

# ***ICT & Business Intelligence & CRM*** **Introduction**

Andrea Vandin

Scuola Superiore Sant'Anna

Paolo Ferragina

Scuola Superiore Sant'Anna

Anna Monreale

Università di Pisa

# General Information

- **Andrea Vandin**
  - Email: [andrea.vandin@santannapisa.it](mailto:andrea.vandin@santannapisa.it)
- **Paolo Ferragina**
  - Email: [paolo.ferragina@santannapisa.it](mailto:paolo.ferragina@santannapisa.it)
- **Anna Monreale**
  - Email: [anna.monreale@unipi.it](mailto:anna.monreale@unipi.it)
- **Learning Material**
  - Slides – github wiki
  - Databases essentials – Antonio Albano
    - Pdf available on github wiki
  - A First Course in Database Systems, Jeffrey D. Ullman and Jennifer Widom
  - Decision Support Database – Antonio Albano
    - Pdf available on github wiki

# Exam

- Assignments to be handed-in during the course
  - Similar to exercises done in class
- Project Work: 3 people
- Final Mark:
  - 100% project

# **Overview of the course topics**

Andrea Vandin

Scuola Superiore Sant'Anna

Paolo Ferragina

Scuola Superiore Sant'Anna

Anna Monreale

Università di Pisa

# Topics

## **Part 1, Andrea Vandin**

- **Databases**
  - What is a DBMS
  - Designing a DB
  - Querying a DB

## **Part 2, Anna Monreale**

- **Business intelligence**
  - Decision Support Databases
  - Data warehouses
  - Designing a data warehouse
  - Querying a data warehouse

## **Part 3, Paolo Ferragina**

- **Graph Databases**
  - Introduction to Graphs: definition, properties and implementation
  - Graph Data Bases (GDB): modeling, principles, and structure.
  - An example of GDB query language: Cypher
  - Hands-on with a Graph DB: Neo4J

# Calendar of Part 1

## Intro to DB, Structure of DB, Create a DB

- 17/02/25 09:00-11:00 2 Vandin Introduction to course, to DB, to DBeaver.
- 19/02/25 09:00-11:00 2 Vandin Conceptual model. Exercises conceptual model.
- 24/02/25 09:00-11:00 2 Vandin Complete ex. conc. model. Logical model (ex.?)
- 26/02/25 09:00-11:00 2 Vandin Complete exercises logical model

## Query a DB

- 03/03/25 09:00-11:00 2 Vandin SQL1 with hands on examples
- 05/03/25 09:00-11:00 2 Vandin SQL2 with hands on examples
- 07/03/25 09:00-11:00 2 Vandin Exercises on SQL.

## Query a DB and Exercises on all topics

- 10/03/24 09:00-13:00 2 Vandin SQL3. Ex. on conceptual & logical model & SQL
- 10/03/24 14:30-16:40 2 Vandin Ex. on conceptual & logical model & SQL

Can we move it to Wed 12 or Thu 13 from 9 to 11?

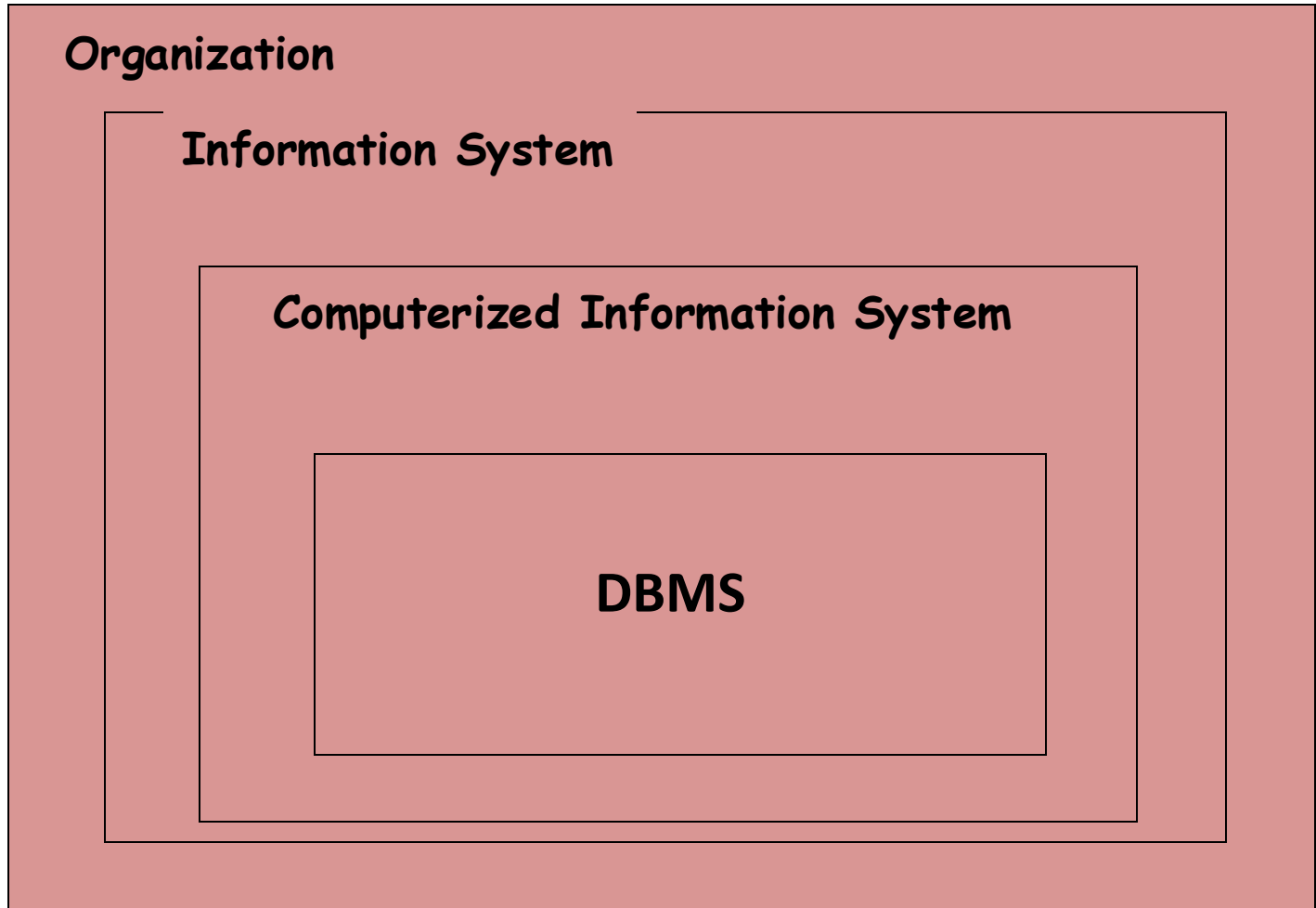
# Calendar of Part 1

Weekly PLAN (SALA CONFERENZE - Palazzo Boyl)

	MON 10	TUS 11	WED 12	THU 13	FRI 14
09:00-09:45	ICT1: Business Intelligence & CRM	Innovation in Cost & Perf. Mgmt.	Here?	Here?	Innovation in Cost & Perf. Mgmt.
10:00-10:45	ICT1: Business Intelligence & CRM	Innovation in Cost & Perf. Mgmt.			Innovation in Cost & Perf. Mgmt.
11:00-11:45	Open Innovation Modelling & R&D	Open Innovation Modelling & R&D	Open Innovation Modelling & R&D	Open Innovation Modelling & R&D	Open Innovation Modelling & R&D
12:00-12:45	Open Innovation Modelling & R&D	Open Innovation Modelling & R&D	Open Innovation Modelling & R&D	Open Innovation Modelling & R&D	Open Innovation Modelling & R&D
13:00-14:30					
14:30-15:15	ICT1: Business Intelligence & CRM	Innovation Lab	Innovation Lab	Innovation Lab	Innovation in Cost & Perf. Mgmt.
15:30-16:15	ICT1: Business Intelligence & CRM	Innovation Lab	Innovation Lab	Innovation Lab	Innovation in Cost & Perf. Mgmt.
16:30-17:15	Innovation Lab				
17:30-18:15	Innovation Lab				

Can we move it to Wed 12 or Thu 13 from 9 to 11?

# We start with some terminology





# Information as a Resource

- **Information** is one of the most important resources of any **organization**
- An intelligent management of the information can help the organization generating **new knowledge**
- It becomes more and more important to learn how to **represent, organize, manage** and **use** information

# Organization

- **Organizations**
  - es: companies, banks, public administrations, ....
  - An organic collection of
    - **resources** (people, materials, **information**),
    - **tools**
    - **procedures**
  - Has the goal to create/offer products/services
    - a bank provides financial services
    - a hospital supplies medical services

# Information System

- Component of an organization finalized to **information management** for supporting the organization's activities
  - Collects, stores, processes and communicates the information
- Any organization has an Information System

# Information System & Computerization

- **Information system** is independent of computerization
- Organizations as Libraries, Banks, etc ... existed and managed information long before computerization...

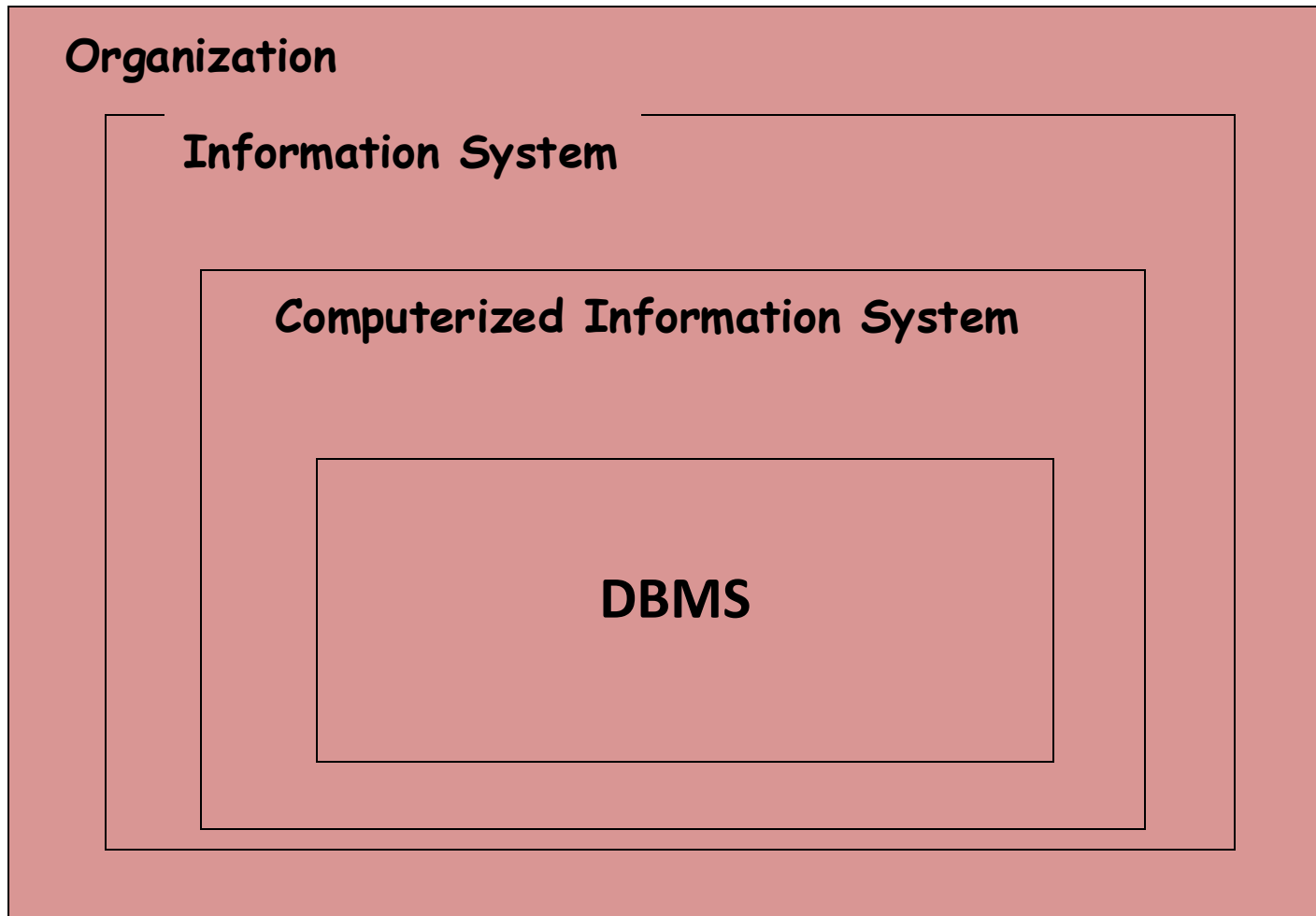


**Library of Alexandria**

# Computerized Information System

- The hardware and software 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**
- Implemented by **database technology**
  - *operational database* and
  - a collection of application programs used to
    - access and update the data **efficiently**

# Computerized Information System

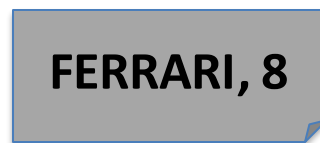


**Goal of DBMS:** processing data for getting information

# 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 instances*) and
  - 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



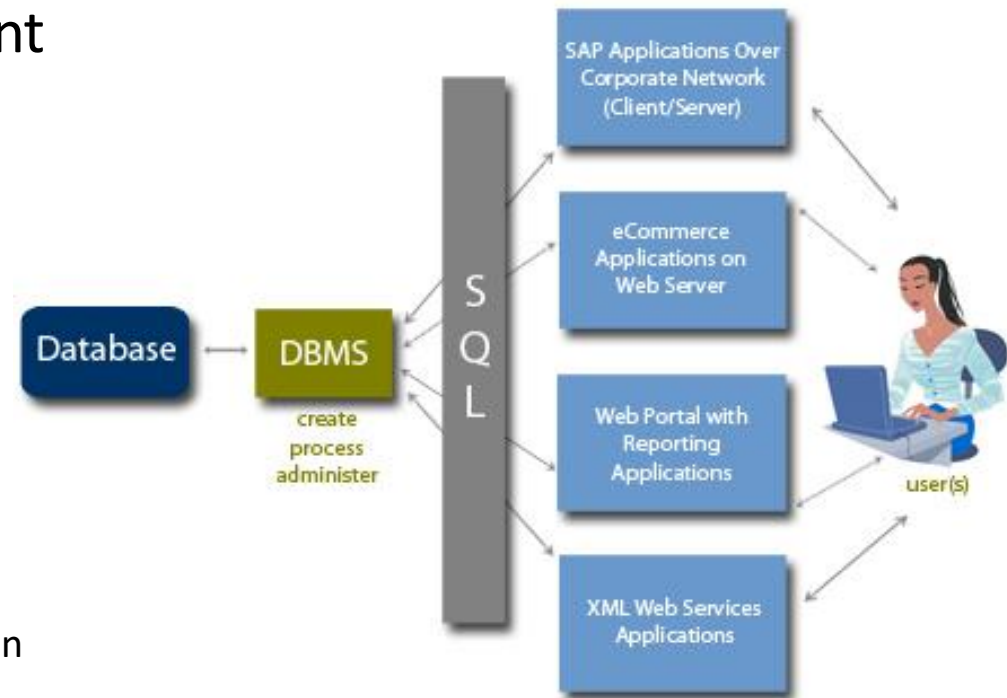
# **DBMS:**

## **Database Management Systems**

- A set of tools to manage homogeneous sets of structured data
  - Databases!
- 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



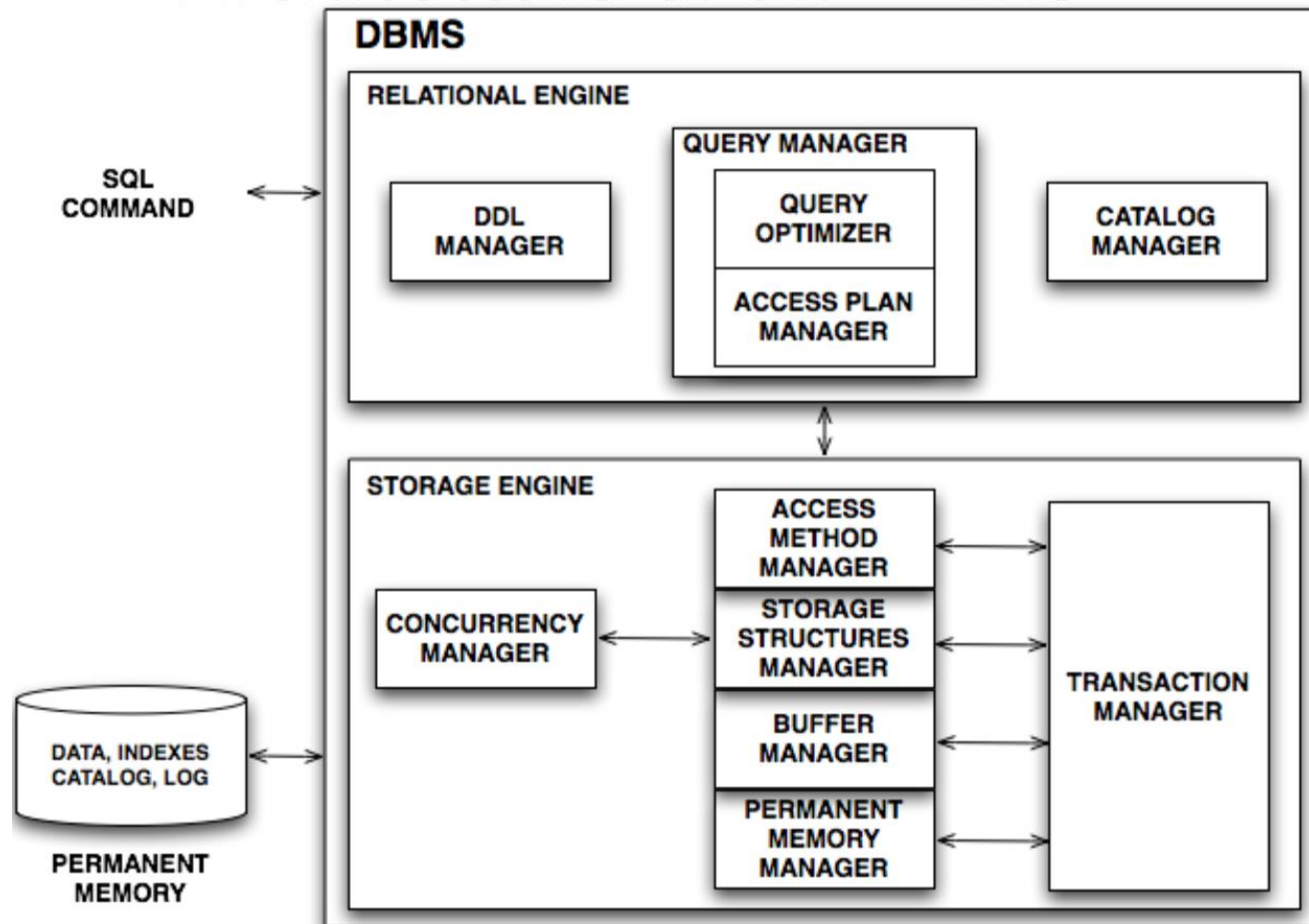
# Some DBMS

- **Commercial**
  - IBM DB2, Oracle, **Microsoft SQL Server**, Sybase
  - Microsoft Access, FileMaker
- **Open Source**
  - MySQL ([www.mysql.com](http://www.mysql.com))
  - PostgreSQL ([www.postgresql.org](http://www.postgresql.org))
  - **SQLite** ([www.sqlite.org/index.html](http://www.sqlite.org/index.html))

# Users of DBMSs

- **Analyst**
  - defines a schema
    - the *structure* of the DB. We will see more on this
- **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)
    - *An atomic transaction is not a 'very strong' transaction :D*
    - It is an action that involves several steps. If any of these steps fails, all the performed steps are reverted
  - **Durability**: transaction effects can be recovered after a failure
  - **Serializability**: no interference by concurrent transactions