding on the final goal of the project it became apparent there would be three main areas to the implementation of the design:

1) To develop a system to obtain real time data from Reddit.
2) To develop an algorithm that can interpolate the Reddit data in real time.
3) To develop a three-dimensional visualization of the Reddit data.

4.2.1 Getting Started

In order to begin mapping Reddit data in real time I began by obtaining a large set of batch data. To achieve this I obtained data from the 25 most popular Subreddits. Of these Subreddits I obtained the most recent 25 commenters (commenter = User) and of these 625 Users;

$$25\ Subreddits\ x\ (25\ users\ for\ each\ Subreddit) =\ 625\ users$$

I obtained the last 25 Subreddits that each User had interacted with. This process gave me a data set containing a maximum of 15,625 Subreddits.

$$625\ users\ x\ (\leq 25\ Subreddits\ for\ each\ user) \leq 15{,}625\ Subreddits$$

The data was stored in two tables in the Firebase database, a key, value table where the Keys are the Subreddits initially obtained (i.e the 25 most popular Subreddits) and the values are the corresponding Subreddits obtained from User comment data [table name = reddit_con];

Figure 10 Structure of reddit_con table

and a table containing each Subreddit with corresponding three dimensional coordinates [table name = reddits].

Associated File: get_data.py.

How these coordinates are obtained is discussed below.

Figure 11 Structure of reddits table

Initially positioning Subreddits:

Subreddit nodes are initially positioned randomly using Pythons random uniform method to achieve a uniform distribution, these positions are later updated as will be discussed in the following section.

Associated File: init_nodes.py

4.2.2 Updating Subreddit positions:

In order to achieve a useful data representation, Subreddits need to be positioned in a manner which reflects how the nodes relate to one another and the environment as a whole. Subreddit nodes also need to be updateable in an online real time manner. To achieve this functionality the reddit_con table was utilized.

To update Subreddits key value pairs in the reddit_con table are used, where keys are 'Alphas', α, and values are lists of 'Betas', β, see below; their corresponding positions can be found in the reddits table which will be updated.

The following method is then used:

Method A:

$$\alpha: [\beta_0, \beta_1, \dots, \beta_n]$$

α and β are considered connected if a β appears in an α keys value list, and will be updated accordingly. If a β appears in an α value list the β will take a step toward the α, the closer (more strongly connected) the α and β Nodes the larger percentage step the β will take, such that as the distance between an α and β decreases, the connection between them increases. Step size is calculated by the following formula;

$$\delta = |\alpha - \beta|$$

$$step = \begin{cases} if\ \delta < 1\ then\ \delta \\ else\ \left(\frac{\delta}{|\delta - e^{-1}|}\right) \cdot \delta \end{cases}$$

By the above, as the connection between two nodes grows stronger the percentage step size between the nodes increases up until a threshold of the connection (distance) between two nodes being less than 1. This is demonstrated in the below graph:

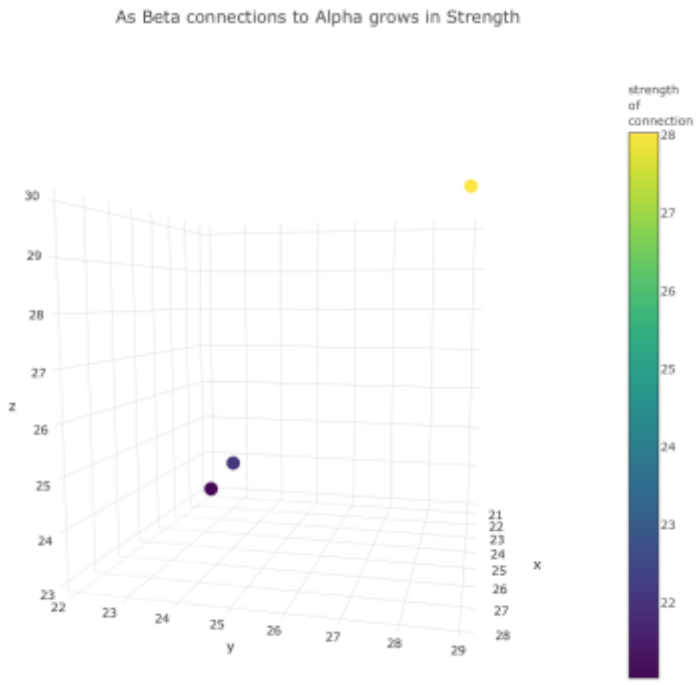


Figure 12 A graph demonstrating that; as the connection between two nodes increases, the potential step size also increases.

Following the calculation of the step size the updated position of the β node is calculated and returned, as below:

$$update = \begin{cases} if\ \alpha > \beta: \beta + (\delta \cdot step) \\ else: \beta - (\delta \cdot step) \end{cases}$$

The returned updated value is a β that has taken a percentage step toward its associated α, this step will never be larger than:

$$\delta\ where\ \delta = \begin{cases} 1 \\ \alpha - \beta \end{cases}$$

The update is then used to put an update to the database and the reddits table as shown in the code snippet below.

```
db.put('/reddits/%s'%(i),k,(update[xco],
                            update[yco],
                            update[zco]
                           )
)
```

Figure 13 A code snippet (source: update_nodes.py) demonstrating an update to the reddits table.

Associated Files: fx.py, update_nodes.py

4.2.3 Real time updates

Thus far only a small portion of Reddit has been mapped. The remainder of the Reddit network will be mapped in real time online as opposed to the batch method used thus far. To achieve this the reddits table is compared to the reddit_con table and for all reddits[keys] not in reddit_con[keys] a similar method to that used in section 4.2.1 is used, where the last 10 Users to interact with the subreddit are obtained, and the last 10 subreddits each User has interacted with are obtained to create a key value of the following form;

$$\alpha: [\beta_0, \beta_1, \dots, \beta_n]$$

Method A, from section 4.2.2 is then implemented. If a β does not have an associated position, it is assigned a position in the same manner seen in the 'Initially positioning Subreddits' portion of section 4.2.1. Once Method A has been implemented, the key, α, is added to the reddit_con table and the process continues.

Associated Files: fx.py, realtime.py

4.2.4 Three-Dimensional Representation

At this point in the process the data has been obtained (or in the case of the real time updates, will continue to be obtained in the background), and a three-dimensional representation can begin to be formed. To achieve this the framework BabylonJS was used.

BabylonJS is a 3D engine based on WebGL and WebAudio as discussed in section 4.1.3. Using the data obtained in the above sections, a three-dimensional Virtual Reality ready representation was developed where Subreddits are represented in three-dimensional space as cubes with a Subreddit name printed on the face. This was achieved using textures in BabylonJS (see below), however as each texture is different and requires a new material each for each node, instancing could not be used thus creating a performance overhead which will be discussed later in section 6.2.

```
let boxTexture = new BABYLON.DynamicTexture("dynamic texture", 200, scene, true);
boxMaterial.diffuseTexture = boxTexture;
boxTexture.drawText(key, null, 100, "bold 40px helvetica", "Black", color);
```

Figure 14 Code snippet demonstrating how to create a texture and add text to a BabylonJS object. (source: view.ts)

Connections between Nodes are displayed as lines, the thicker the line the stronger the connection. To achieve this affect the Euclidean distance, δ, between the node is calculated and used in the following formula to return a value for the line thickness, τ;

$$\tau = 1/\delta$$

and the result is passed as a parameter to the BabylonJS 'CreateTube' method as the line radius. Lines are created between Alpha and Beta Nodes, such that for all Keys in the reddit_con table, a connecting line is drawn between it and all of its corresponding values, see Figure 15.

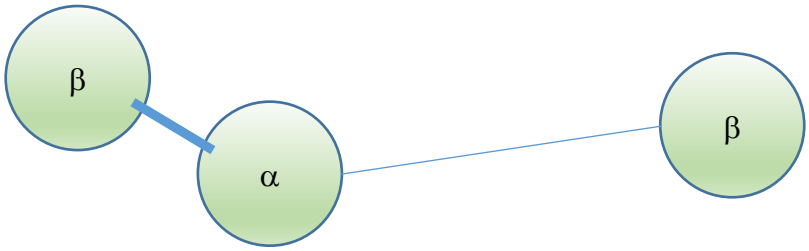


Figure 15 Demonstrating the connections between alpha and beta nodes.

The BabylonJS VRDeviceOrientationFreeCamera [20] is used to provide Virtual Reality for the application.

Finally, aside from being Virtual Reality ready and real time the final functional requirement of the application is to be interactive. Interactivity is achieved through the BabylonJS action manager and is activated by clicking on Nodes. By triggering a Node, information regarding the Node will be served to the User, such as the last ten comments, the Node key, and the Node position. In an early iteration of the visualization User nodes where included as pink spheres, however this functionality was removed after User testing when Users conveyed that the User nodes made the visualization confusing, this will be discussed further in a later section.

Associated Files: view.ts, view.js (Generated from view.ts), index.html

Chapter Five: Software Verification

Summary

During the development of the project a variety of verification and testing tools where implemented in order to guarantee a reliable final product. User and Unit testing where implemented to guarantee the software's viability and robustness whereas linting tools where used to promote the creation of the highest quality code achievable and guarantee standards where followed correctly. The results of

these methods will be discussed throughout the following the sections.

5.1 Unit Testing

To validate the positioning functionality seen in Method A of section 4.2.2 the method *f()* contained in *fx.py* underwent unit testing. The unit tests where implemented to guarantee that;

1. A returned step size will be less than that of the difference between an alpha beta pair.
2. A step size will always be greater than one.
3. Updates will always move toward the alpha.

During the initial phase of development a problem was discovered where an input such that an alpha value is smaller than a beta value would result in a negative delta and an output of infinity or minus infinity causing the above tests 1 and 2 to fail. Due to this behavior a decision was made to update the functionality of the method such that delta was calculated using the absolute value as seen in Method A, see below:

$$version0: \delta = \alpha - \beta$$

update:

$$version1: \delta = |\alpha - \beta|$$

The above update solved the 'infinity' problem however the system then began to fail test 3, as the result of the update was always positive, to deal with this failure another update to the system was necessary.

In order to solve the problem of the system failing test 3 an update to the systems control flow was necessary, such that:

$$version1: update = \beta + (\delta \cdot step)$$

update:

$$version2: update = \begin{cases} if\ \alpha > \beta: \beta + (\delta \cdot step) \\ else: \beta - (\delta \cdot step) \end{cases}$$

The above update to the system guarantees a position update will always be in the direction of the alpha node, and therefore passes test case 3. After the final update was implemented the system passed all three test cases.

5.2 User Testing

In order to guarantee the creation of a coherent usable design the application was User tested and User results were obtained through a survey, the contents and results of which are included in the appendix. The results of User testing are discussed below;

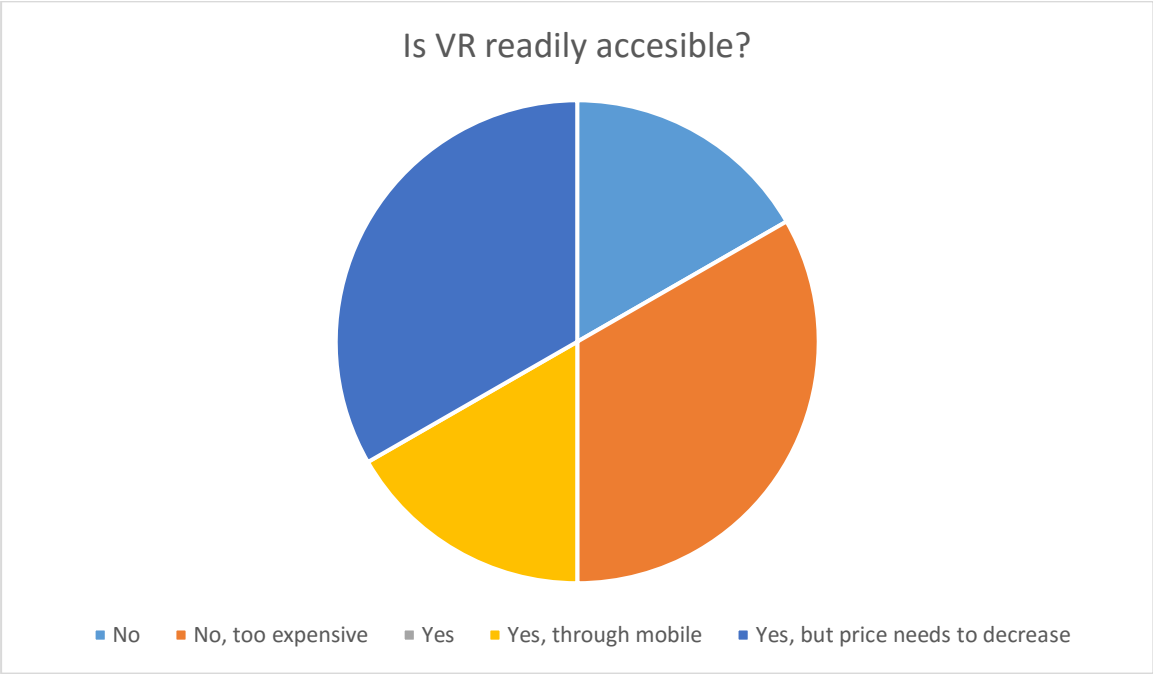
User Nodes

Initially User nodes were included in the application design as pink spheres, however during User testing it was conveyed by the Users that the User nodes made the application confusing or cluttered, hence the decision was made to remove the User nodes from the final application design.

Accessibility

The cost of a high end Virtual Reality headset such as the Oculus Rift, and the machine needed to power the device, makes these devices and the experiences geared toward them, far from accessible. During User testing, 0% of people thought Virtual Reality was readily accessible, a results breakdown follows;

Table 1 Is VR accessible?



Due to the overwhelming User response that VR is too expensive or only readily accessible through mobile, the decision was made to attempt to make as much of the application experience as possible available on mobile and less powerful machines alongside high end machines geared toward Virtual Reality usage.

Ease of Use:

During User Testing, the application was rated highly on usability while on a machine. However, when Users accessed the applications on mobile, usability ratings suffered. In the future mobile VR experiences will benefit from dedicated mobile controllers, this concept will be further discussed in a later section.

5.3 Linting

To guarantee code quality and keep up to date with style guides and code guidelines in a variety of languages, primarily Python and TypeScript, two Linting tools, Pylint [21] and TSLint [22] were integrated into the development environment.

Linting is a process by which code is checked for programmatic and stylistic errors, thus verifying code quality. This process can be particular helpful while developing for web applications, where errors perhaps may not be caught until they are seen in the browser. The combination of TypeScript and TSLint made for a much more efficient and streamlined development process.

The Python lint tool Pylint, deploys a scoring system (out of 10) to rate code, having this system in place allowed me to set a target of 80% code quality at the beginning of the development process. When development was complete, an overall code quality of 94.07% was achieved, a results breakdown follows;

Table 2 Linting: Results breakdown

File	Score/10
__main__.py	10.00
testA.py	8.70
data/get_data.py	10.00
compute/fx.py	9.38
compute/init_nodes.py	10.00
compute/update_nodes.py	9.33
compute/realtime.py	8.44

Unfortunately, TSLint does not provide a scoring system, however the tool does reveal a host of Stylistic and Programmatic errors or suggestions which can directly lead to the creation of better code, and which where adhered to as much as possible throughout the development of the project.

Chapter Six: Conclusion

Summary

In this final chapter, how the project has been approached, future developments and conclusions are discussed.

6.1 Project Approach

In regards to how to approach the project, at the beginning of this process three main goals were set:

1. To learn TypeScript, BabylonJS and PRAW to the best of my abilities.
2. To develop a novel, interesting and enticing representation of the social network Reddit.
3. To write clean, quality code.

As the development stage of the project progressed I became increasingly proficient in TypeScript, BabylonJS and PRAW. I found TypeScript in particular a joy to use and plan to implement the language in future projects.

As I began development on the visualization I heeded to the simple mantra of '*novel, interesting, enticing*'. My approach to creating a powerful visual experience for the User was to achieve three main goals;

1. The representation must be a novel view of Reddit.
2. The representation must be interesting.
3. The representation should be visually enticing to the User.

As the development process progressed I referred to these concepts resulting in the final visualization that is full of colour and presents as much information as possible in a simple visualization.

Finally, throughout the development process I strived to create code of the highest quality, this was partly achieved through the use of linting tools and thorough testing as discussed in chapter 5. Through the use of these mechanisms I approached the project in a manner that would be beneficial to creating a robust, functional and quality product.

6.2 Future Developments

As was mentioned in section 4.2.4, to draw unique text on a BabylonJS mesh, a new material for each mesh is required. Thus instancing of meshes becomes impossible. "Instances are an excellent way to use hardware accelerated rendering to draw a huge number of identical meshes" [23]. For example a forest or in the case of our application, a large number of cubes all sharing the same root mesh. In the future I would contribute to the BabylonJS community by attempting to develop a method by which unique text can be drawn on instanced meshes, thus creating a more efficient visualization that uses less GPU resources.

During User testing it was conveyed that the application was difficult to control on mobile. To deal with this complaint I would implement voice controls through a library such a annyang [24], or

implement the event listener ‘ondevicemotion’ [25] to track device motion and use the output to allow the User to interact with the application.
As the application is now, real-time updates are activated through a python script, *realtime.py,* ran on a personal machine. In the future however it would be ideal to have the script on a server, running perpetually.
A Leap Motion was implemented as a controller throughout the development of this project. However due to some compatibility issues between the Leap Motion and BabylonJS, the controller functionality is not always reliable. In the future I would like to implement a more reliable controller infrastructure, in order to provide a fuller experience to the end User.
Finally, in the future I would implement functionality by which the User can interact with the graph and the physics of the graph as opposed to just the nodes. This functionality could potentially include allowing the User to deselect nodes to see how the graph reacts or, by implementing WebAudio, to make different sounds based on were the User has selected on the graph.

6.3 Conclusions

> “Ninety percent of all of the data in history, by one estimate, was created in the last two years. In 2014, International Data Corporation estimated the data universe at 4.4 zettabytes which is 4.4 trillion gigabytes. That volume of information, the research firm said, straining for perspective, would fill enough slender iPad Air tablets to create a stack more than 157,000 miles high, or two thirds of the way to the moon.” [26]

The above extract is from the 2015 book ‘Data-ism’ by Steve Lohr. Two years on from 2015 I estimate that stack of iPads has long surpassed the moon. With such a proliferation of data, visualizations have become increasingly important in how we consume data, we even turn to visualization in order to understand the sheer scale of the data we are consuming, see that stack of iPads currently shooting further into our solar system.
During User Testing 100% of respondents stated they preferred a data visualization, such as an infographic, to data presented in a plain text format, and this project is proof of just how much information can be fitted into a single visualization.
Virtual Reality is an opportunity for data visualization, an opportunity to enable Users to experience data, and to grasp a new understanding of information from these Virtual experiences. However, the benefits are in no way one sided. Big Data and Virtual Reality are commencing on a symbiotic

relationship, which has the potential to benefit society as a whole. Virtual Reality is already contributing to the world, by aiding surgeons in their training, as evidenced by the Royal College of Surgeons Ireland's 'Medical Training Simulator' [26]; and as the technology improves, society will continue to reap the benefits.

Virtual Reality may need to undergo some changes before it truly proliferates into the zeitgeist, such as decreasing in price or the development of more mobile controllers such as Google Daydreams controller [28]. However, as Parmy Olson of Forbes states;

> "For now virtual and augmented reality gadgets seem best place[d] for use in enterprise settings, where issues like price and discomfort can be forgiven so long as the technology makes itself useful in a work setting." [29]

Virtual Reality is positioned ideally for use in enterprise and professional environments. In the near future doctors will use an Oculus Rift to visualize how a certain bug is travelling through a local community, firefighters will train to deal with blazing forest fires without leaving the confines of the station's HTC Vive, and high up in a New York sky scraper, stock traders will visualize the consequences of a potential trade in India, with a mobile headset and the same iPhones they use to order their lunches from Deliveroo. 'Virtual' may be a key term regarding these examples, however the effects they have on society will be very, very real.

References

1 Loh, Stephanie. 2015. “Virtual Reality QB trainer ‘a game changer’“, The San Diego Union-Tribune. Available at: http://www.sandiegouniontribune.com/sports/sdut-virtual-reality-quarterback-trainer-derek-belch-2015may21-story.html [Accessed March 2, 2017]

2 Mary Aiken & Mike J. Berry: Posttraumatic stress disorder: possibilities for olfaction and Virtual Reality exposure therapy, *Virtual Reality,* June 2015, Volume 19, Issue 2, pp 95–109

3 Redditmetrics.com. “New reddits by month (How Reddit grew over time)”. Available at: http://redditmetrics.com/history/month/ [Accessed March 1, 2017]

4 Redditviz. “redditviz – reddit interest network”. Available at: http://rhiever.github.io/redditviz/ [Accessed October 12, 2016]

5 Marketplace.gephi.org. “openord-layout”. Available at: https://marketplace.gephi.org/plugin/openord-layout/ [Accessed March 2, 2017]

6 Wilson, Mark. 2013. “This Brilliant Visualization Could Build A Better Reddit”, fastcodedesign.com. Available at: https://www.fastcodesign.com/3021066/infographic-of-the-day/this-brilliant-visualization-could-build-a-better-reddit [Accessed March 2, 2017]

7 Redditmetrics.com. Available at: http://redditmetrics.com [Accessed March 1, 2017]

8 Nulldesign.js. Metrogram3d, “Tokyo Metro Realtime Simulation”. Available at: http://nulldesign.jp/metrogram3d/ [Accessed October 4, 2016]

9 Chromeexperiments.com. “Small Arms and Ammunition – Imports & Exports”. Available at: https://armsglobe.chromeexperiments.com [Accessed October 4, 2016]

10 Reddit.com. “API DOCUMENTATION”. Available at: https://www.reddit.com/dev/api/ [Accessed September 22, 2016]

11 Praw.readthedocs.io. “Getting Started”. Available at: https://praw.readthedocs.io/en/v3.6.0/pages/getting_started.html [Accessed October 5, 2016]

12 Pypi.python.org. python-firebase 1.2. Available at: https://pypi.python.org/pypi/python-firebase/1.2 [Accessed November 3, 2016]

13 Firebase.google.com. Documentation “Installation & Setup in JavaScript”. Available at:

https://firebase.google.com/docs/database/web/start [Accessed October November 4, 2016]

14 Sherret, David. 2015. Stack Overflow. Available at: http://stackoverflow.com/questions/32050645/how-to-use-jquery-with-typescript [Accessed November 22, 2017]

15 Docs.babylonjs.com. Tutorials. Available at: https://doc.babylonjs.com/tutorials [Accessed November 18, 2016]

16 Developer.leapmotion.com. JavaScript SDK Documentation. Available at: https://developer.leapmotion.com/documentation/v2/javascript/index.html [Accessed October 5, 2016]

17 Developer.leapmotion.com. "Get Started with our SDK". Available at: https://developer.leapmotion.com/get-started/ [Accessed November 18, 2016]

18 Typescriptlang,org. Available at: https://www.typescriptlang.org [Accessed September 23, 2016]

19 Ssllabs.com. Qualys SSL Labs. "SSL Report: adamedwarddarcy.tech". Available at: https://www.ssllabs.com/ssltest/analyze.html?d=adamedwarddarcy.tech&s=151.101.65.195 [Accessed March 5, 2017]

20 Babylonjsguide.github.io. "Advanced Cameras". Available at: http://babylonjsguide.github.io/advanced/Cameras [Accessed January 27, 2017]

21 Pylint.org. Available at: https://www.pylint.org [Accessed September 3, 2016]

22 Palantir.github.io. TSLint. "An extensible linter for the TypeScript language". Available at: https://palantir.github.io/tslint/ [Accessed March 3, 2017]

23 Docs.babylonjs.com. Tutorials. "Instances are an excellent way to use hardware accelerated rendering to draw a huge number of identical meshes". Available at: https://doc.babylonjs.com/tutorials/how_to_use_instances [Accessed March 5, 2017]

24 Talater.com. Annyang!. "SpeechRecognition that just works". Available at: https://www.talater.com/annyang/ [Accessed March 11, 2017]

25 Mozilla Foundation. Developer.mozilla.org. "Window.ondevicemotion". Available at: https://developer.mozilla.org/en-US/docs/Web/API/Window/ondevicemotion [Accessed March 11, 2017]

26 Steve Lohr, Data-ism, 2nd edition, HarperCollins Publishers, New York, 2015

27 Msurgery.ie. "VR RCSI Medical Training Sim". Available at: http://msurgery.ie/vr [Accessed March 18, 2017]

28 Vr.google.com. Daydream View. Available at: https://vr.google.com/daydream/headset/ [Accessed March 17, 2017]

29 Olson, Parmy. 2017. “The Hype Around Virtual Reality Is Fading”, Forbes. Available at: https://www.forbes.com/sites/parmyolson/2017/03/03/the-hype-around-virtual-reality-is-fading/#353baf7d1344 [Accessed March 15, 2017]

Appendix

User Testing – Results

Timestamp
2017/03/15 1:25:45 PM GMT
2017/03/15 1:29:38 PM GMT
2017/03/15 1:39:09 PM GMT
2017/03/15 1:41:19 PM GMT
2017/03/15 3:06:12 PM GMT
2017/03/15 7:15:56 PM GMT

Have you ever experienced Virtual Reality
Yes
Yes
Yes
No
Yes
Yes

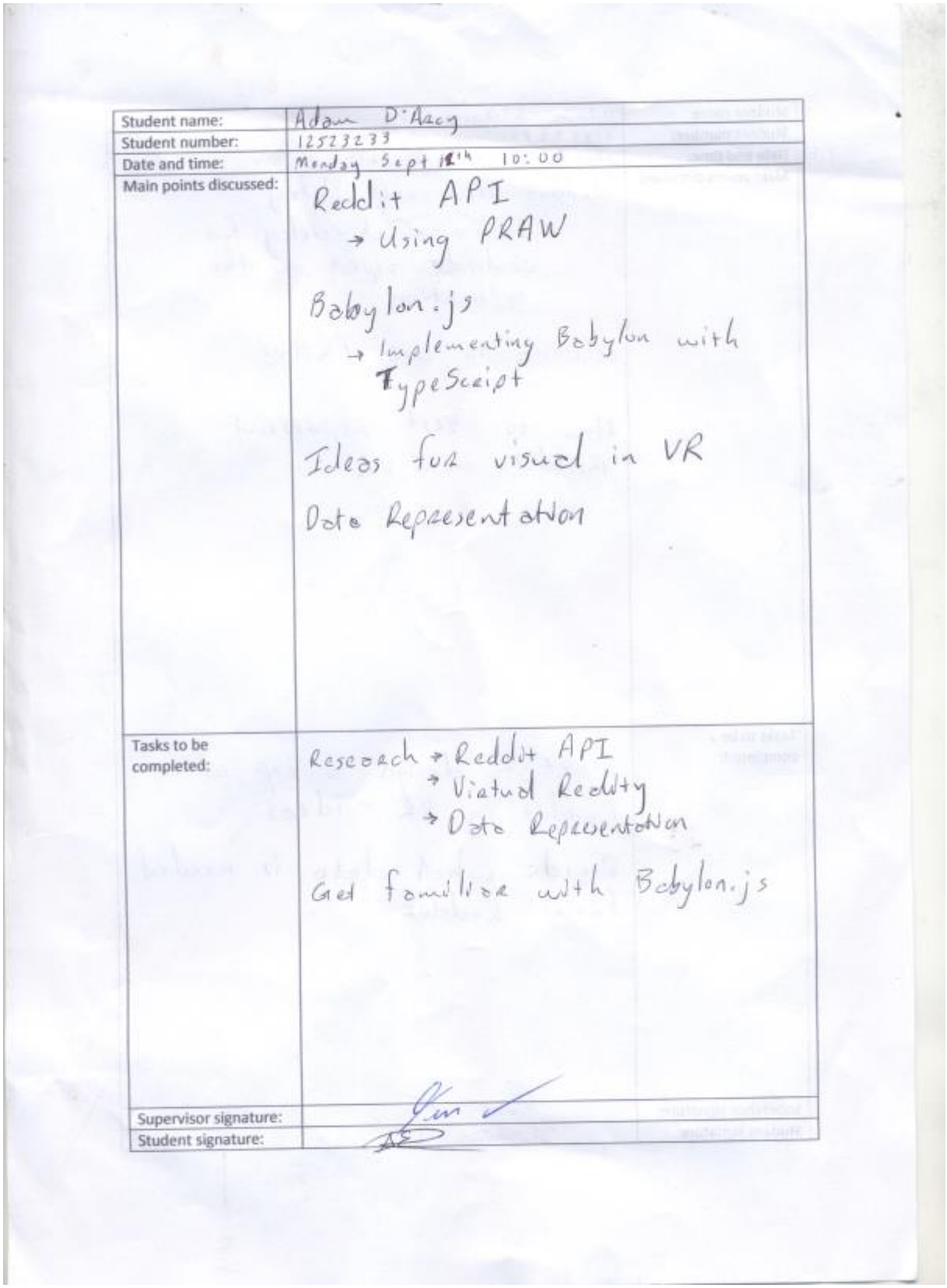
If so, how would you rate the experience?
7
10
8
9
8

At the moment, do you believe good VR experiences are readily accessible?
No
Yes but the price needs to decrease
Yes, but through mobile
No, too expensive
Yes but the price needs to decrease
No, too expensive

Do you prefer data presented in an infographic or a plain table, please give a short reason why?
Yes, easy to read
Infographic
infographical as it feel it displayes the information clearly
Infographic is much easier to follow.
Infographic - Easier to look at and understand
Infographic - it's easier to understand the presented info

How would you rate the ease of use of the application?
9
9
8
9
7
8

Meeting Feedback Forms



Student name:	Adam D'Arcy
Student number:	12523233
Date and time:	Monday Sept 12th 10:00
Main points discussed:	Reddit API → Using PRAW Babylon.js → Implementing Babylon with TypeScript Ideas for visual in VR Data Representation
Tasks to be completed:	Research → Reddit API → Virtual Reality → Data Representation Get familiar with Babylon.js
Supervisor signature:	
Student signature:	

Student name:	Adam D'Arcy
Student number:	12523273
Date and time:	Monday September 26th 11:00
Main points discussed:	Reddit API call Delay → dealing with delay for realtime aspect of the application Decision to use PRAW How to best represent Reddit in VR
Tasks to be completed:	Start to design a rep of Reddit in VR - ideas Decide what data is needed from Reddit
Supervisor signature:	
Student signature:	

e

Student name:	Adam D'Arcy
Student number:	12523237
Date and time:	Monday October 3rd 10:00
Main points discussed:	Idea to develop "map" of Reddit in VR → Based on Reddit Nodes connection to each other through users → Nodes updated in Realtime Implenting a database to deal with Reddit API delay
Tasks to be completed:	Begin to develop node positioning algorithm in Python Investigate databases
Supervisor signature:	
Student signature:	

f

Student name:	Adam D'Arcy
Student number:	12523233
Date and time:	Monday October 17 11:00
Main points discussed:	SQLite } using Firebase to store data Firebase → ease of use for Python and Typescript How to position users for visualisation → weighted average of associated Reddit positions?
Tasks to be completed:	Implement Firebase Implement position - user functionality
Supervisor signature:	
Student signature:	

Student name:	Adam D'Arcy
Student number:	12523233
Date and time:	Monday November 14 10:00
Main points discussed:	• updating node positions • Babylonjs visualisation • using a leap motion to interact with the data representation in VR
Tasks to be completed:	• Begin work on visualisation in typescript + Babylonjs • investigate leap motion for VR
Supervisor signature:	
Student signature:	

h

Student name:	Adam D'Arcy
Student number:	12522273
Date and time:	Thursday February 16 10:15
Main points discussed:	Interacting with the VR representation Fixing vertice issue in Babylonjs Implementing leap motion fixing texture issue
Tasks to be completed:	Fix vertice bug / texture issue
Supervisor signature:	
Student signature:	

Student name:	Adam D'Arcy
Student number:	12523183
Date and time:	Thursday February 23 10:35
Main points discussed:	• implementing heat map • interactivity functionality • thesis write-up • UI clean up
Tasks to be completed:	implement interactivity functionality
Supervisor signature:	
Student signature:	

j

Student name:	Adam D'Arcy
Student number:	12523233
Date and time:	10:00 Thursday 9th March 2017
Main points discussed:	• Info page, FAQ? • Firebase Auth • Thesis write up - Algorithm - Graphs - Adding Technology 2.7.x sections - Diagrams • Style guides
Tasks to be completed:	• Add info button for app • Continuing working on thesis
Supervisor signature:	
Student signature:	

k