

CSCI 3901 Highlevel Project Breakdown

Objective : The course project is the opportunity to demonstrate all of the concepts from the course in one body of work.

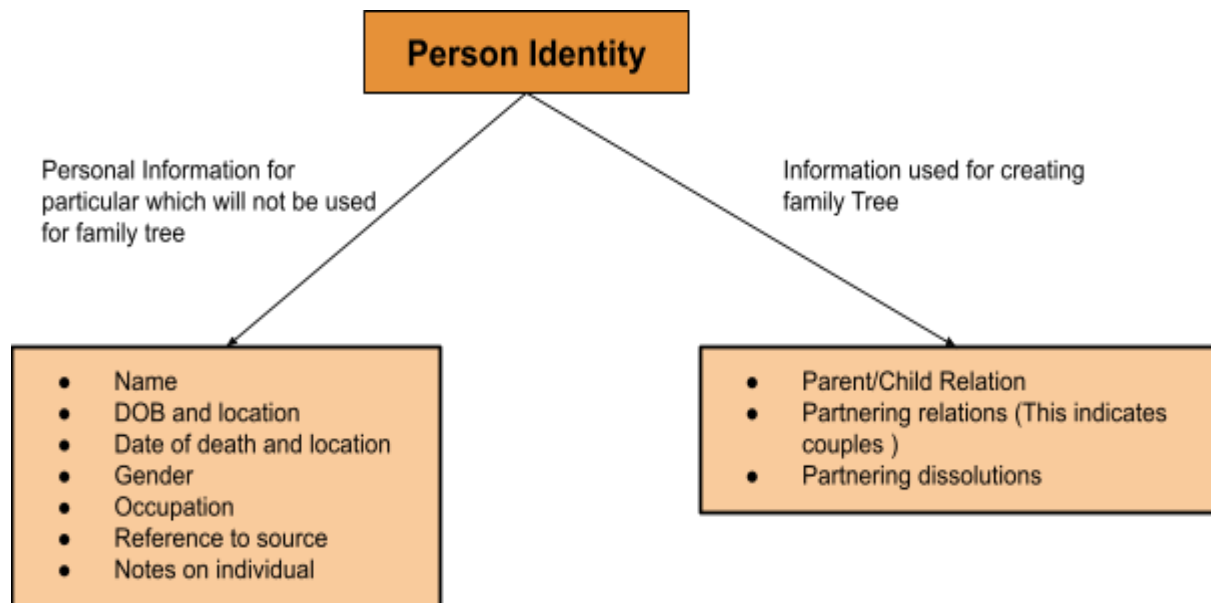
Report by:

- Shivam Bhojani (B00895637) - shivam.bhojani@dal.ca

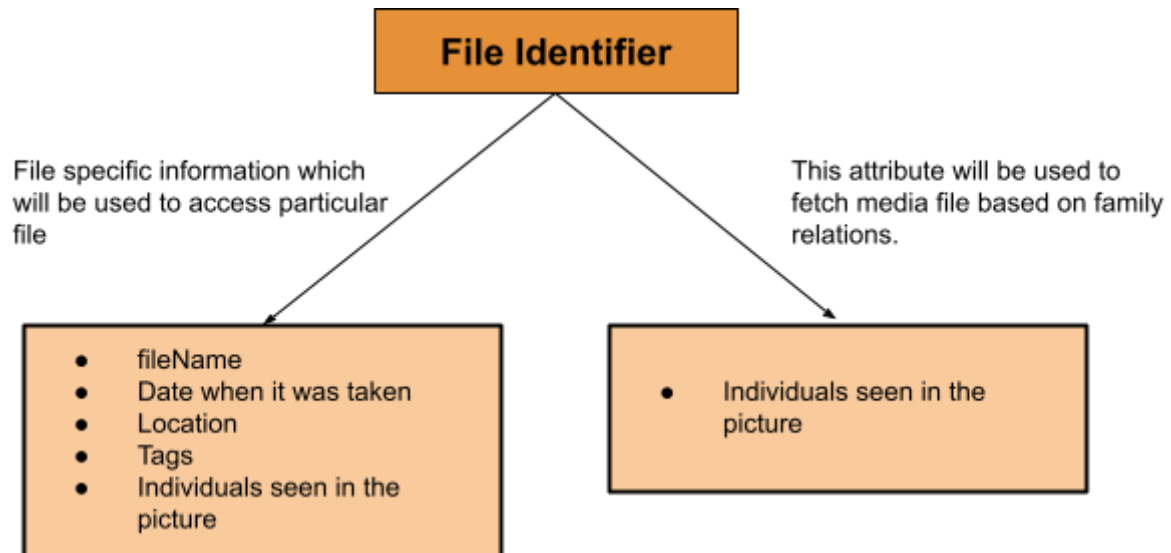
Project Understanding and simplification

- This problem statement deals with easy searching of media files based on the media archive. This is the basic version of how Google photos works. Google photos combines the images of a particular person based on the face recognition system.
- Here in this project definition, we will be creating family tree and apply the same logic in database queries in order to fetch media files
- Suppose, there are thousands of media document files in a system and also we have data which states that who are the people involved (seen) in the particular file.
- We also have a person database which states multiple attributes about a person, including their children and parents
- Based on these database tables, a user will be able to find the media file based on the people seen in the particular file.
- Now, we will be creating a family tree during runtime which will be using the data attributes from Person Identity.
- After creating a family tree, it will be possible to find the name of ancestors and descendants of a particular person.
- Suppose the genealogist wants all the media files which includes X and his previous 4 generations. The logic will run in this way in code:
 - Find ancestors of X until previous 4 generations
 - Have that data saved in a variable
 - Execute a query in the mediafile archive table and find the media files which have all X and his 4 ancestors in it.

Data Understanding



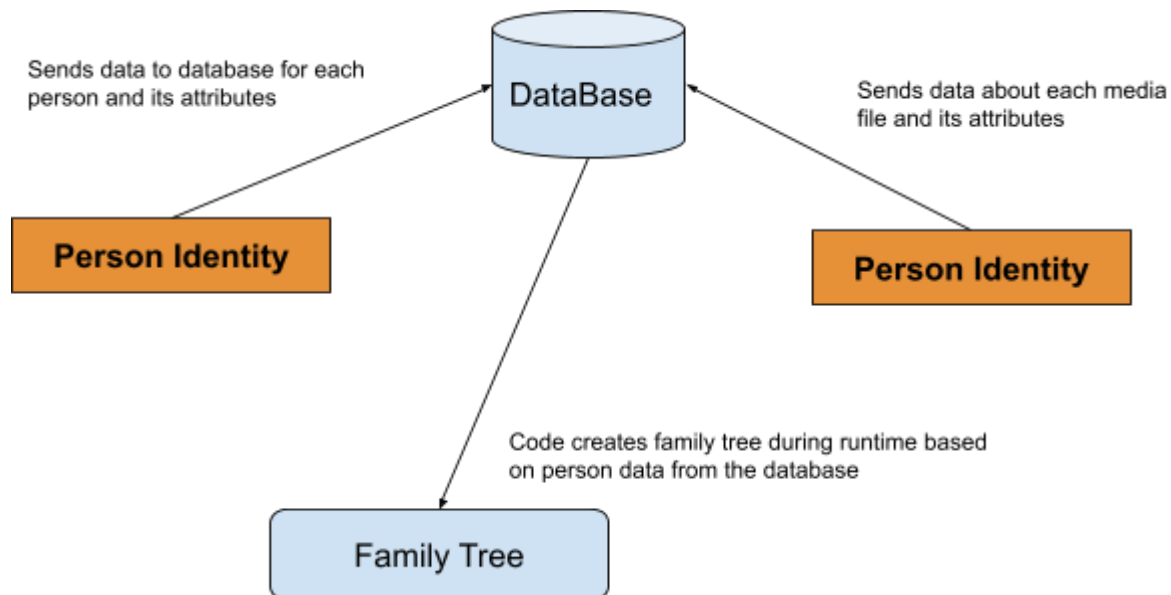
- As per my understanding, I have bifurcated the PersonIdentity data into two parts.
 - One which is personal and it has no contribution in creating family tree
 - Second which
 - As per different data input methods in project description, I understood that Person Identity will receive several type of information from user and that data will be stored in database
 - With reference to the above diagram, some personal level information like Name, DOB, locations, Gender extra are considered as personal level information, which has no relation in creating family tree.
 - This data will be used to fetch data from the database when conditions require specific personal level attributes as conditions.
 - Fetch Notes based on person Name
 - Fetch Media files where person seen in media == personName (given in argument)
 - Fetch person details by Name
-



- In File Identifier, the methods will be used for taking inputs to some individual attributes such as fileName, Date, Location, Tags, etc. This attribute does not support getting data with respect to personal details.
- The data about individuals seen in the picture actually relates with the person Identity database.

Data Flow

1. How different data is stored and how the family tree is created.



High Level - Step by Step Implementation Plan

(It does not contains data types details but it explains my step by step approach to achieve the solutions)

1. Have database Schema ready
2. Add Person to database
3. Add Attributes by using Map as Argument in the particular method
4. Test the Person Identity database queries with Insert, Select and Update queries
5. Create fileIdentifier database based on decided schema
6. Add fileNames with help of method
7. Add Attributes by using Map as argument in the particular method.
8. Test fileIdentifier with with Insert, Select and Update queries
9. Test by fetching media files data by putting validation on Person Identity attributes in validation
10. Add validations for duplicate person entry
11. Add validations for duplicate filename entry
12. Add validation on invalid argument to fetch person data
13. Add validation on invalid argument to fetch media files data

14. Add Parent/Child Relation, Partnering relations (This indicates couples), Partnering dissolutions
15. In runtime, a family tree should be created based on which different types of relationships, ancestors and descendants will be identified.
16. Create logic to count cousinship and difference in generations, in order to identify relationships between two person
17. Fetch relationship details between two person by using BiologicalRelation class

Breakdown Points

What comes into the program?

- PersonIdentity data
- Relationship data between two people
- Media archive data

What transformations do I need to make to the data?

- Based on relationship data like child - parent relations, partnering relations, partner dissolution relations, there needs to be a family tree generated on the back-end.

What part of the data is processed right away?

- Personal information of Person Identity like Name, gender, DOB, Date of death, location, Tags are such data which are unique with respect to a particular person. So that data will be stored directly in the database as per the schema.
- All media identifier data will be stored directly into database without any processing in it

What part of the data do I need to keep longer and what part can be deleted soon?

- The data about a person, their relations with others and medical file data will remain in the database for a longer amount of time.
- Adding Attributes to the database via Map. Hence Map data can be flushed out once database has completed filling the attribute data from Map
- During runtime, a family tree will be created using the information about

What goes out of the program?

- Person Data which is unique to particular person
- Person to Person relationship. (Siblings, Child, Grand Child, 1st Cousin, etc)
- Find media files based on people seen in the media file == person X
- Find media files based on people seen in the media file == Relationship of person X
- Media files with validation of specific mediaIdentifier attributes

What assumptions can I make?

- In the database, validation with any given attributes will be for finding exact matches and not similar data.

- Understanding the project description I assume that editing of database attributes column data is not allowed.

What constraints exist?

None

Are there strange cases to handle?

- There are many negative scenarios which need to be handled either with proper return values or by throwing exceptions. For example:
 - Adding duplicate person Identity
 - Adding attributes for a person, whose data already known to database
 - Fetching media files for person whose data is unknown to database
 - Fetching relationship data whose details are unknown to database
 - And similar negative scenarios.

Next steps

1. Understand and create Database Schema
2. How Family tree can be created using relationship data stored in database
3. Class Diagram and data Structure

References:

- [1] "Cousin Chart—Family Relationships Explained." *FamilySearch Blog*, 23 July 2019, www.familysearch.org/en/blog/cousin-chart-family-relationships-explained. Accessed 4 Nov. 2021.