

Definitions

Data

- **Known facts stored and recorded**

Information

- **Data presented in context**
- **Data that has been processed increasing users knowledge**
- Data is known and available, information is processed and more useful

Knowledge

- **Human understanding of a subject matter that has been acquired through proper study and experience**
- Knowledge can be derived from information the same way as information is derived from data

Data Definition Language (DDL)

- Define and set up database

Data manipulation language (DML)

- Maintain, use database

Data Control Language (DCL)

- Control access to database

Meta data

- Structure, rules, constraints, definitions
- Needed for consistency, meaning

Database Management System (DBMS)

- Software allowing definition and control
- Allows **CRUD (Create, read, update, delete)**
- other commands:
 - administer database
 - BACKUP TABLE, RESTORE TABLE, ANALYZE TABLE
 - transaction control
 - Misc
 - DESCRIBE tablename
 - USE db_name

Databases

- **Collection of logically related data and its description**
- Integrity, concurrency (shared access), recovery, data descriptions
- **Advantages**
 - **Data independence**
 - Separation of data and program, application logic
 - Central data repository, central management
 - **Minimal data redundancy**
 - Redundancy can be controlled (normalisation)
 - **Improved data consistency**
 - Single store, no disagreements, update problems, less storage space
 - **Improved data sharing**
 - Multiple, arbitrary views

- Can allow access to external uses
- **Reduced program maintenance**
 - Data structure can change without application data changing
 - Views can be maintained despite internal physical structure changes
- **Increased productivity of application development**
 - Data already collected and structures already known
 - DBMS provides many tools to access, manipulate data
- **Enforcement of standards**
 - Centralised data management
 - Documented policy for data management
 - Data definition and dictionary (metadata)
- **Improved data quality**
 - Constraints built into database
 - 'scrubbing' data if necessary
- novel ad hoc data access 'without programming' ie SQL
- **Drawbacks**
 - **Difficulty in analysing and designing good databases**
 - **Conversion costs**
 - **Not that easy to manage**
 - **Organisational planning/conflict**
 - **Backup, recovery strategy crucial**

C/f file systems

- Each program access a different file
- Problems
 - **Program-data dependence**
 - Hard to change
 - **Duplication of data**
 - Wasteful, inefficient, loss of data integrity
 - Loss of metadata integrity (same name different data, same data different name)
 - **Limited data sharing**
 - Tied to application, hard to create ad hoc reports
 - **Lengthy development times**
 - App has to do low level data management
 - **Excessive program maintenance**

Integrity constraints

- Referential Integrity
 - Each value of a non null foreign key must match a value of the primary key in the related table
 - Can use SET foreign_key_checks=0
 - Rules for update and delete (SQL CREATE statement)
 - **ON DELETE RESTRICT**
 - Don't allow deletes or updates of the 'many' side of a one-to-many relationships if the related rows exist in the 'one' side
 - **ON UPDATE CASCADE**

- Automatically delete/update the 'many' side of a one to many relationship if related rows are affected in the one side
- Data Integrity constraints
 - Data field integrity
 - Valid values and domain
 - Selection of data type is the initial constraint on the data
 - Default value
 - Range control
 - Allowable values limitation
 - Null value control
 - Entity Integrity constraints
 - Primary key cannot be null
 - No component of a composite key can be null
 - Primary key must be unique

Relation (in a relational database): corresponds to a table. Cf relationship

Database Development Lifecycle

- **Database planning**
 - How to do project
 - How does enterprise work
 - Enterprise data model
 - Planning other stages
 - Outside scope
- **Systems definition**
 - Scope and boundaries, major views, users, application areas, how it fits into other organisational system
 - Slightly outside scope
- **Requirements definition and analysis**
 - Develop a data dictionary to be extended in design
- **Conceptual**
 - Construction of a model of the data used in database – independent of all physical considerations
 - Data models
 - ER, EER diagrams
- **Logical**
 - Construction of model of data based on conceptual design
 - Independent of a specific database and other physical considerations
- **Physical design**
 - Description of implementation of logical design for a specific DBMS
 - Describes
 - Basic relations
 - File organisations
 - Indexes
 - Integrity constraints