Sunny

About database management systems

Database management systems are specialized software to store and manage large scale data, and provides the ability to query data with optimized algorithms, and more.

2 main types of databases: relational and non-relational.

About relational databases and schemas
One key distinction between them is that relational databases have a **schema**, which outlines the **relations** of the database. In other words, the schema is just the collection of tables, with all the attributes (columns) the database will have. **Schemas** offer guarantees, but it requires careful design to be effective.

Why relational databases for us?
We imagine that our product prioritizes **accuracy** - students and researchers need the right connections. The emphasis is on the matching of people to projects. I think relational databases will serve us well regarding our priority.

How the database will be designed and implemented

1. I need to start talking to potential users about the data they'd expect the system to have. I need an English description first.
2. Then, I'll spend a lot of time drawing an ER Diagram.
3. Diagram will eventually give us the **schema.**
4. From schema, I'll implement the database system
5. Use cases collected by the project will inform development of commonly used queries
6. Testing data insertion and querying. Testing integration of database into website environment
7. Be awesome

Why schema first?

Knowing and designing a good schema early on is quite important. On one hand, besides the benefits of schemas, the schema outlines the database itself - without schema, I don't know what to put in the database. On the other hand, relational schemas tend to organize data in unnatural (normalized) ways - at least that's how I wanna do it - and thus data collection and insertion into the database can be made easier if we know what to look for.

PHASE I: collecting information for database

I need to know what people want in a database that stores research project information.

There are 3 user groups: professors, graduate students and undergraduate students. Our product connects them to projects. Actually, I think we can group up undergraduate and graduate students into 1 student group.

So we should ask some questions to each group.

The goal is to form an English description of the database. ***Keep in mind 2 things: Entities and Relationships between them. For more details, see entities_and_relationships document***

Questions we ask them

Professors:

- What research projects do they offer?
- Prerequisites for each project: sometimes coursework, sometimes experience -> hard to figure out database structure
- What are research projects about? How can we represent that in a database?
- Major? Available to graduate students only, undergraduates only, or both?
- Compensation?

Students

- What do they look for in a research project?
- Major, matching prerequisites, compensation, availability, compensation?