CSCI 3901: Course Project

Milestone 1: Breakdown Analysis of the problem

Inputs for the System:

To generate Family Tree Database:

- Name of the individual.
- Attributes for the individual: Date of Birth, Gender, Occupation, etc.
- Identifier of an individual to store their Reference or Note.
- Identifiers for two individuals to record a relation between them.
- Type of relation between two individuals.

To generate Media Archive:

- Location/name of a media file to add to the archive.
- An identifier for the media file and its attributes: Year, Location, Tags, etc.
- List of individuals that appear in a media file.
- Tags for a media file.

For Reporting:

- Name of an individual and location/name of a media file to locate them.
- Identifier of a person and media file to fetch their names.
- Number of Generations in the past or future for an individual to find their ancestors or descendants respectively.
- Tags and locations to filter a media file.
- Set of identifiers of people to find their corresponding media files.
- A date range specified by a start and end date to filter content.

Transformations for the input data:

- Attributes for a person (Date of Birth/Death, Gender, Occupation, etc.) must be converted to key-value pairs in a Map. The key is the type of attribute as a String. However, the value must be converted to an appropriate format based on its type.
- The range of dates is provided in the form of Start and End dates as Strings. They must be converted to their corresponding Date types.
- Convert the data of people and the relationships between them into a Biological Tree.
- Connect each individual with their corresponding media files in the media archive.

Data that is processed right away:

For recording information of a person or media file in the database, only two mandatory fields are required: name of the person and filename respectively. So, this data must be recorded immediately.

Long-Term Data:

All the input data mentioned above contribute to the record of a person or media file. This information is stored in the database for future access.

Tasks to be achieved with this data:

- Maintain a record of details of people and media files.
- Record the relationship between pairs of people. Could be parent/child relation, partnership, or dissolution of the partnership.
- Find the corresponding images of the people and their ancestors or descendants, based on a time range.
- Find how a pair of individuals are related and their corresponding images, based on a time range.
- Filter images based on locations or tags.
- Fetch personal details of a person.
- Fetch the ancestors or descendants of a person based on the number of generations provided.

Data Structures that can be used:

- A Map for storing the attributes for a person or media file.
- A Tree for generating the Biological Tree. The tree structure must be stored in the database in some way.
- A List or Set for storing person identifiers and media file identifiers.
- A List for storing notes and references for an individual.

Output from the System:

- Show how a pair of individuals are related.
- Find details about an individual.
- List the descendants and ancestors of an individual for a given number of generations.
- List images based on tags or locations for a given time range.
- List images corresponding to the provided individual, a list of individuals or family members of the individual, based on a time range.

Assumptions:

As mentioned in the assignment text, we are assuming that an individual or a media file will not have all the information mentioned above in the input section of the document. Only the name of a person and filename for a media file is required to store them in a database. Also, the dates can have the following information: Only Year, Only Year and Month, or Full Date.

Constraints:

As mentioned in the assignment text, the system will only deal with biological family relations. Also, the family relations defined will be based on English genealogy as they have precise definitions for them.

Purpose of the system:

The system must accurately provide the information to the genealogists about how a pair of individuals are related, list the ancestors or descendants for an individual as well as fetch the corresponding images of the family members. The tracking will be done based on various parameters supplied.

Users of the system:

The users are genealogists and will help them in tracking the family relations for a person as well as get the images for those family members.

Target Environment:

The system will be used as a Command Line Interface Program. The input will be provided through the keyboard and the output will be displayed in the terminal. As it will be programmed in Java, it will be able to run on any operating system.

Queries about the problem:

- The biological tree only depicts the relation of a child with one of the parents. So, it must be made clear that the ancestors we are tracking are based either on the mother's or father's side.
- A proper method must be found to store the tree structure in a database that can be easily accessed.