INST327

Section 0202

Project Final Report

12/12/22

Group 5

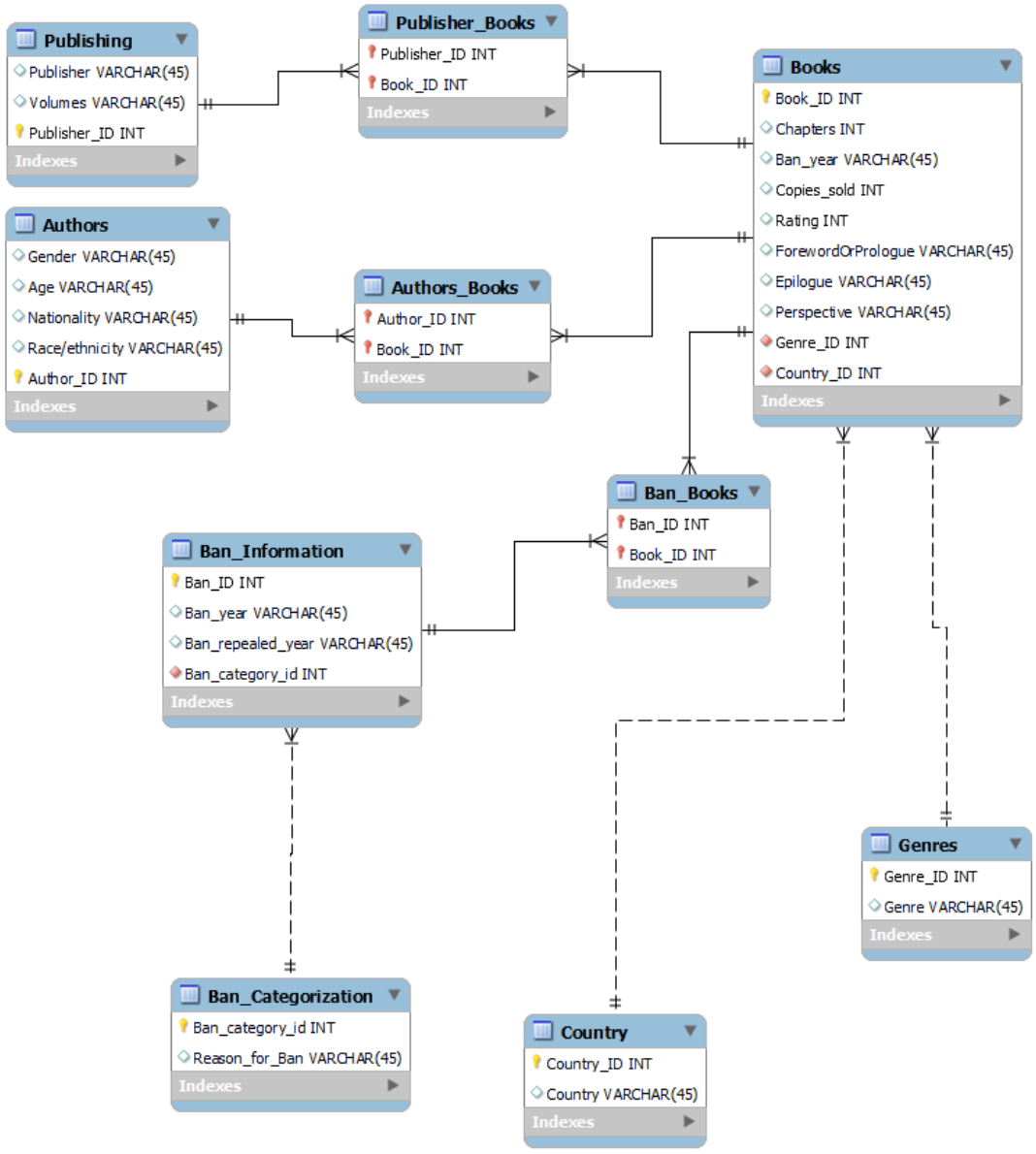
Tyler Dulin, Joe Puchalla, Tobi Arewa, Will Heffernan, Kenny Tran

Introduction

The topic we chose to work on for our group project is a database of books that were banned in Nazi Germany during World War II. Our database encompasses a large variety of genres, authors, countries of origin, and more. We are specifically targeting the time period just before World War II, specifically 1933, when the first reported book burning occurred, until the end of the Nazi regime in 1945. This is a twelve-year timespan in which we cataloged a small subset of the books that were burned. We selected this topic as it was interesting to all of us, in addition to the historical significance of this topic. It is fascinating to consider what dialogue and discussion could be so jarring as to receive a ban. We also felt it was important to delve into the history of the thought process and culture of the Nazi regime and examine why these books in particular were considered dangerous in their eyes.

Our database consists of ten tables: books, authors, publishing, genres, ban information, ban categorization, authors_books, publisher_books, ban_books, and country. Our tables are connected by one-to-many relationships, but the main parent table in our database is the books table as it holds most of the database together. This database will be useful to anyone interested in World War II and its aspects.

Logical Design



Physical Database

Our database consists of 10 tables. The main table will have to be the Book table. This table references Publisher_book, Authors_book, Ban_Books, Country, and Genres. Publisher_book, Authors_book, and Ban_Books are all linking which connect the main Book table to the Publishers, Authors, and Ban_Information tables, where the information about those aspects can be found. Ban_Information is linked to Ban_Categorization, which is specifically for the rationale behind the ban.

Sample Data

The data used in our database is sourced from goodreads (https://www.goodreads.com/list/show/101688.Books_Banned_by_Nazis). This website had all the historical information we needed to compile our database, specifically, the book titles, authors. We used the book titles and authors to remain historically accurate in creating the tables. The rest of the data we created ourselves, so we attempted to keep everything feasible for the time period, such as the ban years of the books falling between 1933 – 1945. The author's country of birth was chosen from the United States, whose free speech policy directly opposed that of the Nazis, in addition to many Western European countries. Other data such as chapters, reason for ban, author age, and all remaining columns were carefully considered and data was added that would be appropriate for each.

Views / Queries

Query	A	B	C	D	E
1	X	X		X	
2	X	X	X		
3	X	X		X	X
4	X	X		X	
5	X		X		

Query 1: Creates a view that displays the books that sold more than 100,000 copies, ordered by the year they were banned.
Query 2: Creates a view that returns only books that were published in Austria with columns titled “Country of Origin”, “Number of Books” and “Average Number of Chapters”.
Query 3: Creates a view that returns books that had 1000 or more volumes published.
Query 4: Creates a view that returns books that were banned from 1939-1945, ordered by the year they were banned and the reason for ban.
Query 5: Creates a view that counts the number of banned books that each author has, ordered by the number of banned books.

Changes from original design

At the time of our project proposal, we included a table for data on topics of each book which we had decided to exclude for its lack of unique contribution to the substance of our database. Additionally we removed the “country of origin” and “genre” attributes from our books table and moved these attributes to tables of their own in order to keep the data in the books table more succinct. The countries table that we ended up creating also meshed well with our authors table as it allowed us to link the two by author nationality as well. Since our proposal, we also changed our database conceptually by deciding to focus our books table more on physical aspects of the books themselves and leave the historical and social aspects to the other tables such as author information and ban information.

Database Ethics Considerations

We as a team have stayed attentive to how we observe and manage our database in considerations to ethics. There are possible procedures we have discovered throughout the semester that could potentially not follow up to ethics. We found that our main ethical concern going into this project would be potential copyright and fair use concerns, since the database includes information about the author and their intellectual property. To combat this ethics issue, we have only collected data from authors who have their work made out to the public, rather than

acquiring data from any proprietary or “closed” data sources. The information we collected for our database does not infringe any of the author's work but describes the physical characteristics of the book itself.

Another ethical consideration arose when importing data into our database. With such a small sample of the thousands of books and their authors who were censored in Nazi Germany, there is a clear underrepresentation of authors belonging to minority groups, particularly by gender and race or ethnicity.

Lessons Learned

Throughout this semester we faced many challenges, but tackled each of them through communication amongst ourselves and the instructional team. One main challenge we faced in the beginning was determining which topic to work with. Originally we chose to develop a database that would track the COVID-19 cases in each Maryland county, but we were prompted to change this topic due to its simplicity and unoriginality. After receiving this news we were quite discouraged, mainly because we felt like we were on the right track and everything was moving smoothly, but this discouragement didn’t last for long! Tyler soon proposed the idea of creating a database on books banned during World War II, and since this was a topic that interested all of us, we decided to move on with it. We also had to make some modifications to this topic along the way, with some suggestions from TAs to merge unnecessary tables into the Books table, add several linking tables, ensure every table has a primary and foreign key, etc. Another challenge we faced was with our constraints. When developing our database we discovered that our authors were only assigned to one book each, which was a major problem. After debugging our database we discovered that the constraints for the columns in our “books_authors” table were set to “UNIQUE”, which caused the issue. During this project we experienced many roadblocks but got over them through teamwork, constant revision, and plenty of help from the instructional team.

Potential Future Work

The main thing we'd work on in the future in regards to the database would be expansion. As it stands, it is nowhere near an exhaustive list of books that would qualify to be added (the Nazis banned *a lot* of books) and there may be some blind spots as we focused on some of the more popular ones. This obviously impacts our diversity, equity and inclusion considerations, and while we think the spread of authors and books gives a good representation of what types of things the Nazis would ban a book for, fleshing out the database to include a more complete list of books would be good. This would both help flesh out those requirements and make our different views more complete, as right now in our queries when you look up how many books H.G. Wells had banned for example it probably gives a different number than what the actual amount of his books banned was, which is inherently a misrepresentation. The main other thing would be ensuring the accuracy of the data, for things like publishers we changed some of the data to ensure that project requirements were met, where in reality finding 15 books with 15 different publishers would be tricky, and would not give an accurate picture for what a fully fleshed out database would look like.