# Main Phases of Database Design

In the picture below there are the main phases of database design. Database design is connected with application design.

The requirements and the collection analysis phase produce both data requirements and functional requirements. The data requirements are used as a source of database design. The data requirements should be specified in as detailed and complete form as possible.

In parallel with specifying the data requirements, it is useful to specify the known functional requirements of the application. These consist of user-defined operations that will be applied to the database (retrievals and updates). The functional requirements are used as a source of application software design. Of course some functions may produce also needs for database design.

Note that some phases are database management system independent and some are dependent. The idea is to design first the database without thinking about the actual database system - just to concentrate on the data.
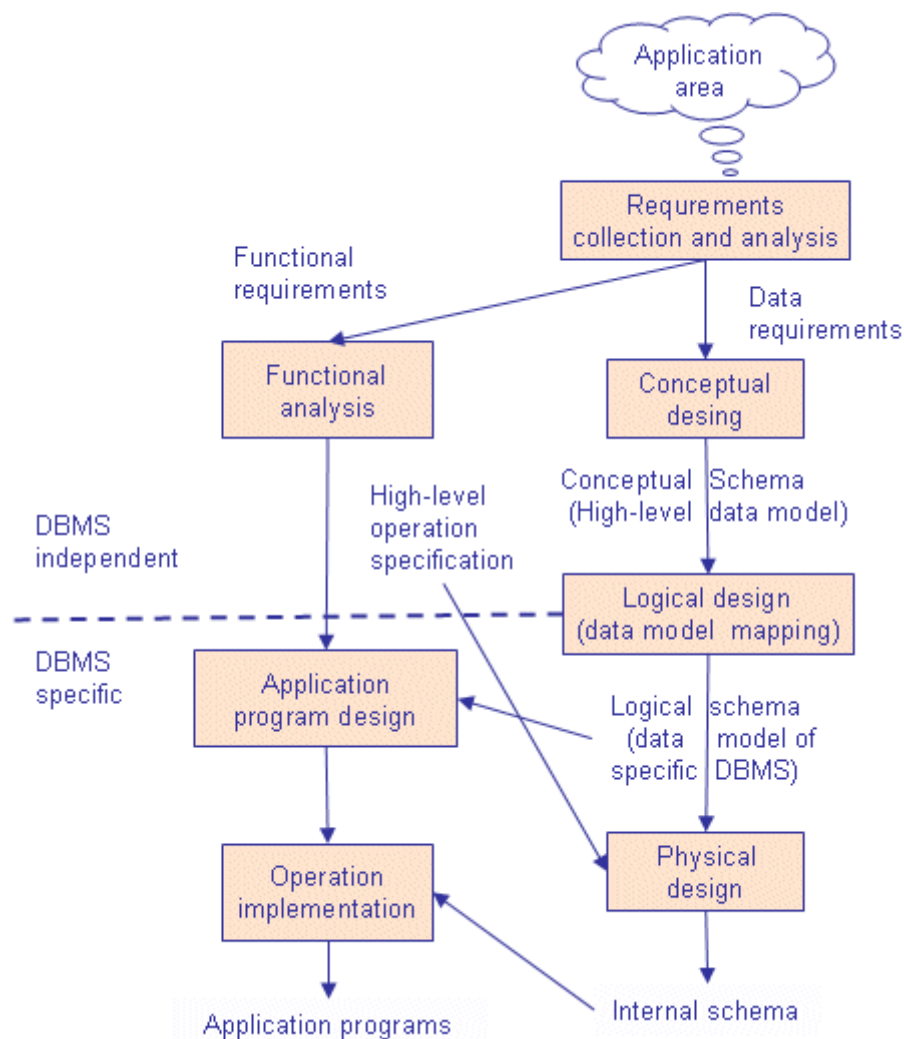
Conceptual Design
Once all the requirements have been collected and analyzed, the next step is to create a conceptual shema for the database, using a high level conceptual data model. This phase is called conceptual design.

The result of this phase is an Entity-Relationship (ER) diagram or UML class diagram. It is a high-level data model of the specific application area. It describes how different entities (objects, items) are related to each other. It also describes what attributes (features) each entity has. It includes the definitions of all the concepts (entities, attributes) of the application area.

During or after the conceptual shema design, the basic data model operations can be used to specify the high-level user operations identified during the functional analysis. This also serves to confirm that the conceptual schema meets all the indenfied functional requirements.

There are several notations to draw the ER diagram.

Logical Design
The result of the logical design phase (or data model mapping phase) is a set of relation shcemas. The ER diagram or class diagram is the basis for these relation schemas.

To create the relation shemas is quite a mechanical operation. There are rules how the ER model or class diagram is transferred to relation shemas.

The relation schemas are the basis for table definitions. In this phase (if not done in previous phase) the primary keys and foreign keys are defined.

Normalization
Normalization is the last part of the logical design. The goal of normalization is to eliminate redundancy and potential update anomalies.

Redundancy means that the same data is saved more than once in a database. Update anomaly is a consequence of redundancy. If a piece of data is saved in more than one place, the same data must be updated in more than one place.

Normalization is a technique by which one can modify the relation schema to reduce the redundancy. Each normalization phase adds more relations (tables) into the database.

Physical Design
The goal of the last phase of database design, physical design, is to implement the database. At this phase one must know which database management system (DBMS) is used. For example, different DBMS's have different names for datatypes and have different datatypes.

The SQL clauses to create the database are written. The idexes, the integrity constraints (rules) and the users' access rights are defined.

Finally the data to test the database is added in.

In parallel with these activities, application programs are designed. The implementation of the programs can start when the database is created and data has been added in.