PROJECT DELIVERABLE 01

Group 13

Individual Contribution			
CWID	**Name**	**Contribution (Description)**	**Percent Contribution**
A20593079	Akshada Ranalkar	Defined 4 entities and their Attributes, Designed ERD, Relational Schema, Documentation	33.3%
A20563287	Anuja Wavdhane	Defined 3 entities and their Attributes, Designed ERD, Relational Schema, Documentation	33.3%
A20560966	Suhasi Gadge	Defined 3 entities and their Attributes, Designed ERD, Relational Schema, Defined Bridge Entities, Documentation	33.3%

Project Title:

ACADEMIC PUBLICATION DATABASE SYSTEM

(REFERENCE WEBSITE: IEEE Website

https://ieeexplore.ieee.org/abstract/document/344065)

1.1 Introduction

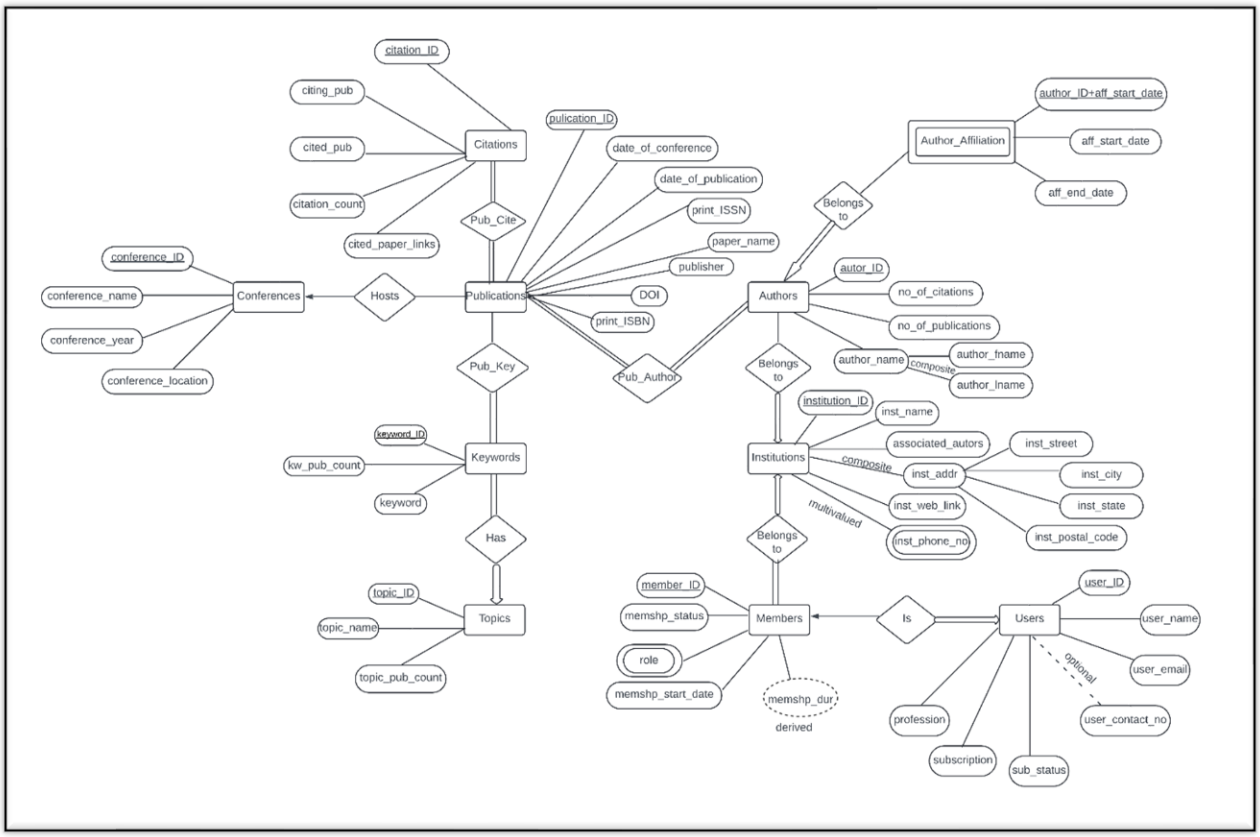
In this project, we are designing a database system similar to the IEEE platform to manage and organize academic publications, authors, conferences, institutions, and more. The goal is to create a structured and efficient way to store and retrieve information related to research work just like the one within the IEEE system. We have started by analyzing the various entities involved in this environment, such as publications, authors, conferences, and institutions, to understand their roles and how they interact with each other.

Using an Entity-Relationship Diagram (ERD), we mapped out these entities and their relationships to capture the system's complexity. The ERD is guiding us in designing a relational schema that includes all the necessary attributes and keys to ensure data integrity and efficiency.

1.2 Entity-Relationship Diagram

We have utilized the online tool(Lucidchart) to create our Entity-Relationship diagram from scratch. We have attached the screenshot of the ER diagram and the link to it as well:

(Link: https://lucid.app/lucidchart/17f8773e-383e-469b-a448-a9698e210e19/edit?viewport_loc=-50%2C199%2C1900%2C896%2C0_0&invitationId=inv_b4314af0-027c-4220-afd4-52045217fb0d)



1.3 Relation Schema

The following are the entities and the attributes that correspond to them:

1. **Publications** (publication_ID, author_ID, conference_ID, keyword_ID, paper_name, publisher, DOI, date_of_conference, date_of_publication, print_ISSN, print_ISBN)
2. **Authors** (author_ID, institution_ID, topic_ID, author_name, no_of_publications, no_of_citations)
3. **Institutions** (institution_ID, inst_name, inst_addr, inst_web_link, inst_phone_no, associated_authors)

4. **Keywords** (keyword_ID, topic_ID, keyword, kw_pub_count)
5. **Topics**(topic_ID, topic_name, topic_pub_count)
6. **Citations**(citation_ID, citing_pub, cited_pub, cited_paper_links, citation_count)
7. **Conferences**(conference_ID, conference_name, conference_year, conference_location)
8. **Users**(user_ID, user_name, user_email, user_contact_no, sub_status, subscription, profession)
9. **Members**(member_ID, user_ID, institution_ID, roles, memshp_status, memshp_start_date) [derived attribute: memshp_dur]
10. **Author_Affiliation** (author_id, aff_start_date, aff_end_date) [weak entity]

1.4 Entity Description

Entity	Description	Primary Key	Foreign Key
Publications	These are the academic papers or articles published within the IEEE system. This entity keeps key information about each publication, like the title, authors, conference details, and identifiers such as DOI, ISSN, and ISBN.	publication_ID	author_ID, conference_ID, keyword_ID
Authors	This holds details about the people who have written the publications. It includes their affiliations, research topics, and metrics like how many papers they've published and how often they've been cited.	author_ID	institution_ID, topic_ID
Institutions	Represents the universities or research organizations the authors are linked to. It has details like the institution's name, address, contact info, and which authors are associated with it.	institution_ID	
Keywords	Contains keywords related to publications. These helps categorize the papers and make them easier to find by topic. Each keyword is connected to specific research topics.	keyword_ID	topic_ID

Topics	These are the broader research areas or subjects that group publications and keywords. This helps organize and find papers based on specific academic fields.	topic_ID	
Citations	Tracks which publications cite others. This is useful for analyzing the impact and influence of different research papers.	citation_ID	
Conferences	Includes information about the conferences where the papers are presented, like the conference name, year, and location.	conference_ID	
Users	Represents everyone who interacts with the IEEE system, like researchers, students, and professionals. It includes their personal details and subscription status.	user_ID	
Members	A specific group of users who have certain roles or memberships within the IEEE system. This tracks their membership status, roles, and duration.	member_ID	user_ID, institution_ID
Author_Affiliation	A weak entity that logs the history of authors' affiliations with institutions. It includes start and end dates, linking authors to the institutions they've been associated with over time.	author_ID + aff_start_date	author_ID

1.5 Business Rules/ Relationship Explanations

The following is the description of Business Rules for the related entities. We have described the cardinality and participation of each one of them in the relational schema.

1. **Publications & Authors**

Cardinality: Many-to-Many

Participation: Publication = 1 (total), Authors = 1 (total)

Business Rule: Each publication must have at least one author. An author can write multiple publications.

2. **Publications & Conferences**

Cardinality: Many-to-One

Participation: Publications = 0 (partial), Conferences = 0 (partial)

Business Rule: Each publication is presented at one conference. A conference can hold multiple publications.

3. **Publications & Keywords**

Cardinality: Many-to-Many

Participation: Publications = 1 (total), Keywords = 0 (partial)

Business Rule: A publication can have multiple keywords. A keyword can be associated with multiple publications.

4. **Publications & Citations**

Cardinality: Many-to-Many

Participation: Publications = 1 (total), Citations = 1 (total)

Business Rule: A publication can cite many other publications. Similarly, citation of a publication can be done in many other publications.

5. **Authors & Institutions**

Cardinality: Many-to-One

Participation: Authors = 1 (total) and Institutions = 0 (partial)

Business Rule: An author is affiliated to at most one institution at the present time. But an institution can have many authors at any given time.

6. **Institutions & Members**

Cardinality: One-to-Many

Participation: Institutions = 1 (total) and Members = (total)

Business Rule: An institution can have many members. A member is usually associated to one institution at the given moment.

7. **Topics & Keywords**

Cardinality: One-to-Many

Participation: Topics = 1(total) and Keywords = 1 (total)

Business Rule: A topic can have many keywords. One keyword belongs to at most one particular topic.

8. **Users & Members**

Cardinality: One-to-One

Participation: Users = 0 (partial) and Members = 1 (total)

Business Rule: One user may be a member. Each member should be a user.

9. **Authors & Author_Affiliation**

Cardinality: One-to-Many

Participation: Authors = 1 (total) and Author_Affiliation = 0 (partial)

Business Rule: An author can have several affiliations over the duration of their research work. One affiliation points to a particular author only.

1.6 Bridge Entity, Weak Entity, and Complex Attributes Description

1. Bridge Entity:

a.) Pub_Key

Description: There is a many-to-many relationship between Publications and Keywords entities. Thus, to deal with that, we have introduced a bridge entity, Pub_Key such that the combination of primary keys from Publications and Keywords i.e. publication_ID

and keyword_ID becomes the composite primary key for Pub_Key bridge entity. Right now, we don't need any other attributes in these bridge entities. However, if needed, we will make the necessary changes to the schema design.
Schema: **Pub_Key (publication_ID, keyword_ID)**

b.) Pub_Cite

Description: There is a many-to-many relationship between Publications and Citations entities. Thus, to deal with that, we have introduced a bridge entity, Pub_Cite such that the combination of primary keys from Publications and Citations i.e. publication_ID and citation_ID becomes the composite primary key for Pub_Cite bridge entity.
Schema: **Pub_Cite (publication_ID, citation_ID)**

c.) Pub_Author

Description: There is a many-to-many relationship between Publications and Authors entities. Thus, to deal with that, we have introduced a bridge entity, Pub_Author such that the combination of primary keys from Publications and Authors i.e. publication_ID and author_ID becomes the composite primary key for Pub_Author bridge entity.
Schema: **Pub_Author (publication_ID, author_ID)**

2. Weak Entity:

a.) Author_Affiliation

Description: The authors are affiliated to various institutions over the period of time and this entity is introduced to keep track of this record. It holds the details of when a particular author began associating with an institution and when it ended. However, without the particular author, the affiliation won't be present. So it is dependent on Authors entity for its existence and thus, is a weak entity.
Schema: **Author_Affiliation (author_id, aff_start_date, aff_end_date)**

3. Complex Attributes:

a.) Composite Attribute: From the entity Authors, we have a composite attribute called **author_name (author_fname, author_lname)** which can be further elaborated as the author's first and last name. Another composite attribute is from the Institutions entity. It is the institution address that can be further divided into the street name, city, and so on. **inst_addr (inst_street, inst_city, inst_state, inst_postal_code)**

b.) Multivalued Attribute: The Institutions entity has an attribute called **inst_contact_no** wherein one institution can have several phone numbers associated with them. The **roles**

attribute in the Members entity is also multivalued. A member can be a user, author, or even a committee member.

c.) **Optional Attribute:** In the **Users entity**, we have an optional attribute of **user_contact_no** wherein it is the users' preference if they want to share their phone number or not.

d.) **Derived Attribute:** Among the **Members entity**, we have a derived attribute called **memshp_dur** i.e. the duration of membership that each of them have been a part of the IEEE community. It can be calculated with the help of another attribute in the entity i.e. memshp_start_date.