Matthew Edwards – me4bb

Colin Warner – ynq9ya

**Progress Report #1**

One Shared Story’s mission is: “to expand what we know about the past so we can better understand the present and work together to create compassionate, inclusive communities for the future.” They do this by encouraging research to reveal hidden histories of individuals and communities, provide support to amplify underrepresented stories, and foster dialogue within communities, and the general public to expand our common narrative. They have been around for a few years and have many spreadsheets of data consisting of transcribed historical records. They contacted UVA in need of some help with their data. Through multiple meetings we have come up with three objectives: Standardize the way volunteers transcribe data, store transcribed data in a database, create an interactive search tool.

The data is the most important part of this, or any other, project. One Shared Story has many excel sheets with data transcribed from historical documents from the local area. These have differing schemas depending on what the document is: wedding records, personal property tax info, slave laborer records, etc. One Shared Story has shared the following four variables that will have some info in every dataset: Person, Event, Source, and Location. There will be more data depending on what was being recorded, but these will be our starting point as we work with the data.

One issue that One Shared Story has run into is the fact that they have many different types of datasets they are transcribing. They search out historical documents, gain access to them, and then transcribe the data from them. Each document tracks something different and therefore will have different data types from other documents. The volunteers transcribing this data currently do so on excel sheets. The first thing we need to do is create a web portal that standardizes the way volunteers gather data. This web portal will also lead to a uniform storage location for the data.

One question we need to answer as we get more of the data is what type of database is going to be best. With the differing types of data, and the fact that each new dataset could potentially be new and different as well, we are going to approach this problem with a NoSQL type database like MongoDB. This will allow for more flexibility as data is stored, and will give volunteers a chance to transcribe all sorts of records in the future.

One Shared Story often visits churches on the weekend to allow individuals to learn about family ancestry. This is where the third objective comes in for our group. They need an interactive search tool to help this process go smoothly for new users. The goal for this tool would be to allow these families to search for their ancestors and view information as it relates to the sources, events, people, and roles that are defined in our data model.

We have built a first pass at a database and transcribing tool. This was built off the schema of two datasets sent by One Shared Story. We presented this tool to One Shared Story and they were very impressed and pleased with the initial step. Future work will be to decide what type of database will be best for all the data One Shared Story has. Once the database type has been decided we will be able to incorporate our database structure into the transcribing tool. We are excited to learn more about what types of databases are being used by others in the field, and have an email to set up a call with other leaders in the space that will help us decide on database type and set up.