Trentoniana Room Database | Group 9

Michael Mongelli, Ally Rizzo, Ashley Bennett, Zack Rich

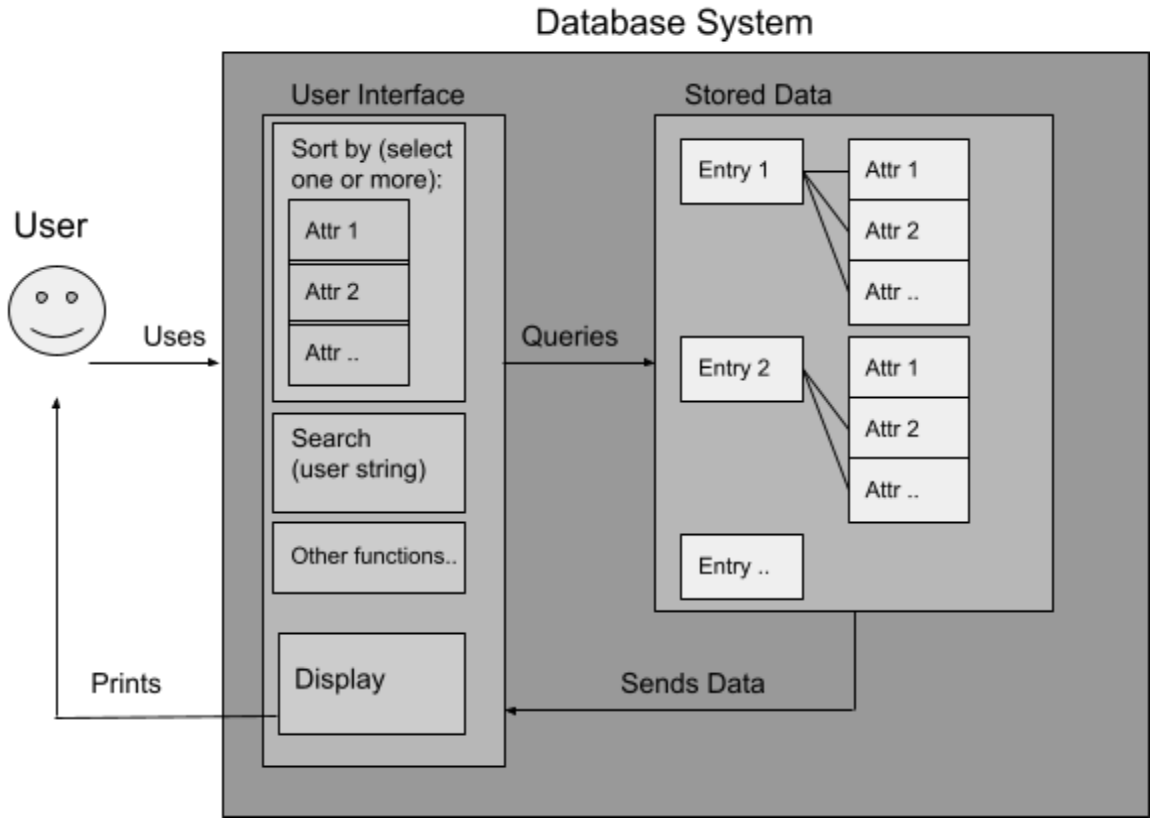
- **Problem Statement:** The Trentoniana Room audio archive has recently acquired a number of audio recordings from residents of Trenton capturing an oral history of the local area, however all the audio files are all saved in one directory and presently have little descriptive information for one to quickly and easily locate the file they are looking for.
- **Objective of the module:** To create a more organized storage of these files and an easier way to navigate them. We will accomplish this by categorizing the files and providing the user an easily-usable interface to search the files. Currently, the files are in categories which are uninformative and redundant (e.g. many files fall into the categories *NJ*, *New Jersey*, and *Trenton*). We will recategorize them by name, date, location of recording, date archived, topic, participant name (if provided), and audio file length, as well as a way to search the content of the transcriptions which will be provided for these files. The new categories to describe the entries will be the following: Jewish history, merchant, oral history, family history, deli, bakery, business, old Trenton, East state street, Market street, arts, family business, immigration.
- **Description of end product:** By the end of the project we aim to have a functional user interface to interact with the database. We will provide the user with the ability to sort the files by category using drop-down menus and to search the files using a textbox to query the files and their attributes for

matching/relevant strings. We may also provide the ability to perform searches based on multiple criteria at once if possible.

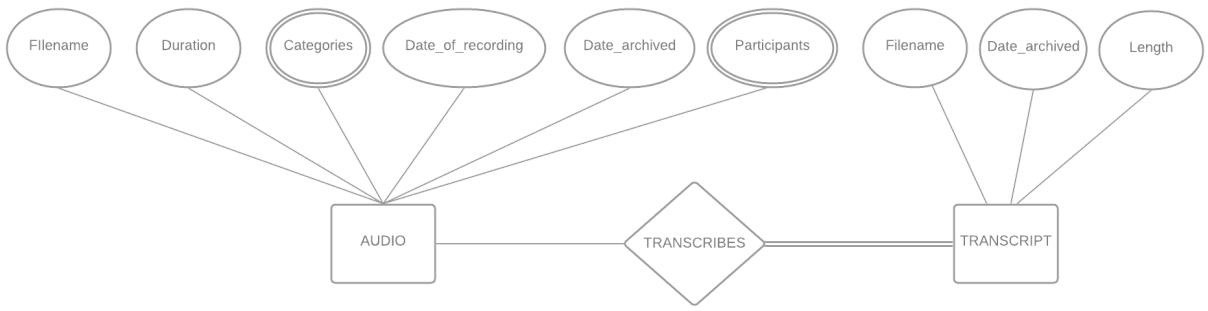
- **Importance of module:** These recordings offer a rich history of the local area, however the files are more difficult to interpret at present because they lack context and are listed in an arbitrary order. By categorizing the files and providing an interface specifically for this problem, users will find it easier to learn from and make sense of the recordings. Thus, this module will help the recordings achieve their full educational potential by providing a more accessible way to view them.
- **Researching problem domain:** Having a linguistics student on the team will help us in our design of the database because they understand the context of these recordings. This will help us to understand how to categorize the recordings based on what will be relevant to the others in the linguistics department. We will also be doing a lot of research on our own while implementing the system, using the textbook and the internet, so we can deliver the best system possible.
- **Similar systems:** many libraries and other sites use similar database systems as we will create for this project, thus we will look at these systems for guidance and inspiration for ours.
- **Other applications of the system:** The system could be utilized by similar those with similar needs, for example, local libraries which may still have not updated their websites to be very user friendly yet. This system may also be useful in other settings, for example, possibly to set up an online menu for a restaurant's site.

- **Performance:** Since there are a small number of entries (less than 100) we should expect extremely fast performance from the system. We will time the operations performed by the database using the system clock to ensure we reach performance we find acceptable. We will keep efficiency and speed in mind while also preparing the database to maintain similar performance even if the number of files increases.
- **Security:** We will try to ensure our system is not vulnerable to common and easily avoidable attacks (such as SQL injections for example). We will research common online security threats and attempt to use them on our system to test its ability to defend against these attacks.
- **Backup and recovery:** We will prepare our system for a full recovery of its data, and we plan for it to update frequently so that files will be backed up as soon as they are uploaded, which will ensure no loss of data in the event that the data needs to be restored if it is corrupted.
- **Technologies and database concepts:** Our team will need to learn how to set up a database, how to upload the data, add attributes and form relationships between the data, how to store the data and make it available for access, how to back the data up, and how to protect it. We will use resources from class as well as online resources to master these concepts so we can use them to build our application and will work as a team to ensure we all understand the tools we use to build this database and its UI.

System boundary diagram:



Database Entity Relationship Diagram:



Relational Schema:

AUDIO

filename	duration	categories	date_of_recording	date_archived	participants

TRANSCRIPT

filename	date_archived	length

The database is estimated to be a size of 74 entries (74 audio recordings, each of which may or may not have a corresponding transcript). The linguistics department estimates it will only add a few more recordings (less than ten) after these initial entries, making the size of the database relatively small and static.

The linguistics members also specified that their searches of the database will most likely be searching the files based on their assigned categories, meaning the most common search type will most likely be done using a drop-down menu selection. There will most likely be a small number of searches to the database at a time, meaning that other than the time to access the files in the database, performance shouldn't be an issue with regards to handling large volumes of queries.

Trentoniana Room Oral Histories Database System - Group 9

Michael Mongelli, Ally Rizzo, Ashley Bennett, Zack Rich

Need	**Approach**
The Trenton Library wishes to have a system to better store and organize their	We will create a model of the database we wish to implement as a team, planning

new dataset containing audio recordings from local Trentonites.	out the interface and what functions we want the database to be able to perform. We will then build out the system as a group by allocating tasks and adding features incrementally. We will also be doing research along the way to ensure that we are building the system using all the expertise the field has to offer us as we learn.
Benefit	**Competition**
The Trenton Library will have a useful tool to give users an easier way to navigate and make sense of this dataset, which will be beneficial to the linguistics department members transcribing and analyzing the recordings as well as normal library-goers wanting to learn from these oral histories. Thus, the system will provide a window into Trenton's history to both academics and to those simply wishing to learn about local history.	Our system will make the data easy to navigate by providing many categories to sort the data, as well as the ability to sort the data by multiple criteria and search by user-entered strings. In addition to a visually pleasing and intuitive UI, the database will store its data safely both by having a strong defense to cyber attacks as well as a backing store in the event of the need of data recovery. The system will also be fast as we wish for there to be only microsecond delay between query

	and output. Between our intuitive UI, informative data attributes, secure data storage, and fast performance, we offer a competitive database system to the Trenton Library.