Author: Shivangi Ajaykumar Bhatt
B00 # : B00863408
Subject Code: CSCI 3901
Date: Nov 4th, 2021
Project Milestone#: 0 (High-level breakdown analysis of the problem)

There would be 3 major classes:

- PersonIdentity
- FileIdentifier
- BiologicalRelation

What comes into the program?

For class PersonIdentity:

- The name of the person
- Map<String, String> of attributes like DOB, occupation, gender, etc. for each person in the family tree database
- Source reference of a person
- Note for the person
- The relations between individuals (Parent/Child, partnering relations, partnering dissolutions).

For class FileIdentifier:

- Location of the file
- Map<String, String> of attributes like year, date, and city for each media file.
- List of people that appear in the media file.
- Tags for any media.

What transformations do I need to make to the data?

Create a family tree for all the available data for people as vertices and their parent/child relation as edges.

What part of the data is processed right away?

- Data related to the personIdentity like name of the person, attributes, reference, notes are processed right away.
- Data related to the FileIdentifier like the location of the file, its attributes, list of people in media file, and tags are processed right away

What part of the data do I need to keep longer?

- Relations between the individuals

What goes out of the program?

- Biological relation between individuals
- Ancestors and descendants of individuals
- Media files belonging to the individual

- Media files belonging to the individual’s children.
- Media files belonging to a certain location.

What assumptions can I make?

- Focus only on biological relation
- Dates can be year, month and year or day, month and year

What constraints exist?

- Dates must atleast include the year.
- File name for media file cannot be empty or null.
- Person name in PersonIdentity cannot be empty or null.

Are there strange cases to handle?

- Finding biological relations between individuals that do not exist in the database.
- Finding media files that do no exist for a number of constraints.

What is important for the solution to do?

- Creating the biological tree and finding the biological relation between the individuals.
- Finding the media file for a given constraint.