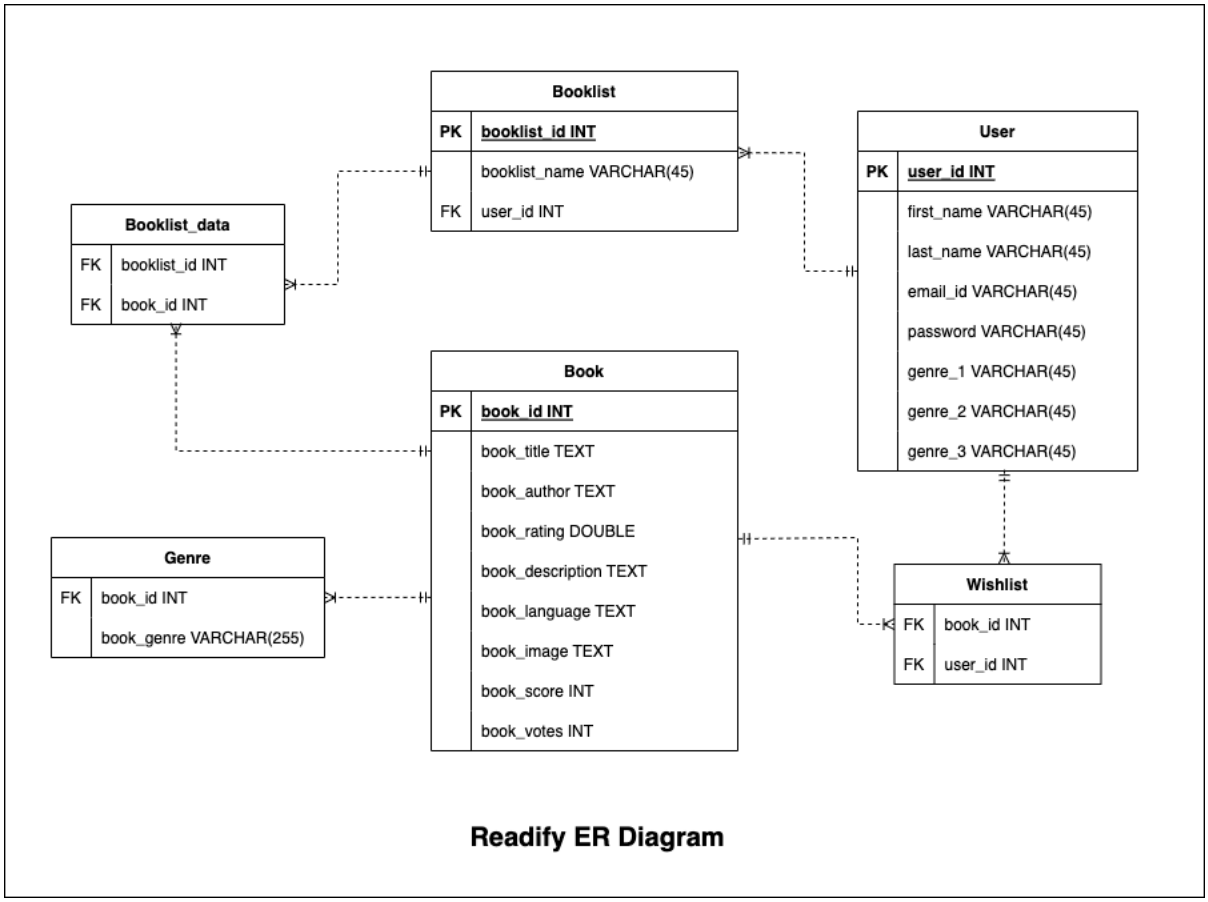
ADT Final Project Part 2

Team 24
Project Topic – Readify

Team: Ameya Dalvi, Shefali Luley, Shubham Bhagat

1] Create ER Diagram



We have 5 entities in total. They are User, Book, Wishlist, Booklist, Booklist_data, and Genre.

Entities:

User

- The User table will have all the information necessary to maintain a user profile in our system.
- Each record will contain information such as user_id, firstname, lastname, email id, password (hashed), and top 3 genres that the user likes.
- User_id will be the primary key in this table.

Book

- The Book table will have records of books we got from the dataset.
- At first there were 25 attributes in the dataset. We selected those attributes which are required for our project. Each record in this table will have fields such as book_id, title, author, description, rating, language, cover image, score, and votes.
- These attributes will be used to apply filters while exploring different books.
- Book_id will be the primary key in this table.

Wishlist

- This table will have books which the user wants to read in future.
- It will have attributes – book_id and user_id which are used to identify a book and a user respectively.
- Book_id will be the foreign key referenced from the book table and user_id will be the foreign key referenced from the user table.

Genre

- Each book is associated with multiple genres. Adding an attribute 'genre' in the book table with comma separated list of genres wouldn't be appropriate.
- Therefore, we are maintaining a different table which will indicate the genre of each book.
- Each record in this table will have book_id and its genre. If one book has multiple genres associated with it, there will be multiple records with same book_id but different genre name.
- Book_id is the foreign key referenced from book table.
- This table will be an important component for designing the recommendation system.

Booklist

- Whenever a user creates a booklist, a new record will be added in this table. Thus, this table is used to indicate which booklist belongs to which user.
- It will have fields such as booklist_id, booklist_name, and user_id.
- Booklist_id is the primary key and user_id will be the foreign key referenced from user table.

Booklist_data

- Whenever a user adds any book to his booklist, a new record will be added in this table. Thus, this table is used to specify which booklist contains which books.
- It will have attributes such as booklist_id and book_id where booklist_id is the foreign key referenced from booklist table and book_id is foreign key referenced from book table.

Relationship

User and Wishlist – *1: Many* relationship

User and Booklist – *1: Many* relationship

Book and Wishlist – *1: Many* relationship

Booklist and Booklist_data – *1: Many* relationship

Book and BookList_data – 1: Many relationship

Book and Genre – *1: Many* relationship

2] Describe Data types

User Table:

Attribute	Datatype
user_id	INT
firstname	VARCHAR (45)
lastname	VARCHAR (45)
email_id	VARCHAR (45)
password	VARCHAR (45)
genre_1	VARCHAR (45)
genre_2	VARCHAR (45)
genre_3	VARCHAR (45)

Wishlist Table:

Attribute	Datatype
user_id	INT
book_id	INT

Book Table:

Attribute	Datatype
book_id	INT
book_title	TEXT
book_author	TEXT
book_rating	DOUBLE
book_description	TEXT
book_language	TEXT
book_image	TEXT
book_score	INT
book_votes	INT

Genre Table:

Attribute	Datatype
book_id	INT
book_genre	VARCHAR (255)

Booklist Table:

Attribute	Datatype
booklist_id	INT
booklist_name	VARCHAR (45)
user_id	INT

Booklist_data Table

Attribute	Datatype
booklist_id	INT
book_id	INT

3] Describe Constraints and Primary Keys

User Table:

- **User_id** is the primary key used to uniquely identify each user. It will be unique, not null, and auto incremented for each user.
- Rest of the attributes will be not null.

Wishlist Table:

- **Book_id** is a foreign key which is referencing 'book_id' attribute (primary key) of book table.
- Similarly, **user_id** is a foreign key which is referencing 'user_id' attribute (primary key) of user table.

Book Table:

- **Book_id** is the primary key used to uniquely identify each book. It will be unique, not null, and auto incremented for each book.
- Rest of the attributes will be not null.

Genre Table:

- **Book_id** is a foreign key which is referencing 'book_id' attribute (primary key) of book table.
- **Book_genre** will be not null.

Booklist Table:

- **Booklist_id** is the primary key used to uniquely identify each booklist. It will be unique, not null, and auto incremented for each booklist created.
- **user_id** is a foreign key which is referencing 'user_id' attribute (primary key) of user table.
- **Booklist_name** will be not null.

Booklist_data Table:

- **Booklist_id** is a foreign key which is referencing 'booklist_id' attribute (primary key) of booklist table.
- **Book_id** is a foreign key which is referencing 'book_id' attribute (primary key) of book table.

4] Code to create Database

Create readify database:

```
1       CREATE SCHEMA `readify` ;
```

Create user table:

DDL for readify.user

```
1   CREATE TABLE `user` (
2     `user_id` int NOT NULL AUTO_INCREMENT,
3     `first_name` varchar(45) NOT NULL,
4     `last_name` varchar(45) NOT NULL,
5     `email_id` varchar(45) NOT NULL,
6     `password` varchar(45) NOT NULL,
7     `genre_1` varchar(45) NOT NULL,
8     `genre_2` varchar(45) NOT NULL,
9     `genre_3` varchar(45) NOT NULL,
10    PRIMARY KEY (`user_id`),
11    UNIQUE KEY `user_id_UNIQUE` (`user_id`),
12    UNIQUE KEY `genre_1_UNIQUE` (`genre_1`),
13    UNIQUE KEY `genre_2_UNIQUE` (`genre_2`),
14    UNIQUE KEY `genre_3_UNIQUE` (`genre_3`)
15  ) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

Create book table:

DDL for readify.book

```
1   CREATE TABLE `book` (
2     `book_id` int NOT NULL AUTO_INCREMENT,
3     `book_title` text NOT NULL,
4     `book_author` text NOT NULL,
5     `book_rating` double NOT NULL,
6     `book_description` text NOT NULL,
7     `book_language` text NOT NULL,
8     `book_like_percent` double NOT NULL,
9     `book_image` text NOT NULL,
10    `book_score` int NOT NULL,
11    `book_votes` int NOT NULL,
12    PRIMARY KEY (`book_id`)
13  ) ENGINE=InnoDB AUTO_INCREMENT=41653 DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

Create wishlist table:

DDL for readify.wishlist

```
CREATE TABLE `wishlist` (
  `book_id` int NOT NULL,
  `user_id` int NOT NULL,
  KEY `book_id_idx` (`book_id`),
  KEY `user_id_idx` (`user_id`),
  CONSTRAINT `book_id_1` FOREIGN KEY (`book_id`) REFERENCES `book` (`book_id`),
  CONSTRAINT `user_id_1` FOREIGN KEY (`user_id`) REFERENCES `user` (`user_id`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

Create booklist table:

DDL for readify.booklist

```
CREATE TABLE `booklist` (
  `booklist_id` int NOT NULL AUTO_INCREMENT,
  `booklist_name` varchar(45) NOT NULL,
  `user_id` int NOT NULL,
  PRIMARY KEY (`booklist_id`),
  UNIQUE KEY `booklist_id_UNIQUE` (`booklist_id`),
  KEY `user_id_idx` (`user_id`),
  CONSTRAINT `user_id_2` FOREIGN KEY (`user_id`) REFERENCES `user` (`user_id`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

Create booklist_data table:

DDL for readify.booklist_data

```
CREATE TABLE `booklist_data` (
  `booklist_id` int NOT NULL,
  `book_id` int NOT NULL,
  KEY `booklist_id_idx` (`booklist_id`),
  KEY `book_id_idx` (`book_id`),
  CONSTRAINT `book_id_3` FOREIGN KEY (`book_id`) REFERENCES `book` (`book_id`),
  CONSTRAINT `booklist_id_3` FOREIGN KEY (`booklist_id`) REFERENCES `booklist` (`booklist_id`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

Create genre table:

DDL for readify.genre

```
CREATE TABLE `genre` (
  `book_id` int NOT NULL,
  `book_genre` varchar(45) NOT NULL,
  KEY `book_id_idx` (`book_id`),
  CONSTRAINT `book_id` FOREIGN KEY (`book_id`) REFERENCES `book` (`book_id`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

5] Explain how data is imported

We have used MySQL CLI to import data from CSV file to the table in the readify database.

Set global local_infile to 1 to use load command for importing csv data to table

```
mysql> SET GLOBAL local_infile=1;
Query OK, 0 rows affected (0.00 sec)
```

Start MySQL with local-infile set as 1

```
shubham@Shubhams-MacBook-Pro ~ % mysql --local-infile=1 -u root -p
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 25
Server version: 8.0.28 MySQL Community Server - GPL
```

Use load data local in file command to import data from csv file to the tables in the database

```
mysql> LOAD DATA LOCAL INFILE '/Users/shubham/Desktop/ADT Project/BookFinal.csv' INTO TABLE book FIELDS TERMINATED BY ',' ENCLOSED BY '"' LINES TERMINATED BY '\r\n' IGNORE 1 ROWS;
Query OK, 41668 rows affected (0.55 sec)
Records: 41668  Deleted: 0  Skipped: 0  Warnings: 0

mysql> LOAD DATA LOCAL INFILE '/Users/shubham/Desktop/ADT Project/GenresFinal.csv' INTO TABLE genre FIELDS TERMINATED BY ',' ENCLOSED BY '"' LINES TERMINATED BY '\r\n' IGNORE 1 ROWS;
Query OK, 353803 rows affected (0.75 sec)
Records: 353803  Deleted: 0  Skipped: 0  Warnings: 0
```

Data Dictionary

User Table:

Table_schema	Table name	Column_name	Data_type	nullable	Example
Readify	User Table	user_id	INT	N	101
Readify	User Table	firstname	VARCHAR (45)	Y	Ameya
Readify	User Table	lastname	VARCHAR (45)	Y	Dalvi
Readify	User Table	email_id	VARCHAR (45)	Y	abdalvi@iu.edu
Readify	User Table	password	VARCHAR (45)	N	rootroot
Readify	User Table	genre_1	VARCHAR (45)	N	Thriller
Readify	User Table	genre_2	VARCHAR (45)	N	Horror
Readify	User Table	genre_3	VARCHAR (45)	N	Crime

Book Table:

Table_schema	Table name	Column_name	Datatype	nullable	Example
Readify	Book Table	book_id	INT	N	102
Readify	Book Table	book_title	TEXT	N	Harry Potter
Readify	Book Table	book_author	TEXT	Y	J.K. Rowling
Readify	Book Table	book_rating	DOUBLE	Y	4.6/5
Readify	Book Table	book_description	TEXT	Y	“A great book”
Readify	Book Table	book_language	TEXT	Y	English
Readify	Book Table	book_image	TEXT	Y	url
Readify	Book Table	book_score	INT	Y	96
Readify	Book Table	book_votes	INT	Y	4

Wishlist Table:

Table_schema	Table name	Column_name	Datatype	nullable	Example
Readify	Wishlist Table	book_id	INT	N	102
Readify	Wishlist Table	user_id	INT	N	101

Genre Table:

Table_schema	Table name	Column_name	Datatype	nullable	Example
Readify	Genre Table	book_id	INT	N	102
Readify	Genre Table	book_genre	VARCHAR(255)	Y	“Mystery”

Booklist Table:

Table_schema	Table name	Column_name	Datatype	nullable	Example
Readify	Booklist Table	booklist_id	INT	N	103
Readify	Booklist Table	user_id	INT	N	101
Readify	Booklist Table	booklist_name	VARCHAR (45)	Y	“My Mystery BookList”

Booklist_data Table:

Table_schema	Table name	Column_name	Datatype	nullable	Example
Readify	Booklist_data Table	booklist_id	INT	N	103
Readify	Booklist_data Table	book_id	INT	N	102