Business Intelligence Technologies

Anna Monreale

Università di Pisa

General Information

- **Anna Monreale**
  - **Email:** anna.monreale@unipi.it
  - Website: http://didawiki.cli.di.unipi.it/doku.php/bdd-infuma/start

- **Learning Material**
  - Slides
  - Databases essentials – AntonioAlbano (pdf)
  - A First Course in Database Systems, JefferyD. Ullman and Jennifer Widom
  - Decision Support Database – Antonio Albano (pdf)

Overview of the course topics

Anna Monreale

Università di Pisa

Topics

- **Databases**
  - What is a DBMS
  - Designing a DB
  - Querying a DB
- **Business intelligence**
  - Decision Support Databases
  - Data warehouses
  - Designing a data warehouse
  - Querying a data warehouse

Information as a Resource

- **Information** is one of the most important resources of any **organization**

- An intelligent management of the information can help organization to generate **new knowledge**

- It becomes more and more important learning how to represent, organize, manage and use information

Organization

- **Organizations**
  - es: companies, banks, public administrations, ….

- An organic collection of **resources** (people, materials, **information**), **tools** and **procedures**, which are finalized to create and offer a product or a service
  - a bank provides financial services
  - a hospital supplies medical services

Information System

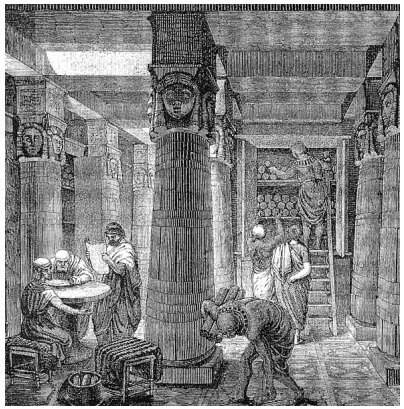
- Component of an organization finalized to the **information management** for supporting the organization activities
  - Collects, stores, processes and communicates the information

- Any organization has an Information System

Information System & Computerization

- **Information system** is indipendent on the computerization

- Organizations as Libraries, Banks, etc … existed before the computerization and managed information

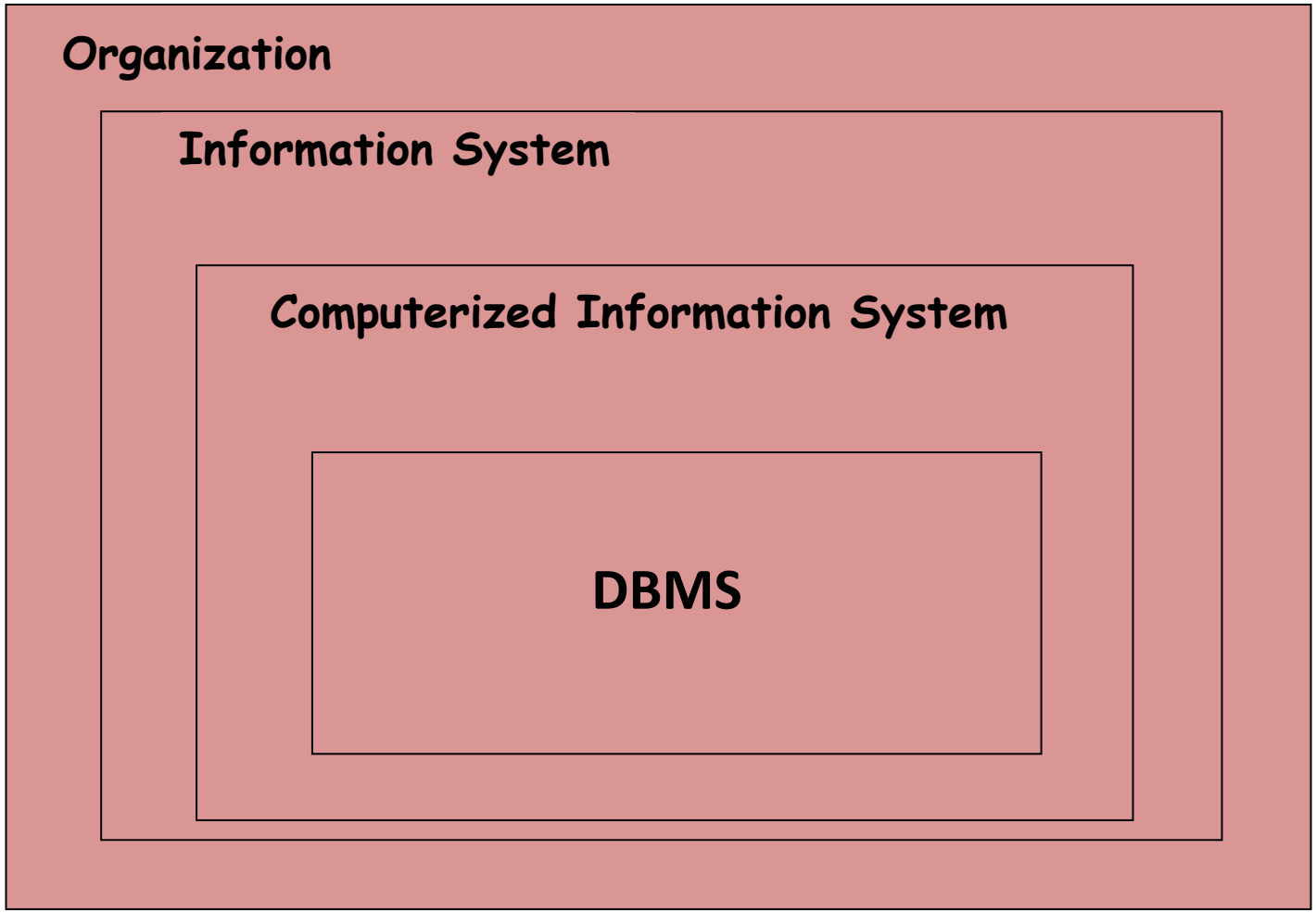


Real Library ofAlessandria d'Egypt

Computerized Information System

- The hardware and sofware used for storing, retrieving, and processing the information which supports the functions of an organization
- In practice, in many cases:
  - **information system** is used as synonym of **computerized information system**
  - But it is not always true
- Implemented by **database technology**
  - operational database and a collection of application programs used to access and update the data quickly and efficiently

Computerized Information System

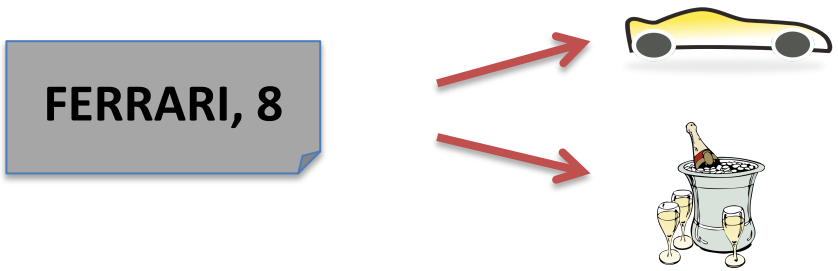


Goal of DBMS: processing data for getting information

Data vs Information

- In information systems **information** is represented by **data**
- **Information** (def): news, or element enabling knowledge about facts and situations
- **Datum** (def): element of an information composed of symbols to be processed

Without interpretation datum is useless



Database

- Collection of permanent data representing facts interesting for an organization (*data or istances*) and facts related to the data organization *(metadata or schema)*

STUDENTS

ID	Surname	Name	BirthofDate
276545	Rossi	Maria	25/11/1991
485745	Neri	Anna	23/04/1992
200768	Verdi	Fabio	12/02/1992
587614	Rossi	Luca	10/10/1991
937653	Bruni	Mario	01/12/1991

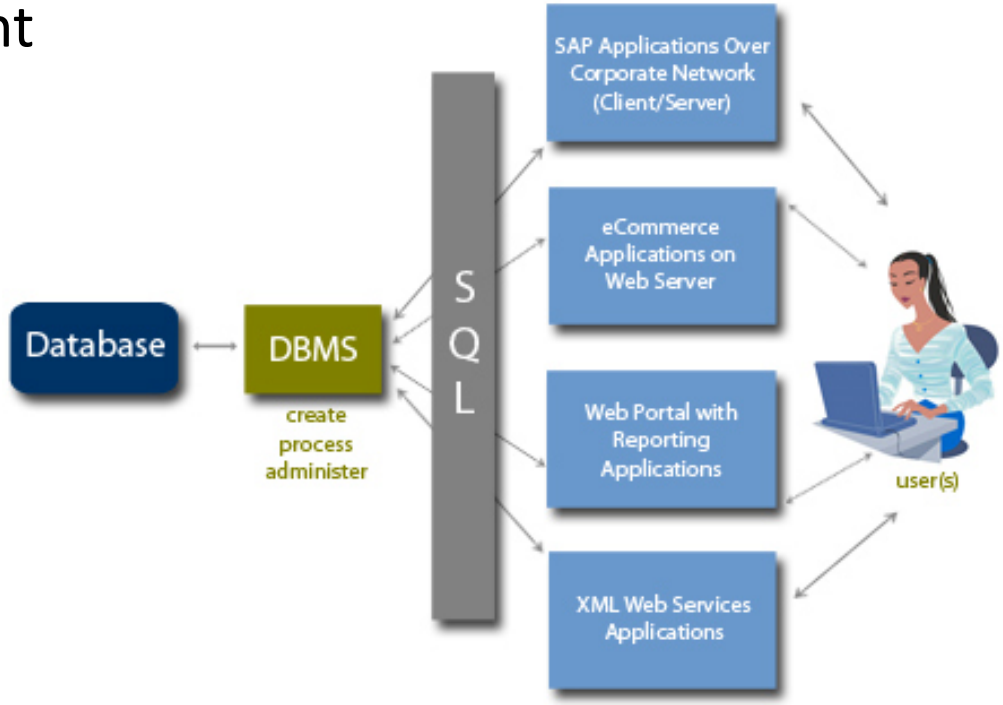
Database Management Systems

- A set of tools, to manage homogeneous sets of structured data

- Able to deal with
  - Huge amounts of data
  - Mission critical data
  - Shared data
  - Queries and updates

Architecture based on DBMS

- DBMS **manages** huge amount of data shared
  - **Efficiently**: optimization, parallelization
  - **effectively**
- And guaranteeing:
  - **Sharing:** concurrency control
  - **Reliability:** Failure resilience, data replication
  - **Security:** authorizations, access control



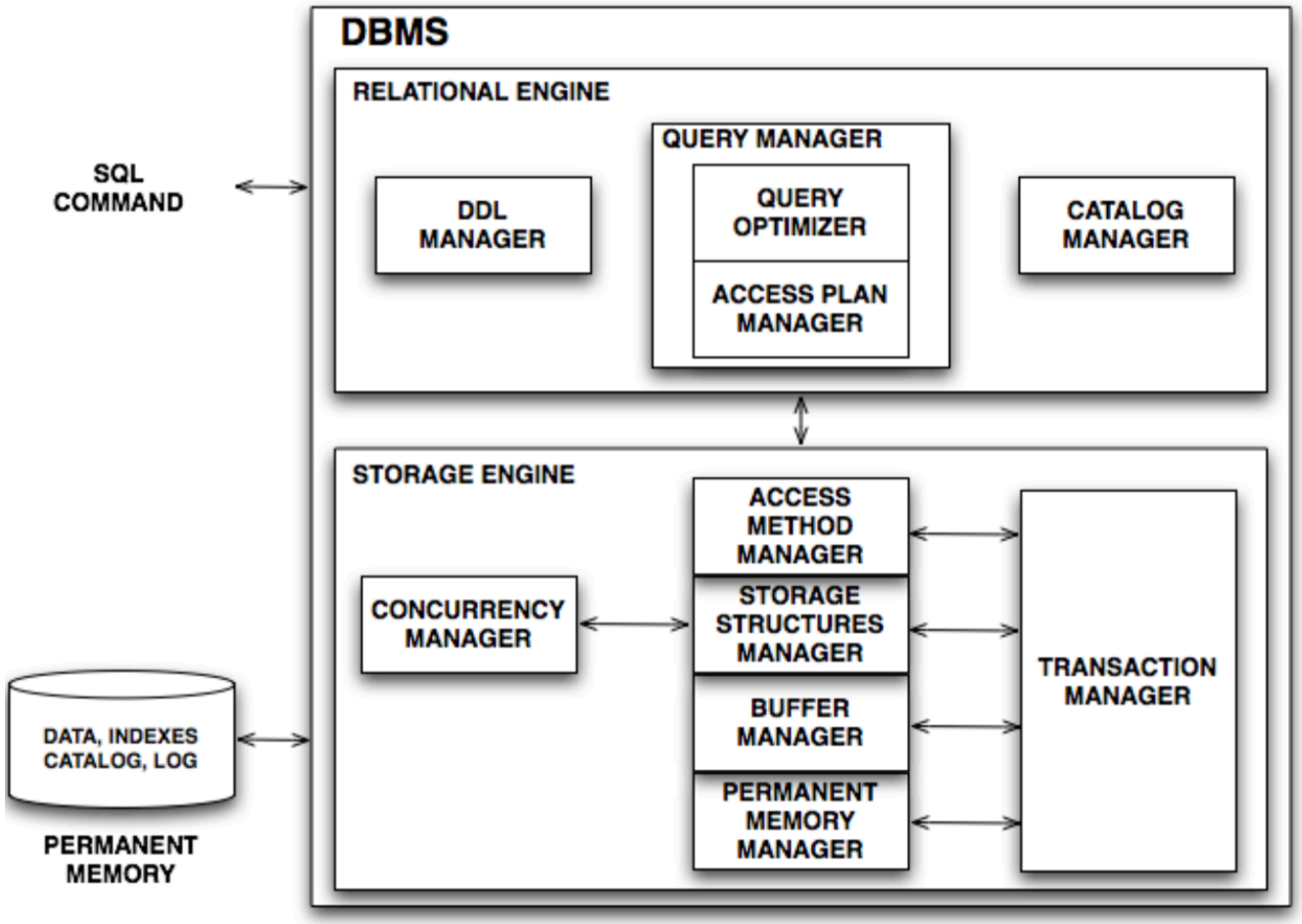
Some DBMS

- **Commercial**
  - IBM DB2, Oracle, **Microsoft SQL Server**, Sybase
  - Microsoft Access, FileMaker
- **Open Source**
  - MySQL (www.mysql.com)
  - PostgreSQL (www.postgresql.org)

Users of DBMSs

- **Analyst**
  - defines a schema
- **Programmer**
  - Writes applications
- **Data Base Administrator (DBA):**
  - Manages data structures
  - Manages user rights
- **Operator (final user):**
  - Uses applications
  - Uses query tools

DBMS Architecture



Transactions

- Transactional execution of a piece of code:
- **Atomicity** in presence of failures (all or nothing)
- **Durability**: transaction effects can be recovered after a failure
- **Serializability**: no interference by concurrent transactions