

Introduction

A library has thousands of books that all people use for their own different purposes. Whether it is a student researching a project, a teacher finding information about a certain topic, or any person looking for a book to keep them occupied, a library is a massive house of knowledge that can provide entertainment or solutions for anyone that is looking to do so.

The database we have assembled contains different resources in McKeldin Library and their online database, WorldCat. We decided to do this because we realize that a library is a center of knowledge, and many people on the college campus utilize the library in order to study and find books and other resources that can aid them in performing better in their studies. We also have interest in reading, and we realize that a library has thousands of books that can be used for personal entertainment, and so it is evident that creating a database for a library will be of great use.

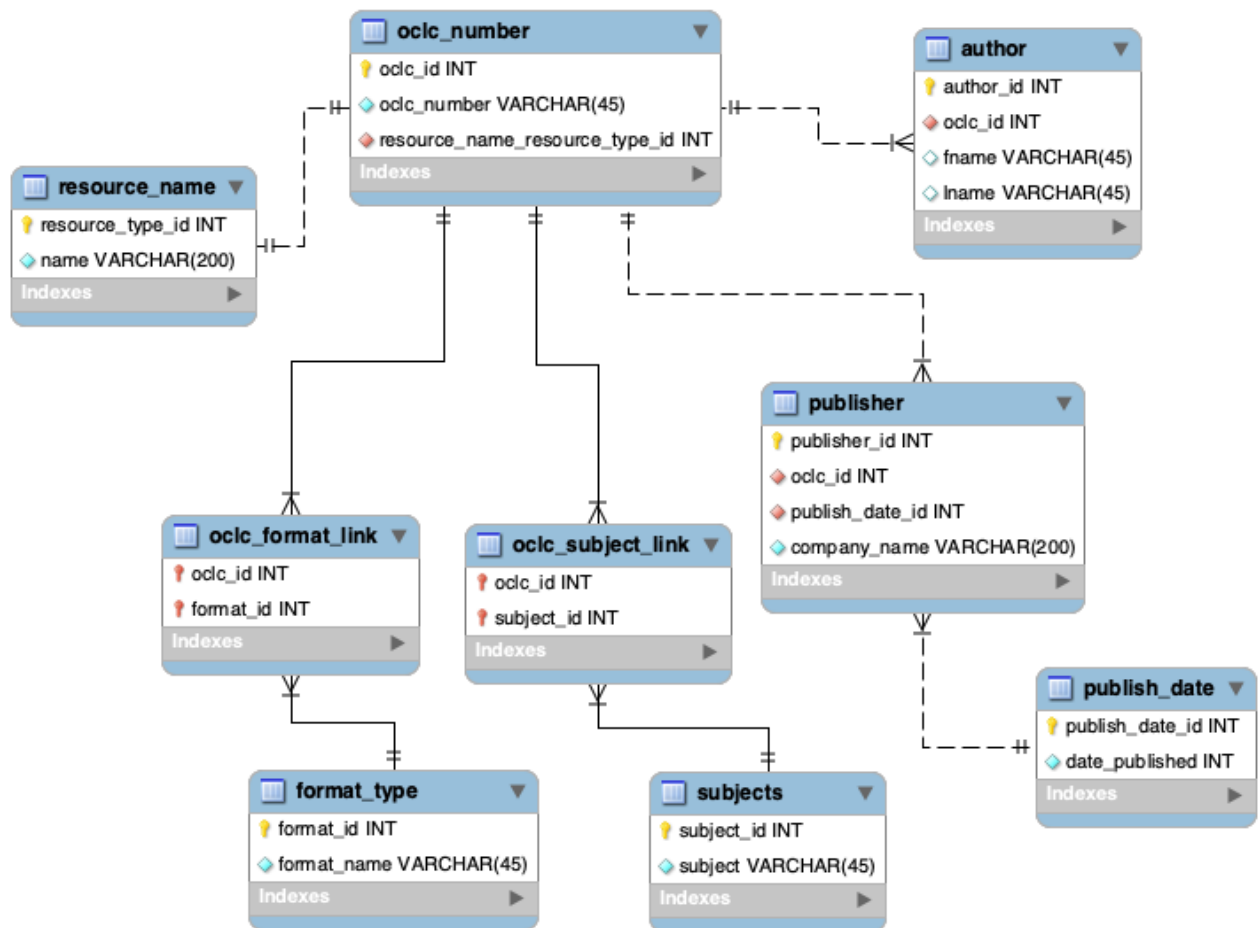
Database Description

The database will consist of books, newspapers, magazines, encyclopedias, and many other items in the McKeldin Library. The database will hold the name of the resource, the subject/genre of the article, author, OCLC number, publisher, date published, and the format type (online or hardcopy) for the article that is being added into the database.

Logical Design:

To create a database that is highly intuitive and technical, it was vital to create and form a logical design that would best demonstrate the complexity of our envisioned database. In order to do so, we created eight different tables that would describe the different resources that are found in McKeldin Library (physically and in their online database).

(ERD Diagram)



One thing that is noticeable is that our logical design is traced back to the oclc_number. We realized that the oclc_number is something that identifies all different types of resources in a library uniquely, and so it is vital that we use this number to connect all of our tables in our database and logical design. For all of the tables (minus the format_type and publish_date tables), it can be seen that oclc_id is a foreign key that connects the tables back to the oclc_number table. This provides for cohesion and connectedness between all of the tables in the database.

We also had to think through the different relations that the tables would have with each other. This was initially something that was very difficult to do, as we had to figure out how different resources differ in the contents that would be put into a table, but we eventually were able to create strong relationships between all the tables. Perhaps the most important relationship we made was between publish_date and oclc_number. We realized that there is no direct

relationship that would be sufficient to draw in our database. However, it was clear that we could create a relationship between the publish_date and publisher tables. From there, we created a one to many relationship between the oclc_number and publisher table, which in turn indirectly connected the publish_date and the oclc_number tables through the publisher table. Following this same process, we were able to draw connections and establish proper relationships between all of the different tables in our database.

One more thing we realized was that we needed to decide which values could hold null values or be left blank. From the sample data that we had gathered, we found that there were many resources where there were not recorded authors (mostly due to how early the resources were written, some of which were in the early 1800s). In order to combat this problem, we made the “fname” and “lname” columns in the author table able to contain NULL or no values.

Physical Design

Our database seeks to meet the needs of UMD affiliates in assisting them find any and all of the resources they are looking for within the McKeldin library. The data we retrieved from “WorldCat” was our first step in physically designing our database. We considered what information we would include in our database and chose to include data like publish dates, publishers, authors, resource types, online or hardcover resources, oclc numbers, etc. We accommodate for each of these data types with a corresponding table. For example, for the publisher and publish date we created a publisher table then a publish date table.

Sample Data

We weren’t able to find a CSV file that contained any of the data that we were looking for so we had to create our own CSV files. We had to use the “World Cat” database to search for data and input the data we were looking for. Filling out these files and finding the data that we were searching for was a very time consuming process. Luckily the database we were using had all of the information we were looking for. For example it listed whether or not the resources were online, listed the publishers, sku numbers, etc. After creating the csv files we transferred the data into the tables in our database.

Views / Queries

<u>View Name</u>	<u>Req. A</u>	<u>Req. B</u>	<u>Req. C</u>	<u>Req. D</u>	<u>Req. E</u>
routledge_authors	X	X			X
hard_copy_books	X	X	X		
no_author_publisher	X	X	X		
oclc_online_works	X	X		X	
recorded_author_works	X	X		X	
Total Queries:	5	5	2	2	1

The following list describes what each query we wrote for our database displays:

Query 1: Our first query will show Which authors had their works published by Routledge.

Query 2: Our second query shows how many works in the library are hard copies and what subjects they are.

Query 3: The third query shows which publishers have published works before 2000.

Query 4: The fourth query will show the OCLC numbers of the works that are found online.

Query 5: Our last query will show which works have a recorded author, excluding those which did not have a name associated with it.

Changes from original design

Since our work on previous database design from the last progress report, we have since updated our ERD diagram in order to address some of the inconsistencies between table relationships and to better accommodate the gathered data into our database model.

Most notable changes were the relationship between OCLC number table and other tables including subjects, author, and format_type table, which was previously named as “format” in our last model. We have found many-to-many relationships between tables to be the best fit for most of our database model and because of that we have modified every table relationship with one-to-many relationships to many-to-many relationships and included additional tables to create

a link between the tables with the said relationship. The author table however, was an exception in this case and the relationship between the author table and OCLC number table was changed from many-to-many relationships due the fact that a single instance of author corresponds to multiple instances of OCLC numbers.

There were additional minor changes to some of the tables as well, such as the name of the table and their columns to abide by the naming convention and changes to columns to accept null values in its dataset in case of additional author or book data being included into our database model. For instance, 'lname' column in the author table, which represents the last name of the author, has been modified to accept null values to accommodate for a special case where the author name is not present or available in the newly added dataset of books.

Database Ethics

In the early stages of the project, something we needed to take into consideration was having non-bias data since we would only be including a small sample of Mckeldin's resources. We kept this in mind and made sure to have a diverse sample of different formats, authors from different backgrounds, and different genres and topics to have the most inclusive data that had the best representation for our target audience.

Lessons Learned

Overall, throughout the creation of the database, we had some obstacles and were able to find a solution either by collaborating as a team or with the help of TAs and professors. Towards the end of the project, we had trouble with our ERD diagram and the relationship between tables. We had to take the time to rework after receiving feedback from peers and TAs to better represent what we were trying to create. This was a learning experience that we had to further improve our understanding of the ERD diagram.

Potential Future Work

For starters, if we had more time we would add much more data to the database to give a better representation of all the works that the library has to offer, what we included was only a small fraction of the published materials that can be found. We could also include more than published reading materials such as the technology and other miscellaneous items and services

that they offer. As far as what we could change, including more description of the works that we did put into our database would be a useful and practical thing. For the most part we only included the base bones descriptors like author, publisher, date, oclc information and format information. For example we could also include the abstract or a brief description of what the published material like most library databases include.

References:

Worldcat UMD. *Umaryland.on.worldcat.org*, <https://umaryland.on.worldcat.org/discovery>.