**Christina Wofford**

**CSCI-765**

**Final Project Whitepaper**

**Abstract**

GenreBridge is an application that seeks to provide users with an enhanced reading experience by providing them with an optimized soundtrack for their current piece of literature in order to maximize the benefits of both reading and listening to music. Additionally, GenreBridge attempts to present users with literature and music that is directly relevant to their interests, allowing them to partake in a solo hobby while also staying connected to the predominate trends in current media.

Unlike simply searching for playlists on Spotify, GenreBridge goes a step further by collecting additional data to reveal interesting trends and patterns in order to allow users to discover new connections between books and music, as well as to gain insight into the popularity of what they are currently reading and watching. Importantly, GenreBridge does not require any sort of login, providing a low barrier to entry and minimizing security concerns for users.

By understanding the factors that contribute to a successful pairing, GenreBridge can offer more accurate and personalized suggestions and combine the best of two worlds by helping users find the soundtrack for their next read, something valuable to book lovers and music enthusiasts alike.

**Introduction**

“Research has shown that pairing books with music can increase immersion and create a deeper emotional connection with the material, ultimately improving the overall experience for readers” (Boltz, 2001), an idea that is fundamental to the goal of GenreBridge.

Genre Bridge collects data that allows for the calculation of the implied popularity of a book based on the number of followers of its associated playlist and the popularity of the songs contained within the playlist. While this is a very imperfect measurement, there is evidence that the general trend holds true. The “Results” section of this paper provides further insight into the data from which the implied popularity of a book can be calculated (Figure 1).

Before going further, it is important to make mention of why it is even prudent to consider the popularity of any piece of media at all. Individual tastes in both literature and music span a variety of genres and often serve as a means of escapism which is both valid and valuable. That said, there are benefits to stepping outside of one’s comfort zone and intentionally engaging with popular media. "Engaging with popular culture can foster a sense of belonging, promote self-esteem, and offer possibilities for self-reflection and self-discovery" (Van den Bulck, Custers, & Nelissen, 2016, p. 1355) and even the potential of unlocking these benefits makes it worth engaging with popular media, even if only from time to time.

It is easy enough to look a New York Times Bestsellers list or turn on the radio to determine the books and movies that are currently popular, but these sources only provide a one-dimensional view into the best media to engage with. As was previously mentioned, the benefits of reading are heightened by the experience of listening to music and vice versa. In order for one to make the best use of their limited free time while still obtaining a high level of benefit, they must optimize their experience and GenreBridge aids in this endeavor.

**Chart

Description automatically generated**

Figure 1 - Implied Book Popularity

Before moving into the full set of results and conclusions gained during the development of GenreBridge, it is prudent to understand the manner in which it incorporates core principles of database management, as well as a high-level overview of its implementation.

**Course Relevance**SQL Data Definition and Manipulation:

GenreBridge employs the fundamentals of data definition and manipulation. The application is built upon a well-structured data model that was built by creating appropriate tables, attributes, and relationships within the database to represent real-world entities, such as books, authors, and playlists.

The tables for the application were constructed by using SQL commands to specify data types, constraints, and primary and foreign keys to ensure data integrity and ensure efficient data retrieval. GenreBridge also uses SQL to manipulate data by inserting, updating, and deleting records as needed. This process includes querying the Google Books and Spotify APIs to retrieve relevant data, inserting the data into the corresponding tables, and establishing relationships between entities. GenreBridge also utilizes SQL queries to aggregate, filter, and analyze the data in order to discover trends, draw insights, and match users with playlists that complement their reading choices.

Applications:

Multiple aspects of GenreBridge are related to the fundamentals of applications, particularly the components that enable input, either via API or directly from the user. For instance, GenreBridge uses the psycopg2 package to store the data returned by Python queries to the Spotify and Google Books APIs in Postgres, connecting the application layer to the backend. Additionally, the application makes use of several other Python packages (spotipy, configparser, google auth) to establish connections and authenticate with the Google Books and Spotify APIs.

Aside from Python packages, GenreBridge also makes use of HTML forms. The frontend of the application gathers user information through two HTML forms for the author's first and last names. These forms facilitate querying the Google Books API and subsequently the Spotify API before storing the data, matching the results, and displaying the information on the frontend for the user. Lastly, GenreBridge leverages a config file to connect to the DB and the Spotify API, ensuring that the username, password, and secret key remain hidden within the main Python application.

Data Modeling:

Diagram

Description automatically generated

Figure 2 - GenreBridge ER Diagram

The project heavily relies on the core concepts of data modeling (Figure 2). Gathering the requirements for the data model was one of the very first steps in performing GenreBridge’s initial data modeling. As result, before building or designing any tables in the database, it was necessary to collect data requirements to support the project, which involved developing a robust concept as a basis for identifying data points.

Understanding and implementing concepts related to entities, attributes, and relationships was crucial for devising a logical way to store the data, ensuring the database structure accurately represented real-life scenarios. For instance, the relationship between books and authors is many-to-many since a book can have multiple authors and an author may have written more than one book. Consequently, storing books and authors in a single entity is unfeasible. To accommodate this, the information is distributed across three tables (books, authors, and book\_author\_mapping), allowing for storage and access in a normalized manner.

It was also essential to understand and implement concepts related to entities, attributes, and relationships to come up with a logical way to store the data so that the database structure was representative of real life. For example, the relationship between books and authors is many-to-many, as a book can have multiple authors and an author has likely written more than one book. Because of this, there is no clean way to store books and authors in a single entity. Instead, this information is spread across three tables (books, authors, and book\_author\_mapping) to allow this data to be stored and accessed in a normalized way.

**Implementation**

The frontend of GenreBridge is built on HTML, CSS, JavaScript, and Bootstrap and enables users to input a book title and author to receive a complementary Spotify playlist (Figure 3). The application uses Python and Flask to query the Google Books API using user input. In doing so, the application retrieves additional information about the book, including publication year, ISBN, and associated keywords; GenreBridge stores all the book information across several tables in a normalized PostgreSQL database.

**Graphical user interface, application

Description automatically generated**

Figure 3 - Search form on GenreBridge front end

Next, GenreBridge makes another API request to Spotify to retrieve a matching playlist. The application initially searches based on the book title, then the author's name, and finally by keywords if no match is found by title or author (Figure 4).

Text

Description automatically generated

Figure 4 - Playlist search logic

Various pieces of information about the playlist returned by this query are stored in the database. In addition to storing the prudent details of the playlist itself (name, Spotify playlist ID), more granular pieces of information including all tracks, artists, and albums that are present on the playlist. The app collects a robust set of data in order to aid in its aforementioned endeavors.

GenreBridge collects the number of Spotify followers for each playlist and the popularity of each track, allowing the creation of a figure comparing the playlist's popularity with a book's publication year to serve as a metric for determining the correlation between the book's age and the matched playlist's popularity. Collecting the number of pages in the book and the number of tracks in the playlist also enables analysis of correlations between book and playlist lengths.

Once a match has been made, GenreBridge populates the pre-built Spotify playlist widget with the returned playlist ID and presents the widget to the user on the frontend where they can click through and interact with the playlist directly through Spotify, should they so choose (Figure 5).

Graphical user interface

Description automatically generated

Figure 5 - GenreBridge with Spotify widget

Finally, GenreBridge uses matplotlib to analyze the collected data to look for trends and create figures and diagrams to visualize them for better comprehension, providing understanding of how the matching process works and highlighting connections between various attributes, such as book publication year, playlist popularity, and track count.

With implementation details having been accounted for, the foundation necessary for exploring the insights generated by GenreBridge is in place and those findings can be explored in greater depth. Each of the figures in the following section is accompanied by a caption summarizing the associated findings.

**Results**

**Chart, scatter chart

Description automatically generated**

Figure 6 - Book Publication Year and Average Song Popularity

Reading a newly released book does not by default mean that a user will be exposed to the most popular music in their matched playlist. In fact, no book in the dataset that was released within the past five years had greater than average popularity score of 50 based on Spotify’s algorithms.

Chart, scatter chart

Description automatically generated

Figure 7 - Book Publication Year vs. Average Song Release Year

There is a positive correlation between reading newly released books and being exposed to new music, even if it’s extremely popular. This makes sense as older books are more likely to have older playlists that were constructed before the release of music that is currently popular.

**Chart

Description automatically generated**

Figure 8 - Distribution of # of Playlist Followers

Most playlists are niche and have under 500 followers and include a very small number of songs are extremely popular. This means that users are still being presented with the chance to stumble across a hidden gem while being exposed to popular new music.

**Chart, pie chart

Description automatically generated**

Figure 9 - Distribution of Song Popularity Score  
  
The number of songs that are extremely popular at any given point in time is incredibly low, with about 5% of songs achieving a score over 80 per Spotify’s ranking based on their internal data.

**Chart, bar chart

Description automatically generated**

Figure 10 – Number of Playlist Followers vs Average Song Popularity

Music that is not extremely popular is more likely to be niche genres, allowing for exposure to new things. With 65% of playlists having 0-500 followers, the odds of a user being presented with one of these is high. This does not impede exposure to new, popular music.   
  
Even though only 7.5% of playlists have over 10,000 followers, they are almost certain to introduce users to what is popular, particularly given that the number of incredibly popular songs at any given time is relatively small as is evidenced in Figure 9.

**Chart, histogram

Description automatically generated**

Figure 13 - Track Release Year Distribution   
  
The overrepresentation of new music in playlists makes it very likely that a user will not only be exposed to new, popular music, but new music in general, which again provides a growth opportunity.

**Conclusions**

The insights in the figures above provide a solid foundation for refining the book and music matching process in the future, highlighting the power of GenreBridge.

By going beyond merely searching for playlists on Spotify, GenreBridge delves deeper by collecting additional data to uncover intriguing trends and patterns, allowing users to discover novel associations between books and music, as well as to gain insights into the popularity of their current reading and listening choices. By collecting and analyzing the elements that contribute to successful book/music pairings, GenreBridge is able to provide more precise and tailored suggestions, merging the finest aspects of both literary and musical worlds.

In summation, GenreBridge helps users find the connectedness and “group cohesion” provided by music (Chanda & Levitin, 2013) and the gain the knowledge and enhanced empathy that accompany reading (Oatley & Djikic, 2020) in a single, unified experience.

**References**

* Boltz, M. G. (2001). Musical soundtracks as a schematic influence on the cognitive processing of filmed events. Music Perception: An Interdisciplinary Journal, 18(4), 427-454.  
  <https://doi.org/10.1525/mp.2001.18.4.427>
* Chanda, M. L., & Levitin, D. J. (2013). The neurochemistry of music. Trends in Cognitive Sciences, 17(4), 179-193. <https://doi.org/10.1016/j.tics.2013.02.007>
* Johansson, Roger, et al. “Eye Movements and Reading Comprehension While Listening to Preferred and Non-Preferred Study Music.” Psychology of Music, vol. 40, no. 3, 2011, pp. 339–356., <https://doi.org/10.1177/0305735610387777>.
* Oatley, K., & Djikic, M. (2020). Reading and Life: Empathy, Knowledge, and the Retention of Plot. Review of General Psychology, 24(3), 285-296.  
  <https://doi.org/10.1177/1089268020919756>
* Van den Bulck, H., Custers, K., & Nelissen, S. (2016). The psychological benefits of superheros and other popular culture for fans of all ages. The Journal of Popular Culture, 49(6), 1349-1359.

<https://doi.org/10.1111/jpcu.12494>