Overview of the Course
What is a Database ?

Databases and Database Management Systems.

A **Database** is an *organized* collection of data.

A **Database Management System (DBMS)** is ***software*** designed to maintain and support databases.

(Imaginary) FAQ

1. Why study databases?

- Databases are nearly ubiquitous.
- Management of data is a bottleneck for many applications.
- Efficient management of data requires special approaches.
- These special approaches are demonstratably successful.

2. Databases are ubiquitous?

- Think of the last time you interacted with a database. How long ago was it?

- How many different databases have you interacted with yesterday? How many different databases do you interact with during a single week?
- Name 10 different uses of databases.

3. Why is management of data such an issue?

- Sizes of data collections - both electronic and not.
- Organization of data collections: how do we make sure what we collect is *usable*?

4. So, what is management of data, after all?

The term *"management of data"* encompasses support for the following:

- description of the data collection organization format;
- storage of data in organized form;
- retrieval of data to serve user information needs.

5. What is a "user information need"?

Any question, which (a) has an answer, which (b) can be established from the data in the data collection is considered a *user information need*.

User information needs can be expressed in a number of different ways: in natural language, using special formal languages understood by DBMS, using mathematical notation...

6. How can we make data management efficient?

Efficiency cen be interpreted differently (everything works fast; data management is easy to do; data collections are stored in small amounts of memory, etc).

Generally speaking, we seek two properties:

- Uniformity: same methods of management for **any** data collection.

  > *We want to be able to support development of databases/database applications for **any** type of data collections. We want this process to be the same, and we want people who perform this process (database developers, database analysts, database application programmers) to be able to reproduce this process in any area of human activity.*

- Performance: ability to process data management tasks in real time.

  > *User information needs must be addressed in a timely manner. Different data collections and different information needs come with different estimates of what "timely" is, but in some scenarios "timely" does mean "instanteneous" or "almost instanteneous", and therefore our methodology must be able to perform tasks fast.*

To achieve both uniformity and performance management of data must be handled by special-purpose procedures!

7. Who works in the area of data management?

- Customer.

  Management of data is typically done for the benefit of non-computer scientists.

- Database analyst.

  Also known as knowledge engineer or database designer.

- Database applications programmer.

  The target audience of this course.

- Database Administrator.

  Not to be confused with the sysadmin.

- DBMS programmer.

  What compiler developer is to the world of programming languages, the DBMS programmer is to the world of data management.

- Database researcher.

  Mostly found in universities and large database corporations.

Can you give an outline of what each group of specialists does?

8. What topics constitute the study of management of data and databases, and where can we study them?

The table below provides a brief key to various aspects of data management, the "target audience" for these topics (i.e., which groups from the list above this information is primarily needed for), and the courses in the department's database sequence, where this information is typically taught.

Topic	Role	Course	Comment
Requirements Elicitation	DB Analyst	N/A	*mostly taught in SE courses*
Database Modeling	DB Analyst	CPE 366	
Relational Database Model	DB Apps Programmer DB Analyst, all but customers	**CPE 365**	
Querying Databases	DB Apps Programmer	**CPE 365**	
Database/DBMS Administration	DBA	CPE 468	mostly in the labs
DBMS Operating Environment	DB Apps Programmer DBA	**CPE 365**	more advanced material in later courses
Interface between DBMS and programming languages	DB Apps Programmer DBA	**CPE 365**	
Internal organization of DBMS	DBMS Programmer Researcher	CPE 468	
Query processing	DBMS Programmer Researcher	CPE 468	
Transaction Processing (multi-user operation)	DBMS Programmer Researcher	CPE 468 CPE 468	
Security, Recovery	DBMS Programmer	CPE 468	
Non-relational database models	Researcher DB Apps Programmer, Analyst	CSC 560 *CPE 366*	 *planned!*

In short:

CPE 365 studies *databases* and how to work with them.

CPE 366 studies *data* and how to build databases from it.

CPE 468 studies *DBMS* and what they are made of. It also covers aspects of *DBMS Administration.*

CSC 560 studies most recent efforts to make *data management* better.