Project Name: Audio-Transcript Database Implementation
Team Name: Group 3 (Internet Explorers)
Lalima Bhola, Aly Maahs, Summer Martin, Kyla Ramos, Jared Schmidt
10 February 2020
Dr. DeGood

- Problem statement.

The Audio & Visual section of the Tentoniana website is currently a work in progress. There is a way for people to access the audio files and our classes have the transcripts of the audio files but there are several problems with the current system. To start, the audio files are not easily accessible. They are hosted on a page that isn't user friendly and they are not easy to navigate to even when you are looking for them. The audio files are also not linked to the transcripts in any way and the transcripts are not currently accessible to the public. Furthermore, neither the audio files, nor the transcripts, are organized and there is no way to search for specific topics, dates or names. A few of the audio files are titled with names, but there is no other identifying information about the topic and several appear to only have a number. The files are also not tagged or labeled in any way that would make searching easier. Additionally, currently only having audio files makes it difficult or impossible for those with hearing impairments or those with auditory processing and other disabilities to use the files. Finally, there is no way for people to submit suggestions or concerns for newly discovered files, changes or errors.

- Objective of the module.

The objective of our project is to create and provide a more efficient and user friendly system for searching and finding audio files on the topics the user is looking for. It is also our goal to combine the audio field and the transcripts so they are both accessible to the public and have the ability to view the transcript that corresponds to the audio file while the audio is playing. We want to provide a user friendly interface for searching, and viewing these files. Providing an straightforward way to interact with these files and a system that would allow for public users to submit new files to be transcribed would encourage interaction between the user and those behind the scenes to collaborate and improve upon existing systems. Better user interaction makes the user feel like their suggestions are valued and their interaction with our system is important. We also want those with higher access to the database to be able to improve and make changes, like adding relevant tags, information and new files, to the system more easily.

- Description of the desired end product, and the part you will develop for this class.

The audio-transcript files would be sorted and arranged in a neat, elegant user friendly user interface and would be tagged or labeled with topics, names, dates, and other relevant information so they would be able to be searched for more easily, made easy through tagging the

files with relevant information. There would be a search bar at the top of the page so users could easily search for their desired topic or a specific year. If selected, the user would be taken to a different page where they could listen to the audio file and have the option to view the transcript. When viewing the transcript, the user would see the entire transcript at once, maybe with the current part highlighted. On the main page there would be a link to the transcript for each audio file in case the user would rather only view the transcript. There would be some feature on the main page which would be a submission box for new audio files, records, or suggestions. For example, if someone found an error in the transcript or wanted to make a suggestion for improvement or a new submission to be transcribed.

- Description of the importance and need for the module, and how it addresses the problem.

Since the current system only contains the audio recordings, we want to add the transcripts to make information more easily accessible. A user may not want to listen to the entire audio file to find information so having a transcript will be helpful for that reason. For other users, it may be difficult to only listen to an audio file and a transcript will help them follow along better and hearing impared users, listening to the audio file may be impossible or extremely difficult. Our module will allow users to view a transcript as an audio file plays. The module can also include a search function to make it easy to find a specific audio file since, at the moment, there is no way to know what is contained in an audio file until you listen to it. This application addresses the issues of user inaccessibility. We also want to make it visually appealing through our user interface so people want to explore the history of trenton and promote the website itself. Our program will make the files easily accessible, searchable, user friendly, and better organized.

- Plan for how you will research the problem domain and obtain the data needed.

In order to research our problem domain and obtain the needed data we will examine similar applications of audio file use, accessibility and the correlating transcripts to see how these methods work, which work best and how we could improve on existing ones. We will see if some platforms ever have performance issues and do research as to why these issues occur and do our best to avoid them. We will also research other websites that host historical transcripts and see how they present information. We may look at other libraries' websites to see how their user interfaces are set up and presented for public use and make sure we allow for similar ease of access and user friendly methods for viewing historical files and transcripts.

- Other similar systems / approaches that exist, and how your module is different or will add to the existing system.

In sites such as ted.com, they have the audio playing while the corresponding transcript appears below the video. Our site will have a similar feature where users will be able to follow along to the transcript as the audio file plays. They would also be able to click a specific instance

of the transcript, which would take them to that moment in the audio. Our module can also include a feature for displaying views, likes, and comments for the audio files and transcripts. By allowing users to like and comment on data, over time, the searches might be able to be filtered by "popular" or "top" files. The comment section would allow for debate or conversation on varied topics. This would create a better sense of community within the Trenton area and users of the Trentoniana site.

- Possible other applications of the system (how it could be modified and reused.)

Since our module will allow for greater user accessibility, it can be applied to other locations where there is a lack of inclusivity in the distribution of information. For example, sites containing audio/video files without subtitles or transcripts will limit those who are hard of hearing or deaf from proper access to this data. The module will also contain a clear user interface/place to enter searches based on categories. Once implemented, the module can transfer to similar sites to allow for a comprehensive user base. Our module could be modified for any set of data or files with need of an organized, user friendly system.

- Performance – specify how and to what extent you will address this.

In order to optimize performance, we will create precise methods of querying and sorting data. For example, audio files should have the option of being sorted by length as well as content. Transcripts will have concise queries and filters/tags to allow users to find their data more accurately. We can also allow users the option of filtering data by exclusion. This can include excluding specific years or keywords from their search. Another option, instead of a direct search box, is to designate fields for an individual's first name, last name, year, length of video, etc. This would also assist in creating an accessible user interface. We will ensure our module has a reasonable performance level to guarantee an optimal user experience.

- Security – specify how and to what extent you will provide security features.

To provide security features for this application, we would set up a User ID and password function for the managers of the system to give them access to more advanced and private features of the database. These would be stored in the database if or when a user chooses to make an account. To provide security for the user, the passwords themselves would be encrypted so they are not stored in a plain text form (which would be a large security risk should that information be leaked in some way).

We would implement three types of user for the application: basic user, moderator, and administrator. (However, the only two users that actually require the use of a user ID and password are the moderator and administrator.) This basic user would have the least amount of power, and not be able to do more advanced things such as directly edit or delete a transcript. They would however have access to the general functions of the application such as suggesting an edit to a transcript.

The second tier of account would function as a moderator. These have more power than a basic user, but still do not have as much as the administrator of the app. They would have the ability to do more advanced things such as review any edits made to transcripts and either accept or deny them, as well as moderate a basic user's activity in the event they are being harmful towards others on the site. However, they wouldn't have direct access to the database unless given permission to by the administrator of the system.

This administrator would be the highest tier of user. They have all the permissions of the previous tiers, as well as having direct access to all database information. They can also create moderator accounts for other users to utilize, and can choose whether or not those moderators will have access to the database as well.

- Backup and recovery – specify how and to what extent you will implement this.

We will make sure that there are private backup copies of all the files inaccessible to regular users in case of file corruption or deletion. Since they are backed up, they can be recovered and reuploaded by those with higher access. New files will be automatically copied to a secure location which would not be accessible to the public and normal users. In the case of file corruption, which would have to be detected by a user and noted by a system manager or high level user, the copied file would have to manually be uploaded to the site.
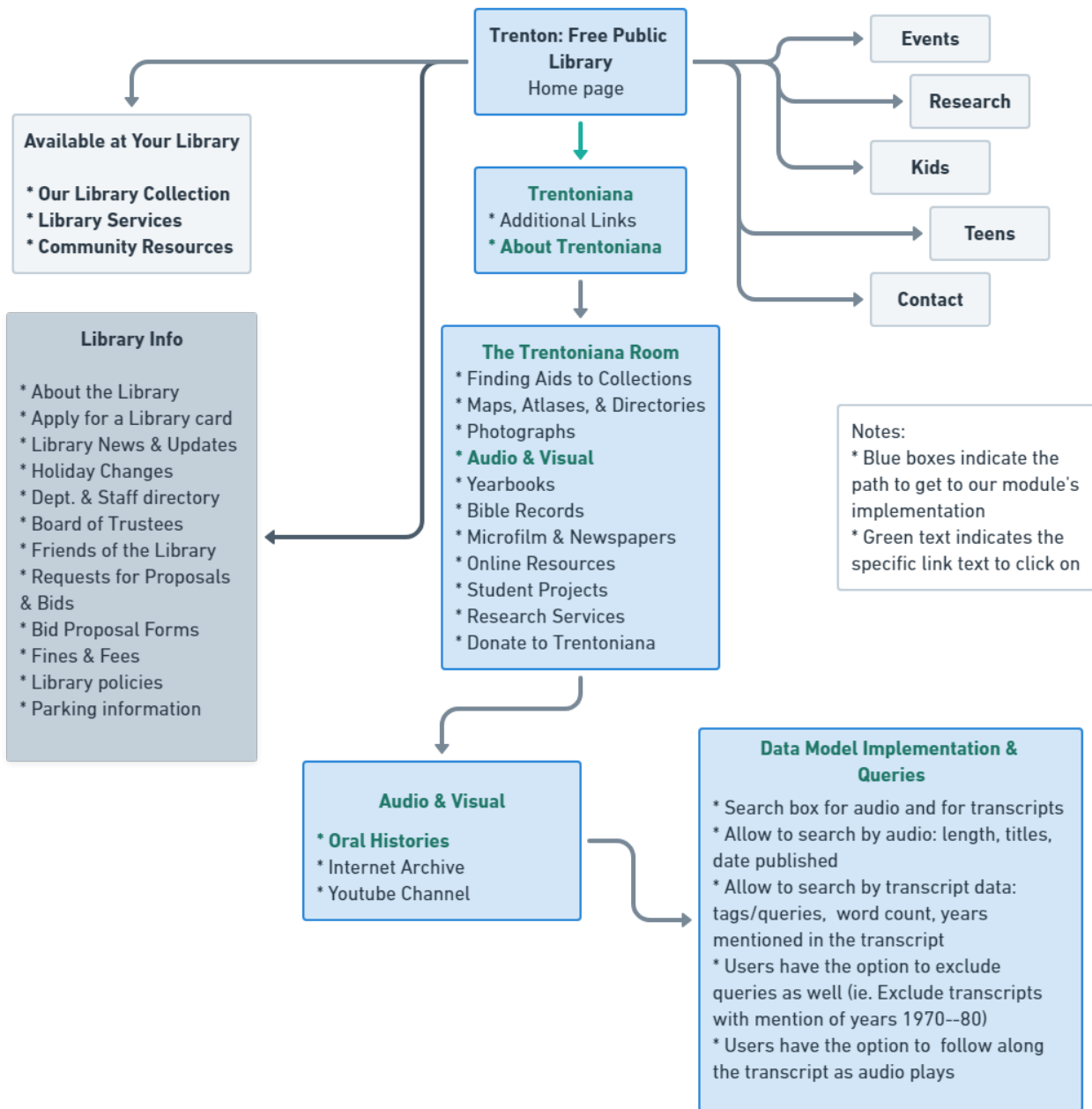
- Technologies and database concepts the team will need to learn, and a plan for learning these.

We will need to utilize PostgreSQL to create the database. In creating the web application for the project, we may also need to utilize PHP to link our database to the website. We will also utilize Figma to create prototype designs of what we want our final project to look like, and following that utilize HTML and CSS to carry out the look and feel we finalized for our web UI.

To accomplish our design goals, we will use online tutorials as well as the prior knowledge or experience of members on our team. A member of our team is currently learning about web design in a class and another completed a class on web design last semester as well as having prior design experience. The rest of the group members can learn from these two team members as well as from any online resources such as w3schools to carry this out.

There are several database concepts our team will need to learn in order to properly implement our program. Some of these concepts include what it means to put something into a database, how to implement security measures in a database, how to backup a database and its contents and make sure they are recoverable. We would also need to learn what the structure of a database is and how to implement one in code.

- A diagrammatic representation of the system boundary that specifies what data you will model and which queries you will implement.

Trenton: Free Public Library
Home page

Events

Research

Kids

Teens

Contact

Available at Your Library

* Our Library Collection
* Library Services
* Community Resources

Trentoniana
* Additional Links
* About Trentoniana

Library Info

* About the Library
* Apply for a Library card
* Library News & Updates
* Holiday Changes
* Dept. & Staff directory
* Board of Trustees
* Friends of the Library
* Requests for Proposals
& Bids
* Bid Proposal Forms
* Fines & Fees
* Library policies
* Parking information

The Trentoniana Room
* Finding Aids to Collections
* Maps, Atlases, & Directories
* Photographs
* Audio & Visual
* Yearbooks
* Bible Records
* Microfilm & Newspapers
* Online Resources
* Student Projects
* Research Services
* Donate to Trentoniana

Notes:
* Blue boxes indicate the
path to get to our module's
implementation
* Green text indicates the
specific link text to click on

Audio & Visual

* Oral Histories
* Internet Archive
* Youtube Channel

**Data Model Implementation &
Queries**

* Search box for audio and for transcripts
* Allow to search by audio: length, titles,
date published
* Allow to search by transcript data:
tags/queries, word count, years
mentioned in the transcript
* Users have the option to exclude
queries as well (ie. Exclude transcripts
with mention of years 1970--80)
* Users have the option to follow along
the transcript as audio plays

- 1-page quad chart; see: Quad_instructions_template.ppt in the Canvas files section

Audio-Transcript Database Implementation

Group 3: Lalima Bhola, Aly Maahs, Summer Martin, Kyla Ramos, Jared Schmidt

TCNJ THE COLLEGE OF NEW JERSEY

Need

- A better way to access the audio files from Trentoniana website as well as access to the transcripts for the audio files.
- Customers with audio and hearing disabilities needs include transcripts for those.
- An organized and sorted view of all the audio files with a search capability.
- User friendly interface with better accessibility.

Approach

- Have the audio files arranged in a unique, elegant user interface and have the audio files labelled and tagged based on topic
- Have a search bar so the users can search for a specific topic or a specific date
- An option for the user to view the transcript by itself or with the audio, potentially highlighting the text to match along with the audio
- Give the user the ability to recommend new audio files to be transcribed or changes to be made

Benefit

- Having a database for these Trentoniana audio files that is much better than the current Internet Archive
- Making the transcripts available to the public
- The program will support those with hearing and auditory processing disabilities by including the transcripts making our system inclusive and comprehensive.

Competition

- Make our database look better and be easier to use than other implementations
- Will be much better than the current Internet Archive's implementation as ours will have the transcripts available and have the audio files sorted, as well as having a search feature (where users can search by audio and/or transcript data)

02/10/21