



College of Engineering

# CS CAPSTONE PROBLEM STATEMENT

OCTOBER 15, 2017

## ANCESTRY DATA VIEWER

PREPARED FOR

CHEAP ROBOTS, INC

ASHLEY McGRATH

\_\_\_\_\_  
*Signature*

\_\_\_\_\_  
*Date*

PREPARED BY

GROUP 22

ANCESTRY DATA VIEWER

LE-CHUAN JUSTIN CHANG

\_\_\_\_\_  
*Signature*

\_\_\_\_\_  
*Date*

MONICA SEK

\_\_\_\_\_  
*Signature*

\_\_\_\_\_  
*Date*

YONG PING LI

\_\_\_\_\_  
*Signature*

\_\_\_\_\_  
*Date*

### Abstract

The goal of this project is to create a program that will open GEDCOM files and show the data contained in a clear and easy to read manner. The user should be able to view their inputted family tree as a whole graph, or be able to zoom in to a specific person. Additionally, the application should be capable of switching to VR if the user has a VR headset and wants to use it. Finally, the application should be capable of finding the nearest common ancestor between two given family members if the ancestor exists.

## 1 PROBLEM

The GEDCOM file format is difficult to read in plain text, yet there is no well known, standard reader to open and view it. This makes it hard to view a familys data when it is contained within a GEDCOM file, which is problematic since GEDCOM files are a fairly common standard for storing a family tree in to a computer file. The GEDCOM file format is composed fairly efficiently, so it is reasonable to not want to replace it with a more readable but less efficient format, especially since it is possible to read it normally, due to the fact that it is possible to write a program that can read it and display it in a visually pleasing and clear format.

Another issue with how the GEDCOM file is formatted is that it doesnt necessarily keep a family tree in order. This is understandable, due to the fact that parents can have children later in their lives, and people can marry late, so if it is updated, a person from a generation might not be included with everyone else of that generation. This would make simply reading a GEDCOM file significantly more confusing than it normally would be, which, due to the formatting already included in the file, is already fairly complicated.

Finally, GEDCOM files dont allow users to search for common points of relation. While it is reasonably simple to search for a specific person in a GEDCOM file by using control + F and typing their name, it is much more difficult to find the nearest point that a person and their third cousin are related, or even if they arent related and their third cousin married in to the family. While this is not necessarily a deal breaker for a viewer application, it would be a good feature to include, due to the fact that it would receive a great amount of usage.

## 2 SOLUTION

In order to solve the problems presented above, a new application dedicated to reading and formatting GEDCOM files in to a readable format should be created. To solve the problem with how difficult reading a GEDCOM file is to read in plain text, the application should be capable of parsing GEDCOM files and outputting all pertinent data in to a human-readable format.

Additionally, in order to maintain readability, and to keep the data clear and visually pleasing, when the data is outputted, the application should output all data in the form of an undirected graph, with the parents being the parent node of the children, and the progenitors of the family tree as the root nodes. Additionally, VR should be implemented along with normal application view, in order to increase interest. For VR implementation, the same undirected graph that was used for normal application view should be generated, but it should generate the graph in 3D, so the viewer can orbit the tree and zoom in to specified nodes as they desire.

To solve the issue of finding where a certain common ancestor is between two given people, the application should include the functionality to travel up and store their parent nodes, until a parent node appears twice, which would indicate that the nearest common ancestor has been found. Once they are found, the application will output the name of common ancestor and allow the viewer to zoom to them. If no such ancestor exists, then a message indicating such should be outputted instead.

## 3 METRICS

The most important metric for the created application should be that every feature is completed, and can be run with no errors. Additionally, the application should be capable of running on Windows 10 and Ubuntu Linux with no issues.

For the graph, it should be capable of generating a graph from a given GEDCOM file. The graph should have every family member in the right order, so that a family's child should be below their parents, and so on, and the graph should have no overlap, so that every node is readable as long as the viewer zooms in on it to read it. The graph view should also be capable of zooming on to specific nodes, so that the user is capable of reading what is written about them in the GEDCOM file.

For searching for the nearest common ancestor, the application should be capable of correctly finding either the nearest common ancestor, and allowing the viewer to zoom in on them, or finding that there is no ancestor, and outputting that there is no connection between the given people. The search should also fail and output that no such person exists if the user attempts to input people who do not exist on the GEDCOM file.

For VR integration, the graph should be in 3D, and have every feature mentioned above. There should be no restrictions on VR that are not also in the application view, barring unforeseen errors with operating systems.

Other metrics that the application should be measured by is ease of use, aesthetics, and speed. The application should be fairly intuitive to use for anyone who is reasonably computer literate, and the application should be, at worst, unintrusive, so that the user isn't concerned about how cluttered or how gaudy the application appears. Finally, the application should be capable of reading, generating, and loading the graph in a reasonable amount of time on a reasonable laptop, here defined as a moderate, \$600 laptop from 2012. The longest delays should follow switching to VR, and initially loading the GEDCOM file.