

IS5102 - Database Management Systems

Assignment 1

Task 1

1.1

1. Intended Use of the Data and Relevant Factors:

a. Staff Information:

- Unique staff ID, name, job title, email, postal address, phone numbers, status (full-time/part-time).

b. Department Information:

- Unique abbreviated name, full name, email, website, postal address, phone numbers.

c. Research Project Information:

- Unique project ID, project name, total budget, start date, end date, funding body, and GBP/non GBP (exchange rate) currency.

d. EU-Funded Projects:

- Task number unique within a specific project, and task description.

e. Employment Information:

- Unique employment ID, role, percentage of time spent working on a particular project.

f. Monthly Reporting Information:

- Unique report ID, unique project ID, month, days worked on each task.

2. Stakeholder Categories:

- 1) The Research Funding Office
- 2) University Departments
- 3) Staff Members
- 4) The Payroll Department

a. The Research Funding Office:

- *Query 1:* “List all active research projects along with their total budgets and funding bodies.”

- *Query 2*: “Provide a summary of staff members' roles and their total time allocation in percentage across all projects.”

- *Query 3*: “List all staff members who are currently working on EU-funded projects and the tasks assigned to them.”

b. University Departments:

- *Query 1*: “Retrieve a list of staff members employed by our department, including their contact information.”

- *Query 2*: “Find all ongoing research projects within our department along with their start and end dates.”

- *Query 3*: “Provide two lists of our full-time and part-time staff members.”

c. Staff Members:

- *Query 1*: “What are my current project tasks, roles, and the percentage of time I'm allocated to each project?”

- *Query 2*: “Show me the tasks I need to report for this month on my part-time research projects.”

- *Query 3*: “List all projects in which I have ever worked, provide a summary of my roles in each, and departments in which my research projects were conducted.”

d. The Payroll Department:

- *Query 1*: “List all staff members and days they worked each month throughout this year.”

- *Query 2*: “Provide a summary of all staff members, their projects and tasks for the month of December.”

- *Query 3*: “Find all staff members who worked more than 20 days on project tasks in summer.”

The aforementioned queries and factors should serve as a starting point for designing the database schema and ensuring that it can efficiently support the needs of the various stakeholders involved in managing research projects and staff in the university.

1.2 Specification

Based on the scenario described, we can create a specification for the relational database. We will break down the database into bullet point lists and describe the entity sets, attributes, relationships, and constraints. Here is the specification:

STRONG ENTITY SETS

1. Staff Members Table:

Attributes:

- Staff ID (Primary Key)
- Name
- Job Title
- Email
- Postal Address
- Phone numbers (multivariate attribute)
- Full-time and part-time entity sets (total disjoint specialisation)

Constraints:

- Staff ID must be unique.
- One or more phone numbers should be expressed with {} since they are multivariate attributes.
- Full-time and Part-time entity sets must be linked to the Staff member entity set via total disjoint specialisation.

2. Departments:

Attributes:

- Abbreviated Name (Primary Key)
- Full Name
- Email
- Website
- Postal Address
- Phone numbers (multivariate attribute)

Constraints:

- Abbreviated Name must be unique.

- One or more phone numbers should be expressed with {} since they are multivariate attributes.

3. Research Projects:

Attributes:

- Project ID (Primary Key)
- Project Name
- Total Budget
- Start Date
- End Date
- Funding Body
- GBP and non GBP (exchange rate) currency entity sets (total disjoint specialisation)

Constraints:

- Project ID must be unique.
- GBP and non GBP entity sets must be linked to the Research Project entity set via total disjoint specialisation.

4. Employment:

Attributes:

- Employment ID (Primary Key)
- Role
- Percentage of Time

Constraints:

- Employment ID must be unique.
- Percentage of Time must be between 0 and 100.

5. Monthly Work Reports:

Attributes:

- Report ID (Primary Key)
- Project ID (Foreign Key)
- Month
- Days Worked

Constraints:

- Report ID must be unique.
- Project ID must appear in another entity set too.
- Report ID and Project ID are a composite key.
- Days Worked must be a non-negative integer higher than 0.

WEAK ENTITY SETS

6. EU Projects:

Attributes:

- Task Number (Discriminator)
- Task Description

Constraints:

- Weak entity set must contain a discriminator – unique Task Number.

RELATIONSHIPS

1. Staff Member – Department (Many-to-Many relationship, total – Staff member side).
2. Department – Research Project (Many-to-Many relationship, total – Research Project side).
3. Part-time & Full-time – Staff member (total disjoint specialisation).
4. GBP & non GBP – Research Project (total disjoint specialisation).
5. Part-time – Monthly Work Reports (Many-to-Many relationship, total – both sides).
6. Monthly Work Reports – EU-Funded Projects (Many-to-Many relationship, total – Monthly Work Reports side).
7. EU Projects – Research projects (Many-to-One relationship, total – EU Projects side, double diamond).
8. Employment – Staff member (Many-to-Many relationship, total – both sides).

9. Employment – Department (Many-to-Many relationship, total – Employment side).
10. Employment – Research Project (Many-to-Many relationship, total – both sides).

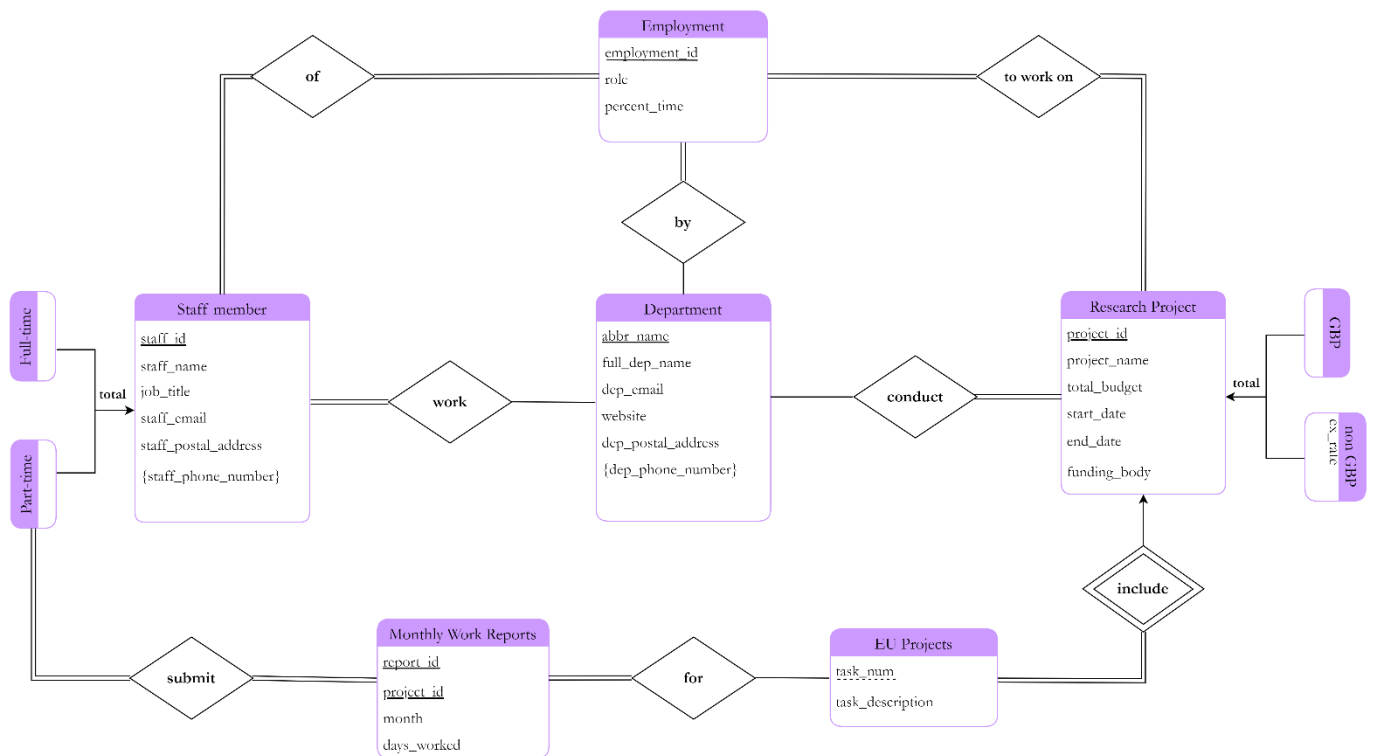
ASSUMPTIONS

- Each staff member has a unique Staff ID.
- Each department has a unique Abbreviated Name.
- Each research project has a unique Project ID.
- Each monthly report has a unique Report ID.
- Each Employment agreement has a unique Employment ID.
- Each task within an EU-funded project is uniquely identified by Task Number within that project.
- Days Worked in the Monthly Work Reports should only be recorded when greater than zero.
- The exchange rate is recorded only for projects which budget is not in GBP.
- The staff can work on multiple projects simultaneously with different roles and in different departments.

This specification forms the basis for creating an Entity-Relationship (E-R) diagram and defining the database schema in SQL or any other database management system. It outlines the structure of the database, relationships between entity sets and constraints on the data to ensure data integrity and consistency.

Task 2

2.1 E-R Diagram for the University Lab



2.2 E-R Diagram Validation

a. The Research Funding Office:

- *Query 1*: "List all active research projects along with their total budgets and funding bodies."

Data Source: The relevant data can be found in the *Research Project* entity set.

Explanation: The *Research Project* entity set stores information about total budgets and funding bodies in *total_budget* and *funding_body* entities.

- *Query 2*: "Provide a summary of staff members' roles and their total time allocation in percentage across all projects."

Data Source: The relevant data can be obtained from the *Employment* entity set.

Explanation: The *Employment* entity set contains information about roles and total time in percentage which are recorded in *role* and *percent_time* entities. The data about staff members (*staff_id*) and projects (*project_id*) can be obtained from the *Staff member* and *Research Project* entity sets since both of them are connected to the *Employment* entity set.

- *Query 3*: "List all staff members who are currently working on EU-funded projects and the tasks assigned to them."

Data Source: The necessary data is recorded in the *Part-time* and *EU Projects* entity sets.

Explanation: Each member of staff who works on EU-funded projects is a part-time worker thus the list of staff members can be obtained from the *Part-time* entity set. The *EU Projects* entity set contains *task_num* attribute where information about tasks assigned to staff members is stored.

b. University Departments:

- *Query 1:* “Retrieve a list of staff members employed by our department, including their contact information.”

Data Source: The data required is present in the *Staff member* and *Department* entity sets.

Explanation: The *Staff member* entity set records *staff_id*, *staff_email*, and *{staff_phone_number}* entities while the *Department* entity set is connected to the *Staff member* entity set and contains information about departments stored in *abbr_name* and *full_dep_name* entities.

- *Query 2:* “Find all ongoing research projects within our department along with their start and end dates.”

Data Source: The necessary data is located in the *Research Project* entity set.

Explanation: The *Research Project* entity set records *project_id*, *start_date*, and *end_date* entities.

- *Query 3:* “Provide two lists of our full-time and part-time staff members.”

Data Source: The relevant data can be found in the *Full-time* and *Part-time* entity sets connected to the *Staff member* entity set.

Explanation: *Full-time* and *part-time* researchers are divided into two separate entity sets and connected to the *Staff member* entity set.

c. Staff Members:

- *Query 1:* “What are my current project tasks, roles, and the percentage of time I'm allocated to each project?”

Data Source: The data required is present in the *EU Projects* and *Employment* entity sets.

Explanation: The *EU Projects* entity set stores information about project tasks in *task_num* entity while the *Employment* entity set records information about roles and percentage of time spent on projects in *role* and *percent_time* entities.

- *Query 2:* “Show me the tasks I need to report for this month on my part-time research projects.”

Data Source: The relevant data can be obtained from the *Monthly Work Reports* entity set.

Explanation: The *Monthly Work Reports* entity set contains information about the months which are recorded in *month* entity and is linked to *EU Projects* entity set with tasks in *task_num* entity.

- *Query 3:* “List all projects in which I have ever worked and provide a summary of my roles in each, and departments in which my research projects were conducted.”

Data Source: The data required is present in the *Research Projects*, *Department*, and *Employment* entity sets.

Explanation: Since the *Research Projects*, *Department*, and *Employment* entity sets are interconnected, a staff member can obtain data about *project_id*, *addr_name*, and *role* using those three entity sets.

d. The Payroll Department:

- *Query 1:* “List all staff members and days they worked each month throughout this year.”

Data Source: The necessary data is stored in the *Month Work Reports* and *Staff members* entity sets.

Explanation: The *Month Work Reports* entity set records *month* and *days_worked* entities and is connected to the *Staff member* entity set where *staff_id* is stored.

- *Query 2:* “Provide a summary of all staff members, their projects, and tasks for the month of December.”

Data Source: The data needed is present in the *Monthly Work Reports*, *Staff member*, and *EU Projects* entity sets.

Explanation: The *Monthly Work Reports* entity set contains data about projects and months in *project_id* and *month* entities. Since the *Monthly Work Reports* entity set is linked to the *Staff member* and *EU Projects* entity sets, we can access *staff_id* and *task_num*.

-*Query 3*: “Find all staff members who worked more than 20 days on project tasks in summer.”

Data Source: The relevant data can be found in the *Monthly Work Reports* and *Staff member* entity sets.

Explanation: The *Monthly Work Reports* entity set records the necessary data in *month*, and *days_worked* entities. The *Monthly Work Reports* entity set is also connected to the *Staff member* entity set with its *staff_id* attribute.

All in all, the E-R diagram satisfies the specification requirements because it organises and relates the data in a way that allows us to retrieve the required information for each query. The entities and relationships within the data model align with the data needed to answer the specified queries effectively.

Task 3

Database modelling is a crucial step in designing a database system that accurately represents the real-world scenario. Reflecting on the process of creating the database model for the university lab's experimental data management system, here are some key observations:

What I Did Well:

Identifying Entities and Relationships: I reckon that I was successful in identifying the relevant entities (e.g., staff members, departments, research projects, etc.), and relationships between entity sets. This helped in creating a structured and logical E-R model.

Primary Keys and Attributes: I think I correctly identified primary keys for entity sets and defined attributes accurately. This is very important for data integrity and uniqueness.

Understanding Business Requirements: I tried to make sure I understood the business requirements thoroughly, which is essential for effective database modelling.

What I Found Challenging:

Complex Relationships: Handling complex relationships, especially in scenarios where a staff member can work on multiple projects, in different departments, and roles simultaneously, posed a challenge. Managing these relationships required careful consideration.

EU-Funded Project Tasks, Monthly Work Reports, roles, and percentage of time: The requirement to create these entity sets and attributes was a bit challenging. It required additional thought on how to structure this part of the database.

Problems Encountered and How I Tried to Resolve Them:

One problem encountered was ensuring that the E-R model accurately represents and meets the business requirements while maintaining data integrity. To address this, I reviewed the schema multiple times, sought feedback from peers, and made necessary adjustments.

Balancing simplicity and completeness was another challenge. While trying to capture all relevant details, I ensured that the model did not become overly complex.

Theory vs. Practical Application:

I did not face any huge gaps between theory taught in class and practical application. The assignment reflects the material taught in class, so I understood what I need to do, although there are still some questions that I am willing to discuss.

What I Would Do Differently:

Collaboration with peers during the modelling process can give valuable insights and alternative perspectives. I would like to have more opportunities for peer review, discussion, and group work to enhance the quality of the database model.

To sum it all up, database modelling is a dynamic and iterative process that requires a balance between theory and practical application. While it can be challenging, careful consideration of business requirements, relationships between entity sets, and constraints, along with ongoing review and refinement, can result in a well-designed and effective database system.