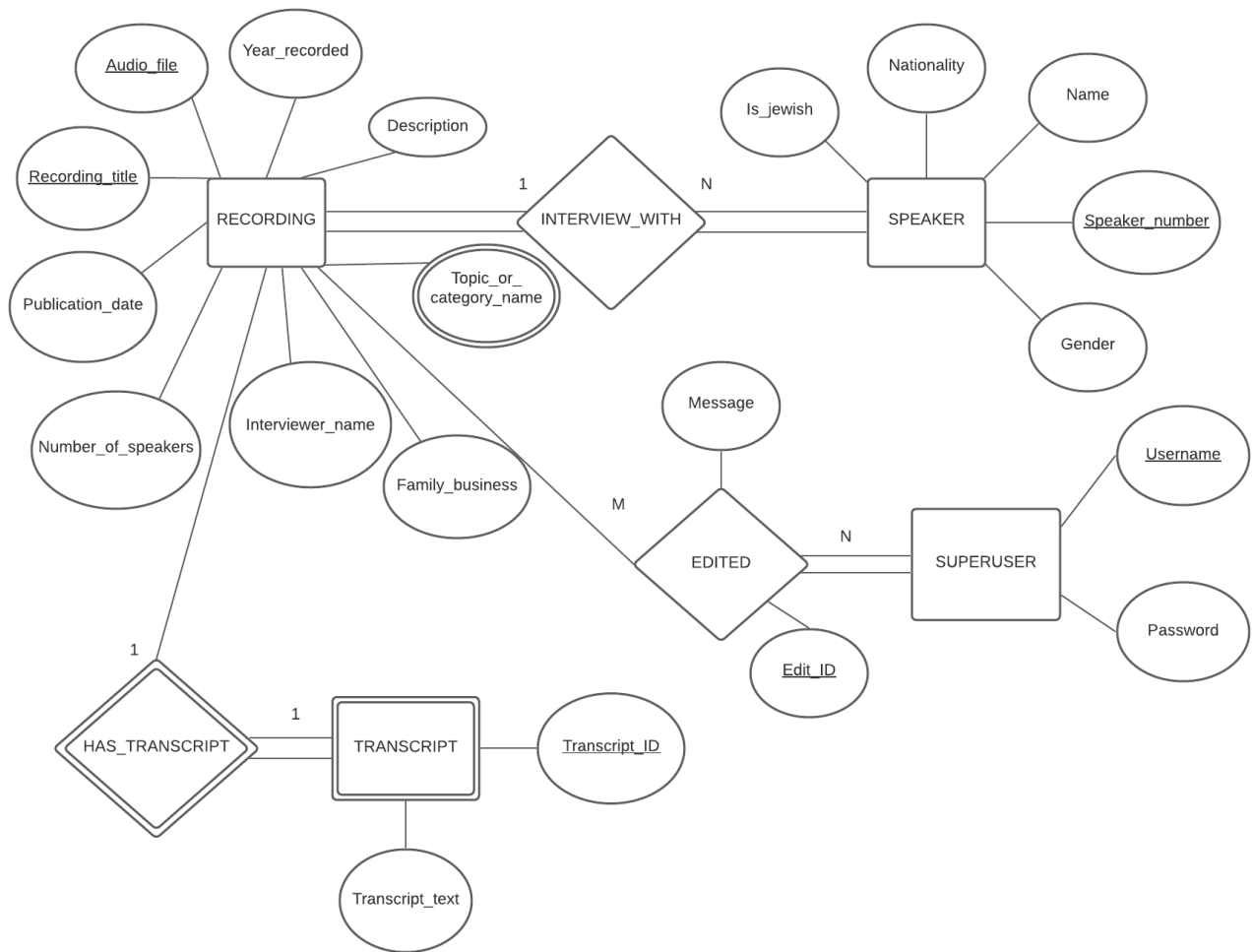Lakaylah Gage, Jabili Gadde, Anthony Scarpa, Nicole Kercado (Group 2)

CSC315-01
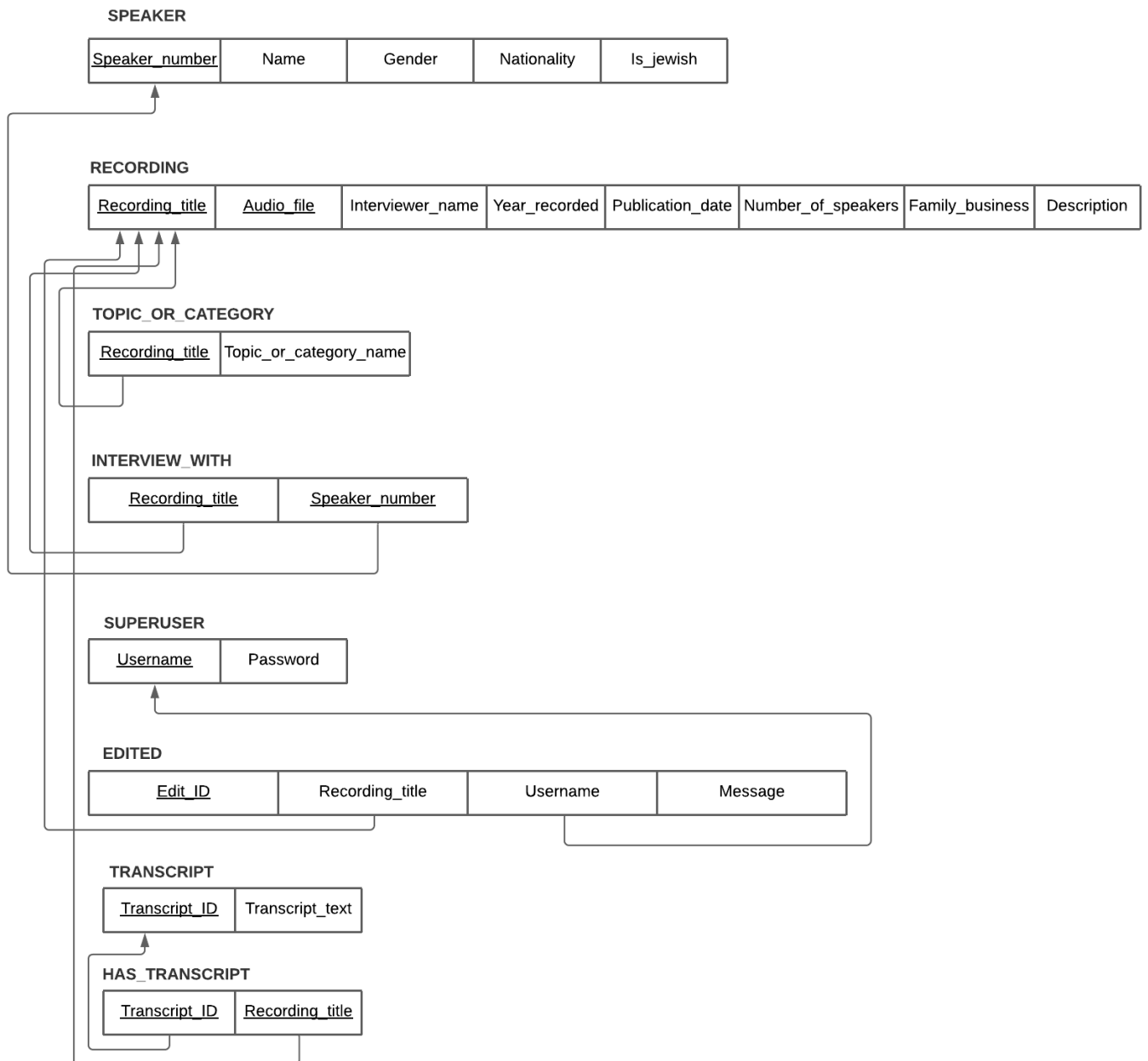
Stage III: Database Model

Link to Project Repository: https://github.com/TCNJ-degoodj/stage-iii-group-2

ER Diagram:

ER Diagram Mapping to Relational Schema:

SPEAKER

Speaker_number	Name	Gender	Nationality	Is_jewish

RECORDING

Recording_title	Audio_file	Interviewer_name	Year_recorded	Publication_date	Number_of_speakers	Family_business	Description

TOPIC_OR_CATEGORY

Recording_title	Topic_or_category_name

INTERVIEW_WITH

Recording_title	Speaker_number

SUPERUSER

Username	Password

EDITED

Edit_ID	Recording_title	Username	Message

TRANSCRIPT

Transcript_ID	Transcript_text

HAS_TRANSCRIPT

Transcript_ID	Recording_title

Calculating the Average Amount of Searches Per Day:

The total number of views for all the recordings combined was 2,215 (from adding up the views of each recording on the Trentoniana site). The collection of the recordings itself had been up for roughly 1,389 days (the site states that the collection of recordings was published on May 23, 2017), meaning that the average amount of searches per day was 2,215 / 1,389, or 1.59 searches per day.

Approximating the Number of Records (Initial Database Size)

There are currently 74 recordings on the site. This means there will be 74 tuples for the RECORDING relation. If we assume that each recording has on average 2 speakers, that means there will be 74 * 2 tuples for the SPEAKER relation (148 tuples). This also means there will be 148 tuples for the INTERVIEW_WITH relationship. If there are 2 topics per recording, there will be 148 tuples for the Topic_or_category relation. If we assume that there are 5-10 super-users who can essentially edit/insert/delete recordings, then we can approximate that there would be 8 tuples for the SUPERUSER relation. We can assume that a super-user would be able to edit/insert/delete every recording, so there would be 74 * 8 tuples for the EDITED relationship. If we assume that there are 24 recordings with transcripts, then that there would be 24 tuples in the TRANSCRIPT relation and 24 tuples for the HAS_TRANSCRIPT relationship.

So to approximate the number of records, 74 + 148 + 148 + 148 + 8 + 592 + 24 + 24 = 1166 records.

Types of Searches:
- Metadata - The data of our data will be searched through.
- Text contents - Any text in the site will be searched through, including any transcripts and descriptions.
- Limit options - On the site, there will be filters available for the user to utilize to narrow down the search. They can filter by topic/category, nationality, gender, and more.