Lecturer's profile

Name: កាក់ សុខគី (Kak Soky)

Education: PhD in Informatics on Automatic Speech Recognition

Experiences:

Researcher: 7 years (2016-2023)
Teacher: 7 years (2012-2019)
Developer: 4 years (2011-2014)

Contact:

E-mail: soky.kak@cadt.edu.kh
Telegram: +85585683680
Site: https://ksoky.github.io
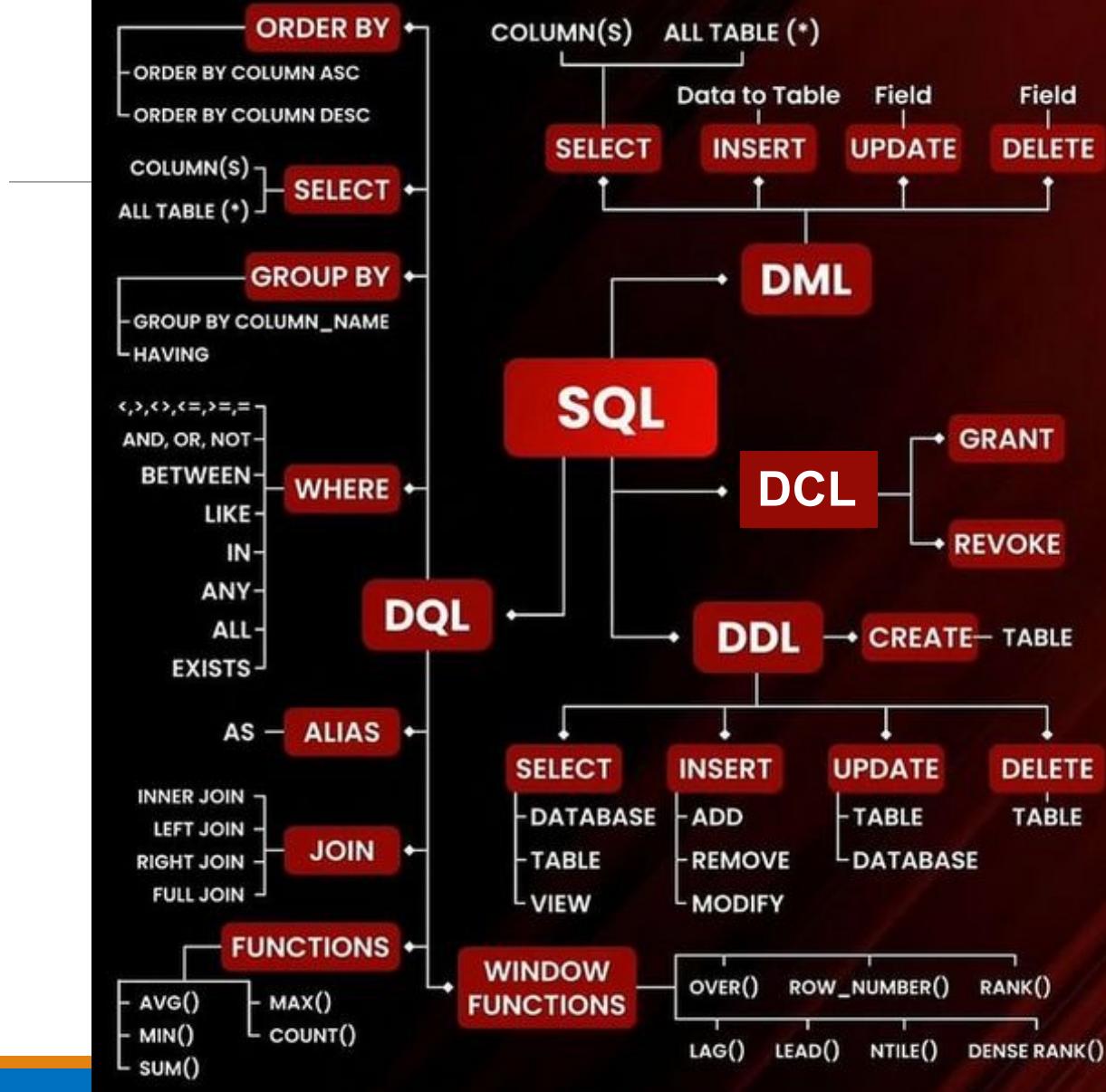
Done is a prior,
Perfection is a later.

Course Outlines

- ✓ Introduction to Database

- ✓ Entity Relationship Diagram

- ✓ Relational Model

- ✓ Database Normalization

- ✓ Structure Query Language (SQL)

- ✓ Project & Presentation

- ✓ Mid-term and Final Exam

Grading Criteria

Activity	Score
Attendance	10%
Case Studies & Assignments	20%
Project	30%
Midterm	20%
Final Exam (+ lab)	30%

Introduction to Database Management System (DBMS)

Outline

- Data and Information

- Database

- Database Systems

- What is DBMS?

- Purpose of DBMS

- History of DBMS

- Advantages and Disadvantages of Database System

- ...

Data
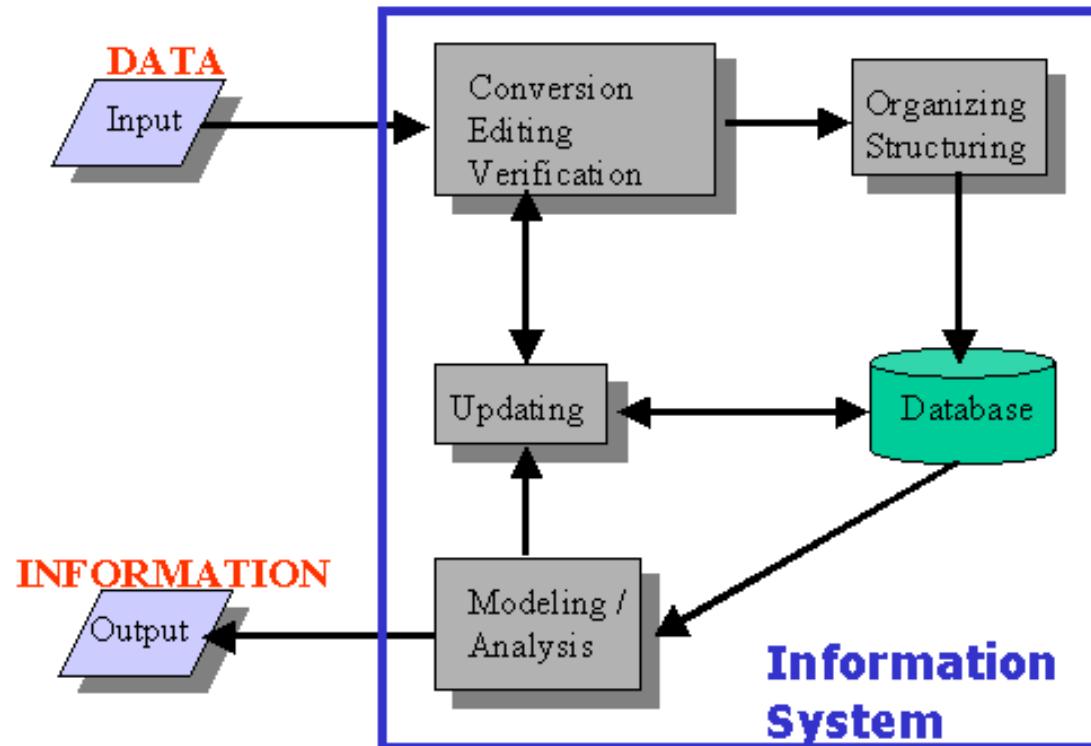
✓ "*Data*" is defined as a body of facts or figures, which have been gathered for one or more specific purposes

✓ data can exist in the forms of
- linguistic expressions (e.g. name, age, address, date, ownership)
- symbolic expressions (e.g. traffic signs)
- mathematical expressions (e.g. $E = mc^2$)
- signals (e.g. electromagnetic waves)

Information

✓ "*information*" is defined as data which have been processed into a form that is meaningful to a recipient. In other words, data has to be interpreted in order to obtain information.

✓ although data are ingredients of information, not all data make useful information

✓ information is only useful to its recipients when it is

- relevant (to its intended purposes and with appropriate level of required detail)

- reliable, accurate and verifiable (by independent means)

- up-to-date and timely (depending on purposes)

Information System

✓ "***information system***" is a system used to change "data" into "information".

Database

✓ A ***Database*** is a collection of related data organized in a way that data can be easily accessed, managed and updated.

✓ In a database, we would be grouping only related data together and storing them under one group name called table.

STUDENT
STUDENT_ID
STUDENT_NAME
ADDRESS
DOB
COURSE

EMPLOYEE
EMPLOYEE_ID
EMPLOYEE_NAME
EMPLOYEE_DOB
EMP_ADDRESS
EMP_PASSPORT_NUM
LICENCE_NUMBER
SSN
DEPARTMENT_NUMBER
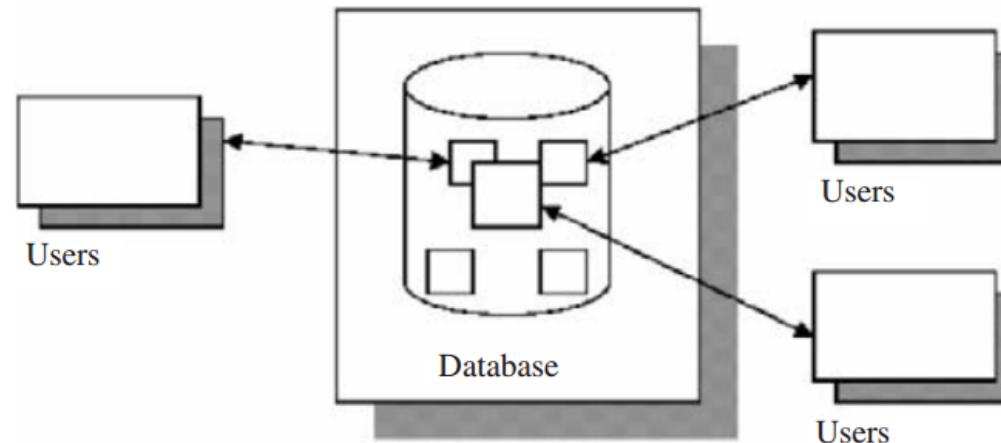
DEPARTMENT
DEPARTMENT_ID
DEPT_NAME

# Applications
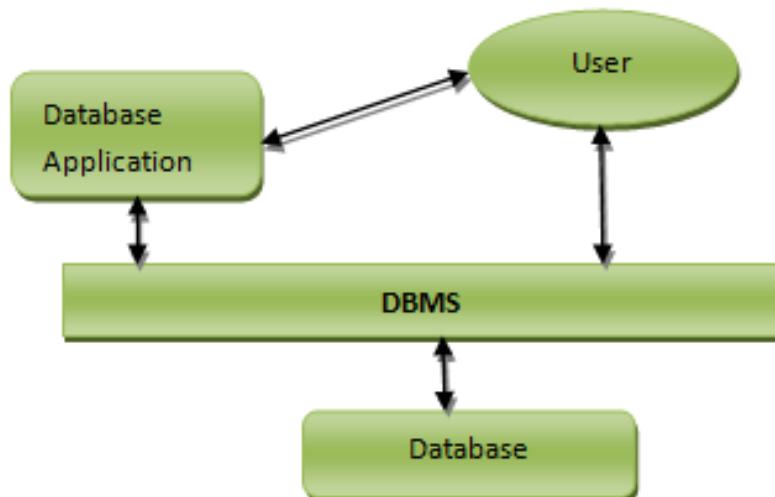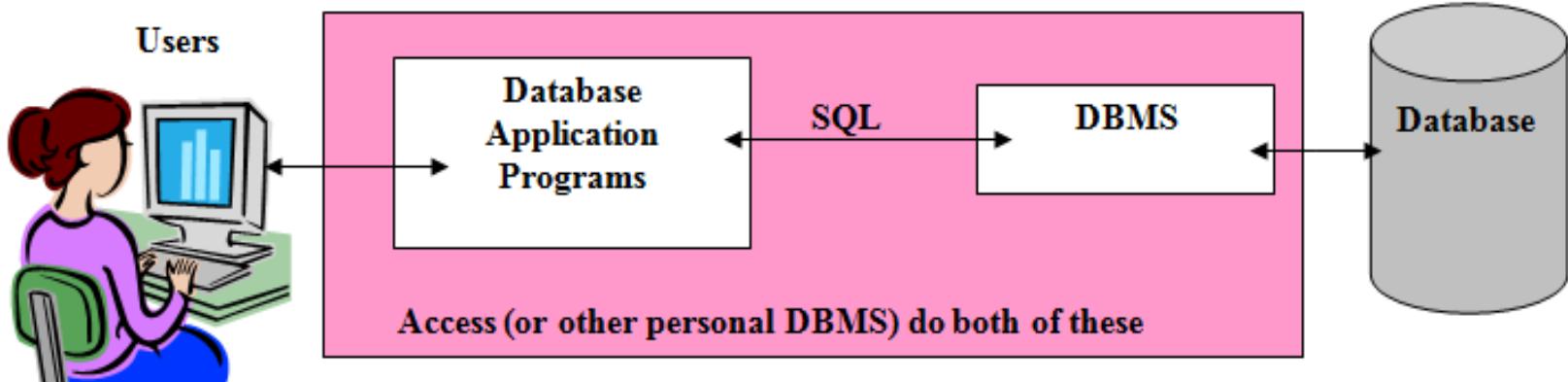
- ✓ Telecom
- ✓ Industry
- ✓ Banking System
- ✓ Education sector
- ✓ Online shopping
- ✓ Transportation
- ✓ Human Resource
- ✓ and so on…

Database Systems

✓ ***Database systems*** are systems in which the interpretation and storage of information are of primary importance.

✓ It is a system that allow users to Store, Update, Retrieve, Organize, and Protect their data.

Components of database system

Components (cont')

✓ **Users:** are the one who really uses the database. Users can be administrator, developer or the end users.

Tasks of users:
- ▪ Input new data
- ▪ Delete data
- ▪ Modify existing data
- ▪ Reads data

✓ **Database Application:** is the application program which helps the users to interact with the database by means of query languages. It serves as intermediary between the user and the DBMS.
Its task includes create/process forms, process user queries, executes application logic.

Components (cont')
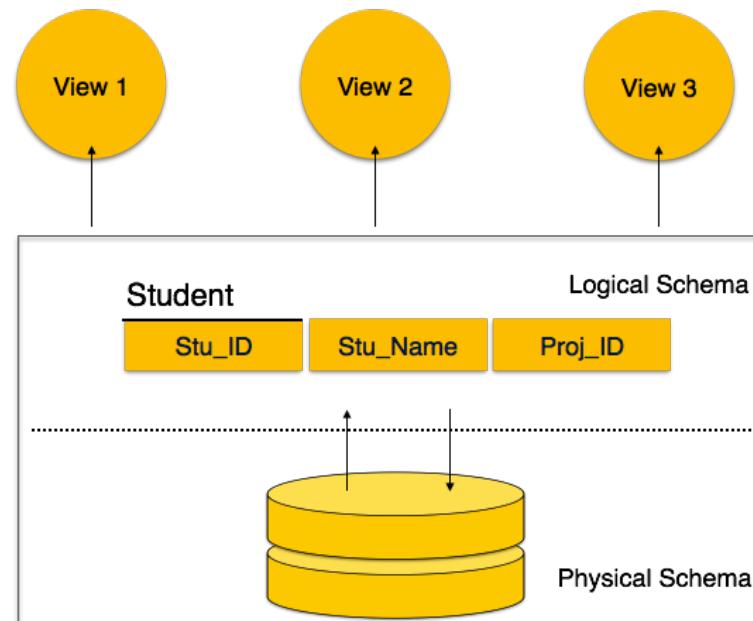
✓ **Database:** is a collection of related data organized in a way that data can be easily accessed, managed and updated.

Two types of data:

- ▪ User data: is data which is stored in the various tables of the database in the form of rows and columns
- ▪ Metadata: is data about data. it stores the information like how many tables and their names, how many columns and their names, primary keys, foreign keys etc.

✓ **DBMS:** is the software helps the user to interact with the database.

Database Schemas

✓ Database schema is the skeleton of database. It is designed when the database doesn't exist at all.

✓ A database schema does not contain any data or information.

Database Instance

- ✓ A database instance is a state of operational database with data at any given time.

- ✓ It contains a snapshot of the database.

- ✓ A DBMS ensures that its every instance (state) is in a valid state, by diligently following all the validations, constraints, and conditions that the database designers have imposed.

What is DBMS?
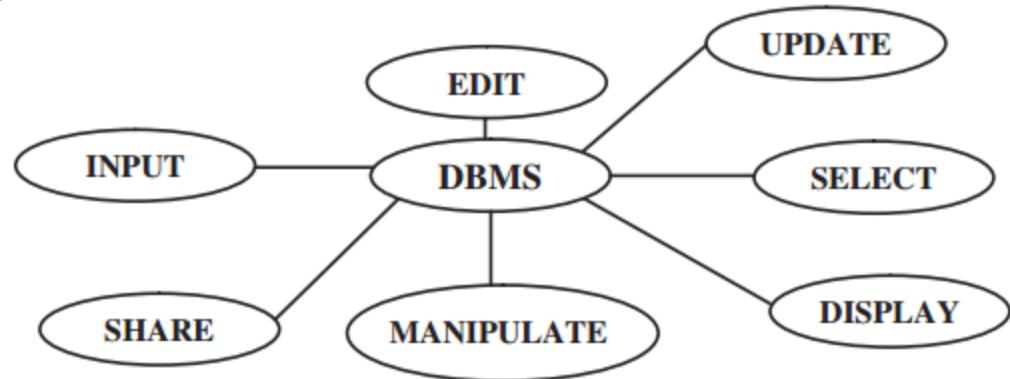
- ✓ DBMS stands for **D**ata**b**ase **M**anagement **S**ystem.

- ✓ We can break it like this DBMS = Database + Management System.

- ✓ Database is a collection of data and Management System is a set of programs to store and retrieve those data.

- ✓ Or DBMS is a collection of inter-related data and set of programs to store & access those data in an easy and effective manner.

What is the need of DBMS?

✓ Database systems are basically developed for large amount of data.

✓ When dealing with huge amount of data, there are two things that require optimization: **Storage of data** and **retrieval of data**.

Simply put, DBMS can:

- ✓ create db structure
- ✓ create table structures
- ✓ data retrieval
- ✓ data update
- ✓ maintain db structures
- ✓ ensure data integrity
- ✓ control concurrency/handle shared update
- ✓ ensure security
- ✓ backup/recovery
- ✓ catalogue/data dictionary (metadata)

History

- ✓ 1960's-1970's: The emergence of the first type of DBMS, the hierarchical DBMS.

- ✓ 1960's-1970's: The emergence of the network DBMS.

- ✓ 970's- 1990's: The emergence of the relational DBMS on the hands of Edgar Codd.

- ✓ NoSQL ("Not only" Structured Query Language) came about as a response to the Internet and the need for faster speed and the processing of unstructured data.

Example

Structured Data

Unstructured Data

Advantages
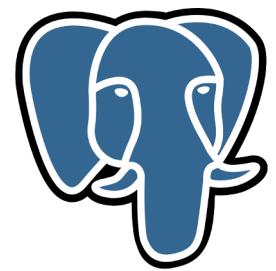
- ✓ No redundant data – Redundancy removed by data normalization

- ✓ Data Consistency and Integrity – data normalization takes care of it too

- ✓ Secure – Each user has a different set of access

- ✓ Privacy – Limited access

- ✓ Easy access to data

- ✓ Easy recovery

- ✓ Flexible

Disadvantages

- ✓ DBMS implementation cost is high compared to the file system

- ✓ Complexity: Database systems are complex to understand

- ✓ Performance: Database systems are generic, making them suitable for various applications, however this feature affect their performance for some applications

DBMS Packages

- ✓ MySQL/MariaDB
- ✓ Oracle
- ✓ SQL Server
- ✓ PostgreSQL, etc.
- ✓ *MongoDB (NoSQL)*
- ✓ Redis *(NoSQL)*
- ✓ ElasticSearch *(NoSQL)*
- ✓ Cassandra *(NoSQL)* , etc.

Relational DBMS
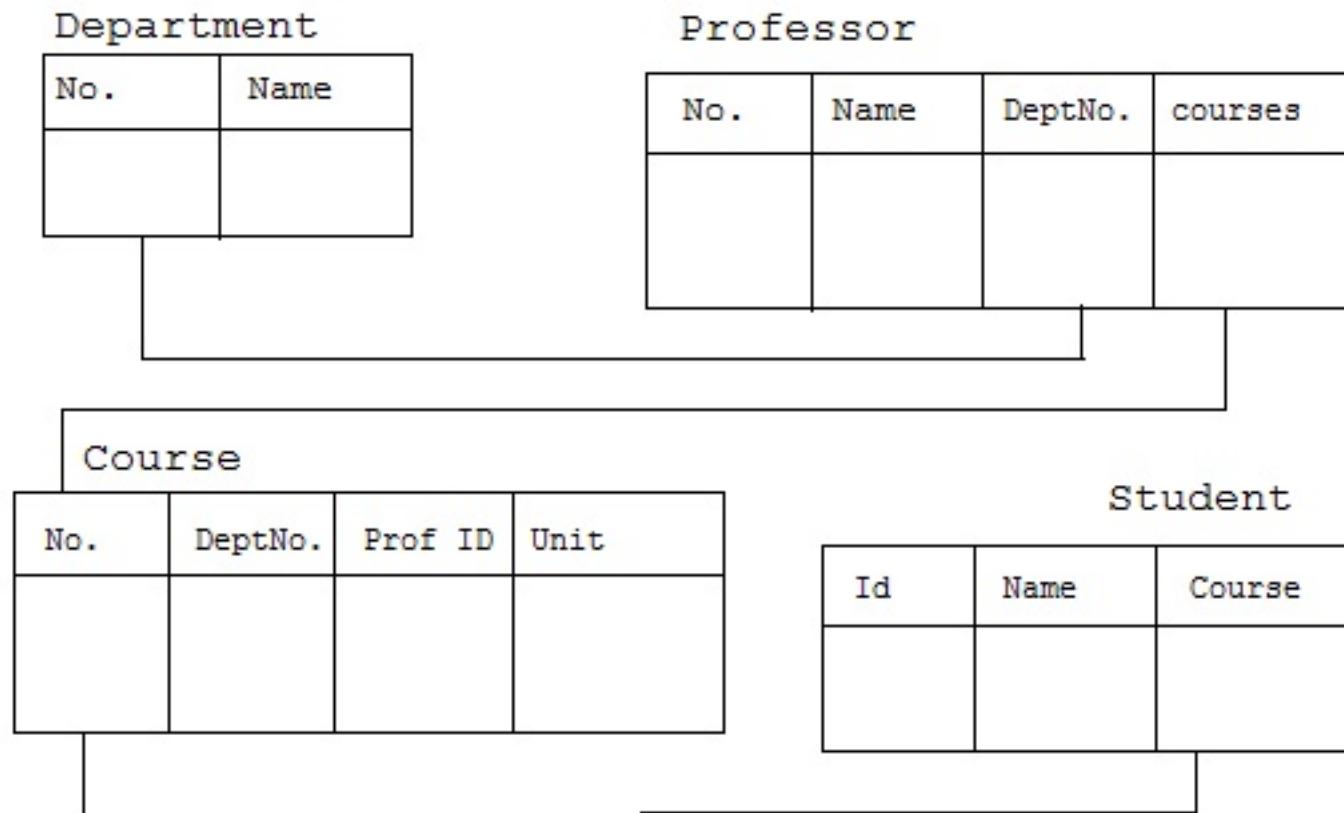
✓ In relational databases, the database is represented in the form of relations.

✓ A relational database has following major components: Table, Record / Tuple, Field & Column /Attribute.
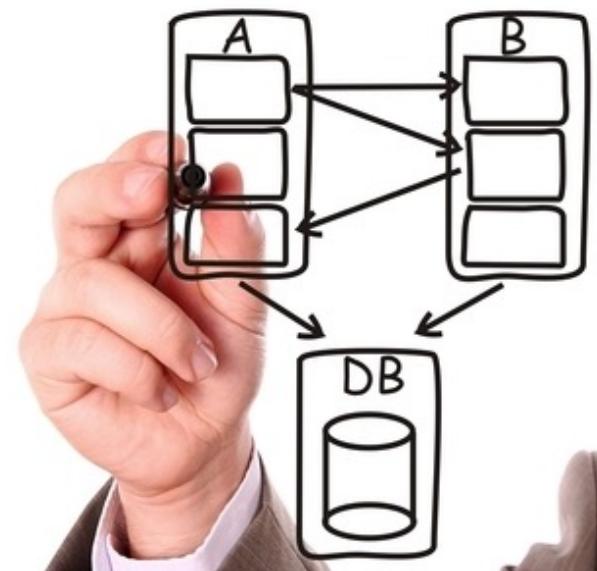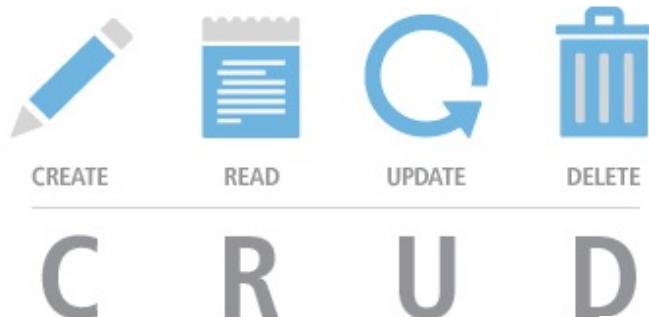
S_Id	Name	Year	Stream
1	Ankit Jha	1	Computer Science
2	Pushpa Mishra	2	Electronics
5	Ranjini Iyer	2	Computer Science

Field → (pointing to Name column)

Tuple → (pointing to rows)

Example
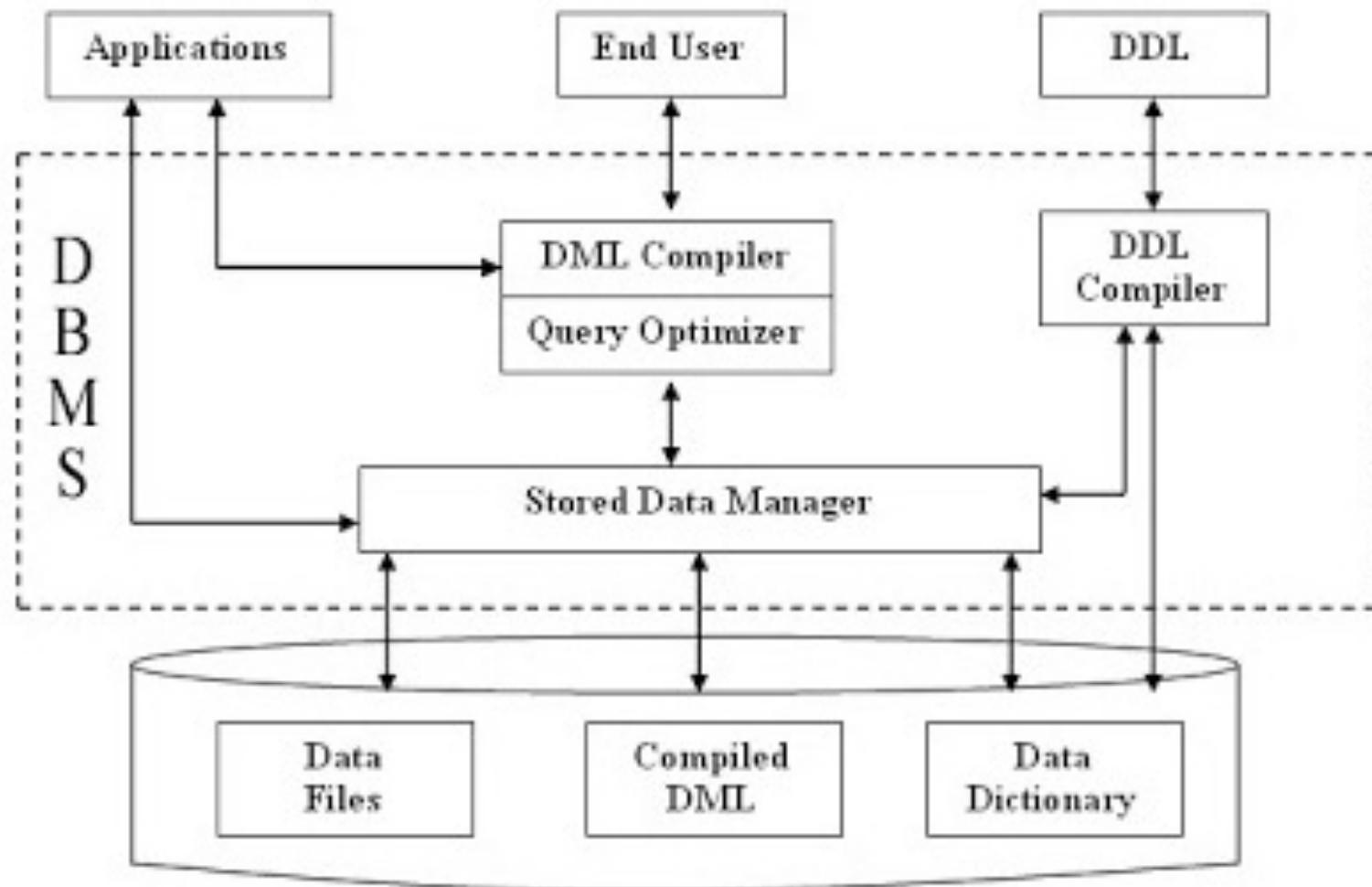
Operations on DBMS

- ✓ Create
- ✓ Retrieve
- ✓ Update
- ✓ Delete

Structure, components and interfaces of DBMS

Structure, components and interfaces of DBMS

1. **DDL Compiler** - **Data Definition Language compiler** processes schema definitions specified in the DDL. It includes metadata information such as the name of the files, data items, storage details of each file, mapping information and constraints etc.

2. **DML (Data Manipulation Language) Compiler** and **Query optimizer** - The DML commands such as insert, update, delete, retrieve from the application program are sent to the DML compiler for compilation into object code for database access. The object code is then optimized in the best way to execute a query by the query optimizer and then send to the data manager.

Structure, components and interfaces of DBMS

3. **Stored Data Manager -** The Stored Data Manager is the central software component of the DBMS also knows as Database Control System.

4. **Data Dictionary -** Data Dictionary is a repository of description of data in the database.

5. **Data Files -** It contains the data portion of the database.

6. **Compiled DML -** The DML complier converts the high level Queries into low level file access commands known as compiled DML.

Data Dictionary

Data Dictionary outlining a Database on Driver Details in NSW

Field Name	Data Type	Data Format	Field Size	Description	Example
License ID	Integer	NNNNNN	6	Unique number ID for all drivers	12345
Surname	Text		20	Surname for Driver	Jones
First Name	Text		20	First Name for Driver	Arnold
Address	Text		50	First Name for Driver	11 Rocky st Como 2233
Phone No.	Text		10	License holders contact number	0400111222
D.O.B	Date / Time	DD/MM/YYYY	10	Drivers Date of Birth	08/05/1956

Further readings

✓ http://www.dataversity.net/brief-history-database-management/

✓ http://kaurchanpreet.blogspot.com/2015/02/data-and-data-warehouse-scope.html

✓ http://dbmsnotes-ritu.blogspot.com/2015/08/overall-structure-of-dbms.html