**ASTON UNIVERSITY**

**BIRMINGHAM, UK**

**CS41DS**

**DATABASE SYSTEMS**

**COURSEWORK**

**INDIVIDUAL ASSIGNMENT**

**DIVYE RAJESWARI PAMULA**

**220318055**

**MSc COMPUTER SCIENCE**

**COLLEGE OF ENGINEERING AND PHYSICAL SCIENCES**

**R1:**

**A: RELATION SCHEMA**

Diagram

Description automatically generated

Fig1: Relational Schema for science archive database system

**B: Decision Discussion**

* The users' table is designed to store information about the users of the system. The client specified generic data, such as the account's user name and password, contact details, contact preference, and user types (reader, author, and author and reader). The UserID attribute is set to be generated always as an identity and is chosen as the primary key.
* Publication\_Author table is used to store information of the authors, who are published. The client mentioned about recommendations of co-authors. In this based on author type(i.e author , co -auhthor) are divided it is used to recommend publications, based on the authors and co-authors.
* Subjects tables is used to store information about scientific subjects . subject id is a primary key each subject has unique id. The client specified that published item may related to more than one subjects. So each subject has unique id to give suggestions. User interest subjects table is used to store information about user interested subjects , publications subjects is used to store information about published items subjects and used to suggest publications based on user interest. The client specified that recommendation based on user interest.
* The client specified the user should be able to review items and score them accordingly. so the reviews table allows users to score and comment on publication items and review id is the primary key, each reviewer information is stored. The bookmarks table allows users to save publications for later reference and it has unique id as bookmarkid.
* based on the client information, the database should keep track of all items a user has read, that is, visited using a hyperlink and/or provided a review/score. User library is used to store user visited information of each user and it tracks all the read publications . Based on review score it suggest the publication to the user.
* publisher tables allows to store information about publisher such as publisher name(IEEE, ACM ect..), publisher email, publisher website and publisher id is the primary key and store each publisher details with unique id.publisher name, pblisher email publisher website all are unique and it doen’t allow the already publisher details.
* Publication table is the main table to store all the information of the publications and each publication has it unique id . User id is used as foreign key , to track who added the publication to the database. The Journal Article, Conference, Books, Thesis tables store information about futher publication details of each publication basde on their type. All the information is stored in these table which are specified by the client.
* The system should be able to make recommendations based on a combition of several factors, including user interest in specific subjects, reviews, and ratings of an item and collaborations.

**C: DDL COMMANDS**

Text

Description automatically generated with low confidence

Fig2: DDL COMMAND

Text

Description automatically generated

Fig3: DDL COMMANDS

Text

Description automatically generated

Fig4: DDL COMMANDS

**SOME OF THE INSERT COMMANDS:**

Graphical user interface, text, application

Description automatically generated

Fig5: user insert

Graphical user interface, text, application

Description automatically generated

Fig6: publication insert

Graphical user interface, application, Word

Description automatically generated

Fig7: Journal\_Articles insert

**D: INTEGRITY CINSTRAINT**

**NOT NULL CONSTRAINT:**

Graphical user interface, text, application

Description automatically generated

Fig8:NOT NULL Constraint

This statement violates the NOT NULL constraint on the phone column and will result in an error.

**UNIQUE CONSTRAINT:**

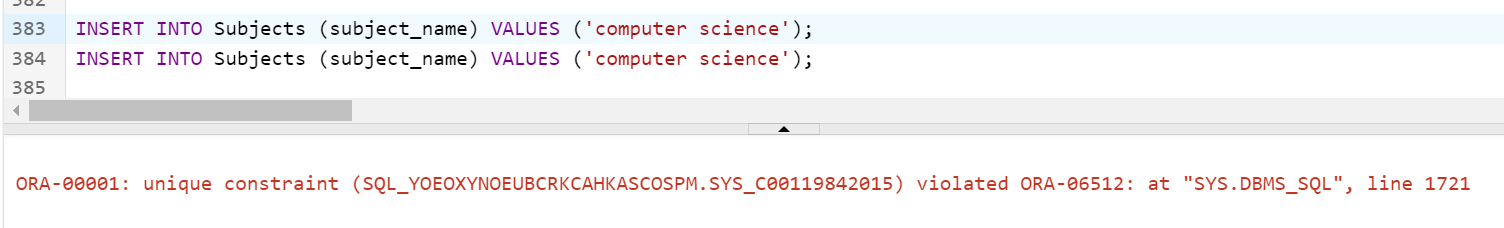


Fig9:UNIQUE CONSTRAINT

This statement violates the unique constraint on the subjectname columd and will result in an error.

**CHECK CONSTRAINT:**

Graphical user interface, text, application

Description automatically generated

Fig10:CHECK CONSTRAINT

This statement violate theCHECK constraint n the user\_type column and will result in an error

**FOREIGN KEY CONSTRAINT:**

Graphical user interface, application

Description automatically generated

Fig11: FOREIGN KEY CONSTRAINT

This statement violates the foreign key constraint on the publisher id column and will result in an error, as there is no publisher with an ID of 999 .

**PRIMARY KEY CONSTRAINT:**

A picture containing diagram

Description automatically generated

Fig12: PRIMARY KEY CONSTRAINT

This statement violates the primary key constraint on the publication id column and will result in an error, as the same id is begin used for two different rows.

**R2: LIST OF IDENTIFIED BUSINESS PROCESS OF USER TYPE**

Based on the client interview process, there are different type of users, that will interact with the science archive system.

* Search Publication details
* Read recommended
* Add review /rating
* Update ,detele user account
* View further publication details based on search
* View the further publication details for the search result
* Store user reading history
* Recommendation based on co-authors:
* Recommendation based on user interest
* filter

**Publication details:**

**Table

Description automatically generated**

Fig13:publication details of user search

**Filter the search items:**

**Graphical user interface, table

Description automatically generated**

Fig14: based on user search filter the publication types

**View further publication details**

**Graphical user interface, application

Description automatically generated**

Fig15: See futher publication details

**Update the user details**

**Graphical user interface, application, Word

Description automatically generated**

Fig16: Update the user details

**Read the user read history**

**Text

Description automatically generated**

Fig17: Read the user history

**View the user history**

**Table

Description automatically generated**

Fig18:View the user history

**Recommendation based on user interest**

**Graphical user interface

Description automatically generated with medium confidence**

Fig19:Recommendations based on user interest

**Can view score and review text for the publication items**

**Graphical user interface, application

Description automatically generated**

Fig20: view score of the publication items

**Recommendation based on co-authors:**

Graphical user interface, application

Description automatically generated

Fig21: Recommendations based on co-authors

**R3: Kimball’s Dimensional Model**

The proposed schema can be reorganised to meet Kimball's Dimensional Model, which has a number of advantages. The Dimensional Model divides information into fact and dimension tables. Dimension tables include descriptive data such as time, location, and product, whereas fact tables offer quantitative data such as sales. Because it provides a simpler and more intuitive structure for data queries, this model is well-suited for analytics and reporting.

We need to determine the facts and dimensions with which the business process interacts in order to reorganise the schema to meet Kimball's Dimensional Model. The fact table in the proposed schema is the "Publications" table. Dimension tables include "Subjects," "Publishers," "Users," "Conference\_Papers," "Books," "Thesis," "Journal\_Articles," "PublicationAuthors," "User\_Interest\_Subjects," "publication\_subjects," "Reviews," "User\_Library," and "Bookmark."

The identified business process interacts with the fact table, "Publications," as well as several dimension tables, including "Subjects," "Publishers," "Users," "Conference\_Papers," "Books," "Thesis," "Journal\_Articles," and "publication\_subjects."

We can add new data marts to the Dimensional Model as the system's capability grows. A data mart is a subset of data devoted to a specific business operation or department. Each business process can have its own data mart that interacts with the fact and dimension tables. We can, for example, establish a data mart for user profiles that incorporates user information from the "Users" dimension table. We may also build a data mart for publication analytics using data from the "Publications" fact table.

Diagram

Description automatically generatedFig22: Kimballs dimensions models