

Database Systems

Introduction (Lecture-01)

Dr. Rudra Pratap Deb Nath

Associate Professor
Department of Computer Science & Engineering
University of Chittagong

rudra@cu.ac.bd

4th Semester 2025

Lectures

- Rudra Pratap Deb Nath

Teaching assistants

- TBD

- PhD (AAU, Denmark) PhD (UPC, Spain) M.Engg (TUT, Japan)
- Specialities: Database, Big Data, Exploratory Business Intelligence, Data Science, Semantic Web, Artificial Intelligence, and Machine learning.
- Vendor Certification: Oracle 10g OCP (DBA)
- Ambassador
 - EU Master Program- Big Data Management and Analytics (BDMA) :
<https://bdma.ulb.ac.be/bdma/>
 - EU Doctorate Program- Data Engineering for Data Science (DEDS) :
<https://deds.ulb.ac.be/>

Tell me about yourselves...

Organization

- Google classroom: [https://classroom.google.com/c/
Nzg0MzM3ODAwNzY3?cjc=zgmmvhis](https://classroom.google.com/c/Nzg0MzM3ODAwNzY3?cjc=zgmmvhis)
- Facebook group:
<https://www.facebook.com/groups/1133121245287047>
- All course materials will be on the group
(slides, exercise sheets, self study, quizzes,...)

Agreements (From my sides)

- Take your own responsibility
- Be part of a group, composed of 2 persons
- Kindly do not copy-paste. I am very strict in this regard.
The score will be minus if plagiarism is proved
- Cooperation instead of competition. Appreciate and acknowledge others
- Maintain the professionalism

Agreements (From your sides)



Book

- By Silberschatz, Korth, Sudarshan
- 7th edition
- The PDF version is in the FB group and Google classroom
- Do not use book as a data, try to earn knowledge

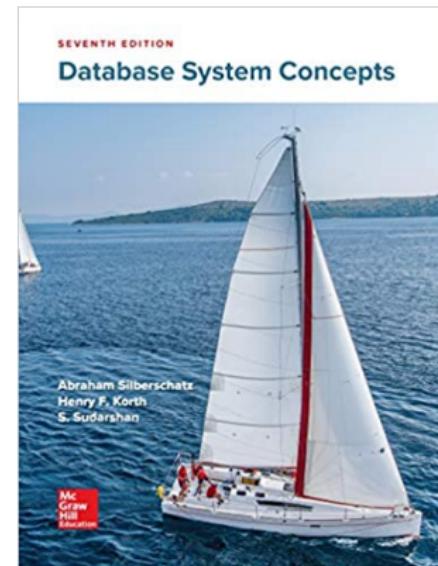


Figure: A reference book

Week plan

- Lectures+ Lab sessions
 - Tuesdays: 9:50 to 13:05
 - Wednesdays: 9:50 to 10:35
 - Thursdays: 9:50 to 12:15
- Assignment
 - once/twice per week
- Self study

Course Setup

Lecture

- We will use slides + white board in lectures
 - More slides than crammed slides
 - Slides override the book
- Make the class interactive by responding, asking
- Questions may be correct or wrong, but you will learn a lot
- Ask questions right away during the lecture
- Actively participate in assignments. Obtained marked will be considered as continuous internal evaluation (CIE) mark.
- For better understanding, read chapters in the book or search on the Web after/before the lecture

Course Setup

Assignment

- Before Lectures ensure that you understand the material covered in the previous week and the applications of the material
- Assignment with submission deadline will be available in the Google classroom
- No extension for the submission deadline

Course Setup

Tutorial & Final Exam

- Expect analytical questions instead of memorization
- 2-4 tutorials based on situations
- See syllabus for the exam question structure

Learning Goal

This is a course that will

- teach you the fundamentals that are common to all DBMSs
- help you make better use of relational database systems.

This is not a course to

- make you a DBA within 4-6 months
- make you expert on any DBMSs.

Learning goal

- Motivation and introduction of Database
- Understand the relational model and apply relational algebra
- Conceptual design of a database (ER model)
- Logical design (create and evaluate a database schema that conforms to normal forms)
- Make use of SQL to create, manipulate, and query relational databases
- Understand how the data is physically organized
- Understand the transactional concept and key issues of concurrency control and recovery
- A bit of hot research topics in databases.

Opportunities

■ Industries

- Data scientist, Data Engineer, Business Intelligence (BI) specialist, Database Administrator, Database developer, ontology specialist, ETL engineer, Data Warehouse specialist, etc.

Opportunities

■ Industries

- Data scientist, Data Engineer, Business Intelligence (BI) specialist, Database Administrator, Database developer, ontology specialist, ETL engineer, Data Warehouse specialist, etc.

■ Academia

- Different academic positions: Research assistant, PhD, PostDoc, (assistant/associate) professor
- 20+ CSECU students were recommended by me in Europe, Japan, Canada, Australia and USA for higher studies
- Last week Ataur Rahman (Shaon) has gotten the acceptance for PhD (DEDS)
- Big fishes: Check Google scholar for top venues-
https://scholar.google.com/citations?view_op=top_venues&hl=en&vq=eng_databasesinformationsystems

Current top publications

The screenshot shows a Google Scholar search results page for the category 'Databases & Information Systems'. The results are ordered by h5-index. The top publication is 'International World Wide Web Conferences (WWW)'. Other top publications include IEEE Transactions on Knowledge and Data Engineering, ACM SIGIR Conference on Research and Development in Information Retrieval, International Conference on Very Large Databases, Information Processing & Management, ACM International Conference on Web Search and Data Mining, ACM International Conference on Information and Knowledge Management, ACM SIGMOD International Conference on Management of Data, Journal of Big Data, International Conference on Data Engineering, International Conference on Web and Social Media (ICWSM), IEEE International Conference on Big Data, Knowledge and Information Systems, ACM Conference on Recommender Systems, IEEE Transactions on Big Data, Information Systems, Semantic Web, Workshop of Cross-Language Evaluation Forum, International Semantic Web Conference, and ACM Transactions on Intelligent Systems and Technology (TIST).

Publication	h5-index	h5-median
1. International World Wide Web Conferences (WWW)	98	163
2. IEEE Transactions on Knowledge and Data Engineering	98	147
3. ACM SIGIR Conference on Research and Development in Information Retrieval	75	115
4. International Conference on Very Large Databases	73	102
5. Information Processing & Management	70	100
6. ACM International Conference on Web Search and Data Mining	69	133
7. ACM International Conference on Information and Knowledge Management	69	114
8. ACM SIGMOD International Conference on Management of Data	68	93
9. Journal of Big Data	55	104
10. International Conference on Data Engineering	55	76
11. International Conference on Web and Social Media (ICWSM)	54	83
12. IEEE International Conference on Big Data	52	93
13. Knowledge and Information Systems	51	76
14. ACM Conference on Recommender Systems	47	111
15. IEEE Transactions on Big Data	45	66
16. Information Systems	44	64
17. Semantic Web	43	66
18. Workshop of Cross-Language Evaluation Forum	43	63
19. International Semantic Web Conference	39	59
20. ACM Transactions on Intelligent Systems and Technology (TIST)	38	72

Dates and citation counts are estimated and are determined automatically by a computer program.

Overview

1 Organization

- Teachers
- About Myself
- Administrative Information
- Course Setup

2 Introduction to DBMS

- Data is Ubiquitous
- Problem with Data
- Architecture
- Available Systems

Computer programs

```
1 public class Student {  
2  
3     int id;  
4     String name;  
5     String birthdate;  
6     String deptName;  
7     double ects;  
8  
9     public Student(int id, String name, String birthdate,  
10                    String deptName, double ects) {  
11         this.id= id;  
12         this.name=name;  
13         this.birthdate=birthdate;  
14         this.deptName=deptName;  
15         this.ects=ects;  
16     }  
17 }
```

Computer programs

```
public static void main(String[] args) {  
    Student st1 = new Student (2101, "Bob", "  
        01-01-2000", "CSE", 3.67);  
    Student st2 = new Student (2101, "Peter", "  
        41-05-2000", "Computer Science", -0.5);  
    Student st3 = new Student (2101, "Peter", "  
        41-05-2016", "Computer Science", 3.80);  
    List <Student> allStudents = new LinkedList <  
        Student>();  
    allStudents.add(st1);  
    allStudents.add(st2);  
    allStudents.add(st3);
```

Computer programs

```
try {
    FileWriter objectOut = new FileWriter("C:/Lecture/
        Database_CU/allStudents.txt");
    objectOut.write("student Id" + " " + "Student
        Name" + " " + "student Birthdate" + " " +
        "student DeptName" + " " + "student Ects" + "\n
        ");
    for (Student student : allStudents) {
        objectOut.write(student.id + " " + student.name +
            " " + student.birthdate + " " + student.
            deptName + " " + student.ects + "\n " );
    }
    objectOut.close();
} catch (IOException ex) {
    ex.printStackTrace();
    System.exit(0); }}
```

Files

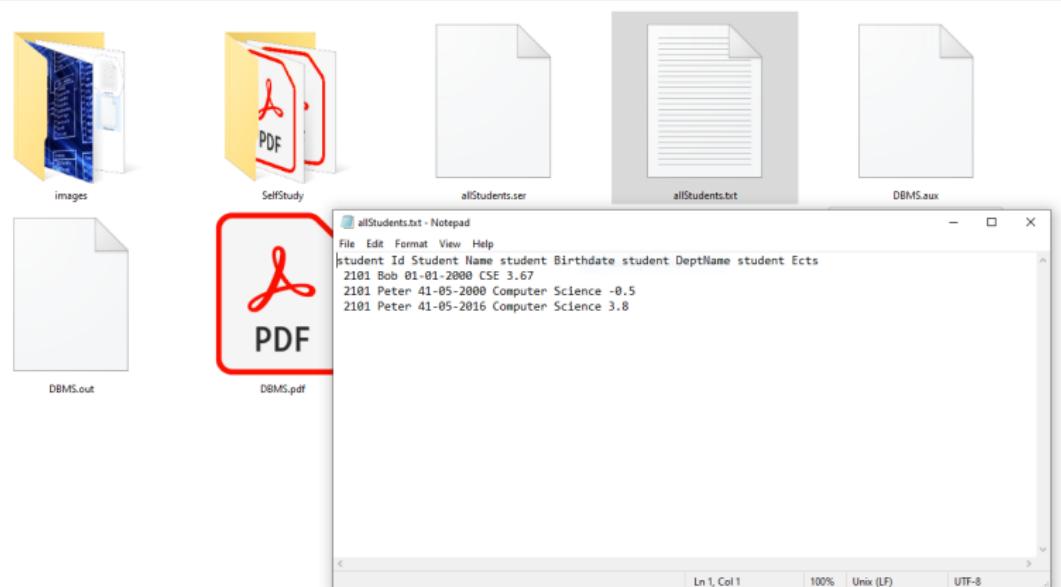


Figure: Data in a file

Mobile phones



Figure: Data in mobile phones

Books

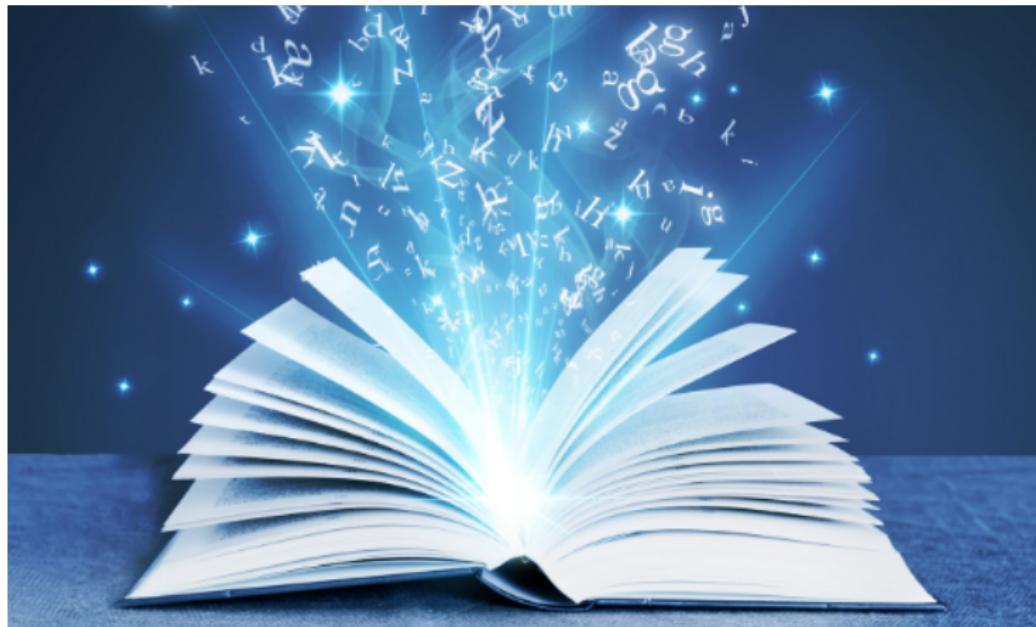


Figure: **Data in books** [<https://www.amnesty.org.uk/>]

Cameras



Figure: Data in images/videos [<https://www.nikonusa.com/>]

Transportation



Figure: Data in transportation [<https://blog.seagate.com/human-only-data-at-the-edge-will-make-driverless-cars-safe/>]

Satellite

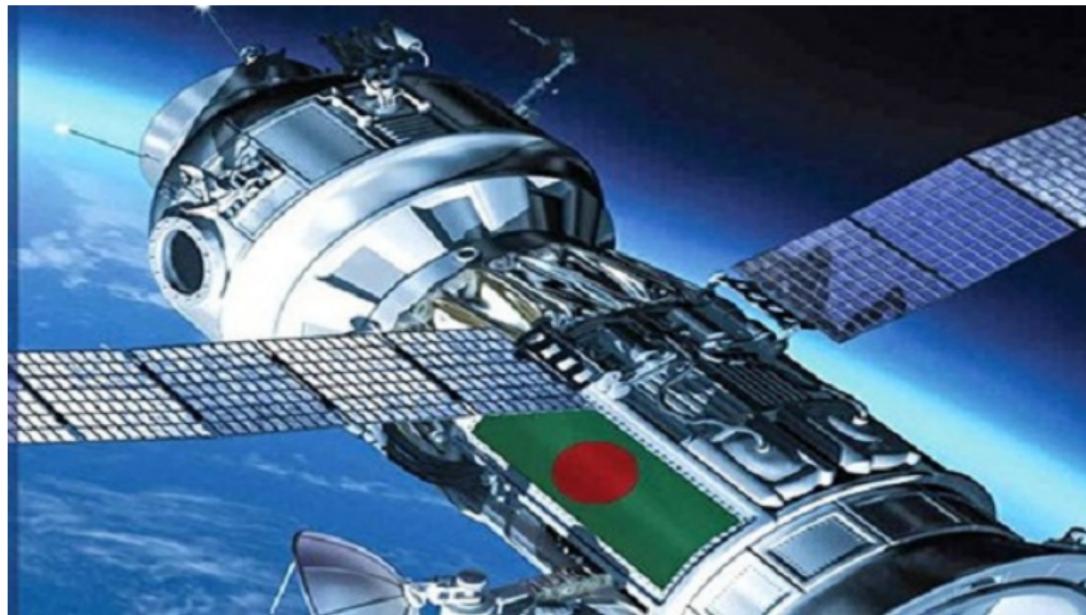


Figure: Data in Satellite [<https://eng.campuslive24.com/tech/article/51/Bangabandhu-satellite-connects>]

Medical devices



Figure: **Data in Medical devices**

Sensors

MEMS sensors & actuators: the 5 senses and many more

(Source: Status of the MEMS Industry report, Yole Développement, 2019)

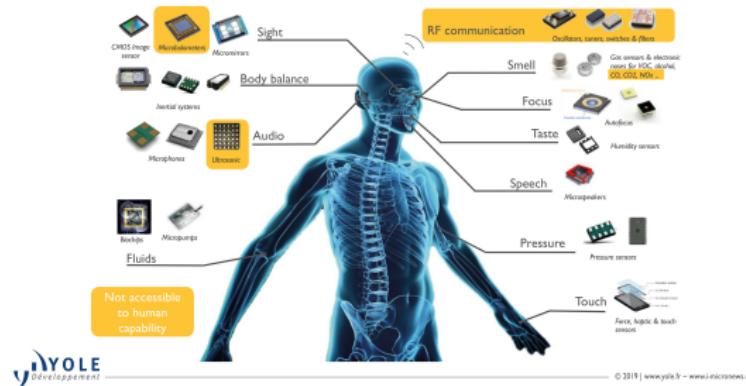


Figure: Data in sensors

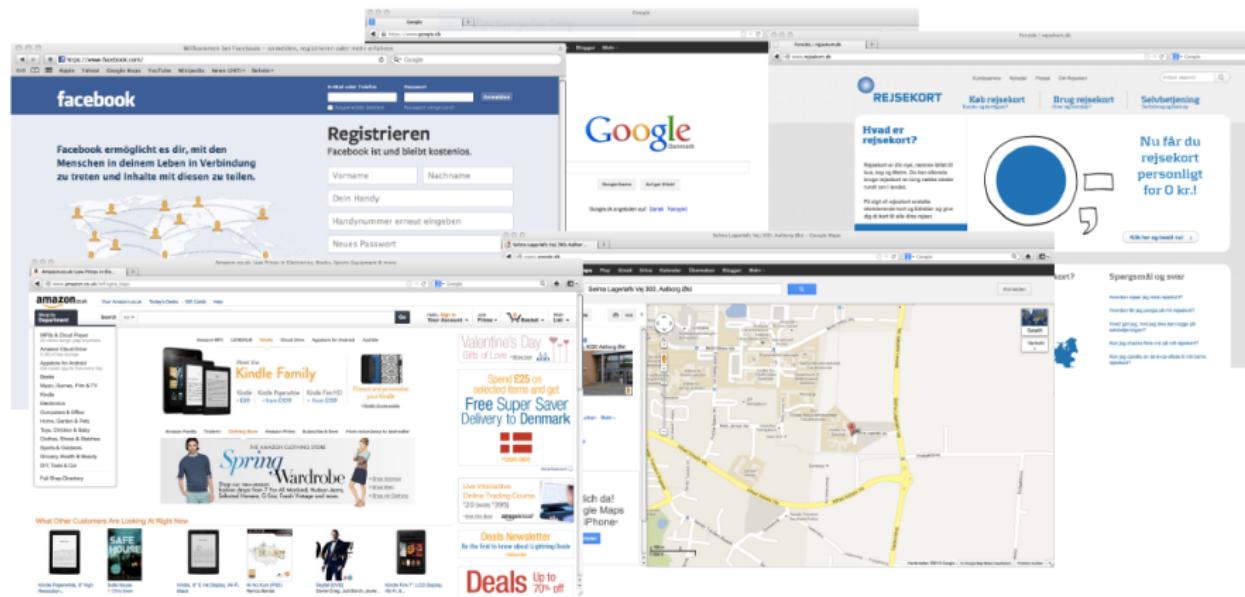
Angels record our data

Data related to our activities are recorded by Angels
(Chittrogupta/Keramon Kate Bin)



Figure: Angels record our data

in many other applications



...in many other applications....

- Banking: all transactions
- Universities: registration, grades
- Sales: customers, purchases, products
- Human resources: employee records, salaries, tax deductions
- Transportations: reservations, schedules
- Online retailers: order tracking, customized recommendations

...in many other applications....

- Banking: all transactions
- Universities: registration, grades
- Sales: customers, purchases, products
- Human resources: employee records, salaries, tax deductions
- Transportations: reservations, schedules
- Online retailers: order tracking, customized recommendations

Databases touch all aspects of our lives

...in many other applications....

- Banking: all transactions
- Universities: registration, grades
- Sales: customers, purchases, products
- Human resources: employee records, salaries, tax deductions
- Transportations: reservations, schedules
- Online retailers: order tracking, customized recommendations

Databases touch all aspects of our lives

Imagine a life without a database...

Problems with data

```
6 Student st1 = new Student  
7 (2101, "Bob", "01-01-2000", "CSE", 3.67);  
8 Student st2 = new Student  
9 (2101, "Peter", "41-05-2000", "Computer Science", -0.5);  
0  
1 id, name, birthdate, department, GPA
```

Problems with data

```
2 Student st1 = new Student
3 (2101, "Bob", "01-01-2000", "CSE", 3.67);
4 Student st2 = new Student
5 (2101, "Peter", "41-05-2000", "Computer Science", -0.5);
6
7 id, name, birthdate, department, GPA
```

Is there any potential problem?



Problems with data

```
3 Student st1 = new Student  
4 (2101, "Bob", "01-01-2000", "CSE", 3.67);  
5 Student st2 = new Student  
6 (2101, "Peter", "41-05-2000", "Computer Science", -0.5);  
7 id, name, birthdate, department, GPA
```

Problems with data

```
3 Student st1 = new Student
4 (2101, "Bob", "01-01-2000", "CSE", 3.67);
5 Student st2 = new Student
6 (2101, "Peter", "41-05-2000", "Computer Science", -0.5);
7 id, name, birthdate, department, GPA
```

Potential Problems

- Both students have same id
- 41-05-2000 is not a valid date
- Is -0.5 is valid GPA?
- is CSE and Computer Science the same?

Problems with data

```
3 Student st1 = new Student  
9 (2101, "Bob", "01-01-2000", "CSE", 3.67);  
0 Student st2 = new Student  
1 (2101, "Peter", "41-05-2000", "Computer Science", -0.5);  
2 id, name, birthdate, department, GPA
```

Blackout



- Are both students written to disk?
- ...or only 1.2
- Possible permanent hard disk failure
- Retrieval of most recent version of data

Problems with data

```
3 Student st1 = new Student  
4 (2101, "Bob", "01-01-2000", "CSE", 3.67);  
5 Student st2 = new Student  
6 (2101, "Peter", "41-05-2000", "Computer Science", -0.5);  
7 id, name, birthdate, department, GPA
```

Potential Problems

- Not enough main memory for all data
- Efficient query processing
- Robust query language
- Access rights
- Adding metadata
-

Towards a DBMS



Analogy

- Data categorization (study, relation, food)
- Data storage
- Data transaction
- concurrent thinking (concurrency control)
- Retrieval of forgotten data

Figure: Human Analogy

Towards a DBMS

In the beginning..

1. Excel sheet, one user

Towards a DBMS

In the beginning..

1. Excel sheet, one user

Next....

more and more Excel sheets, more and more users

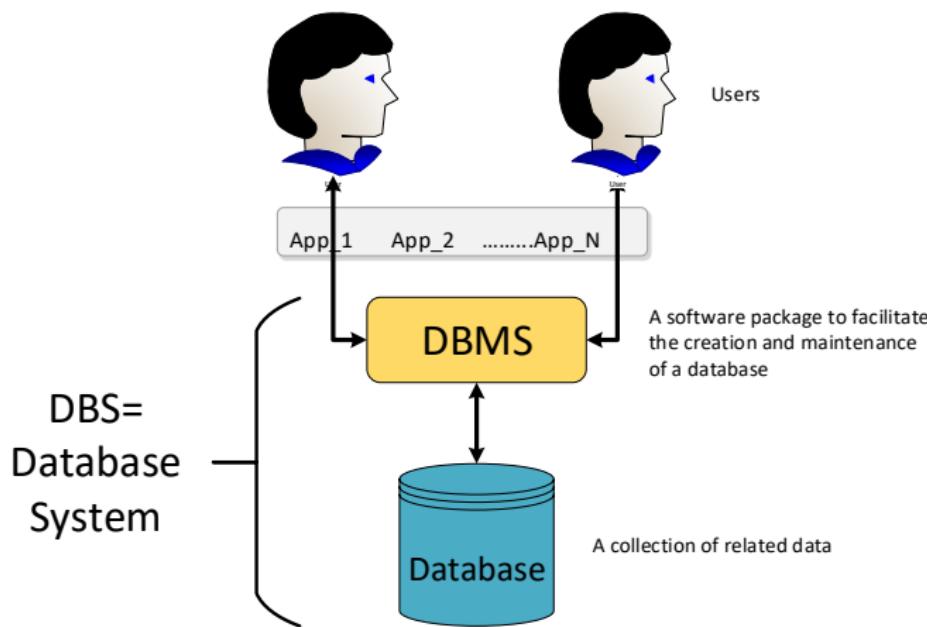
Problems, problems, .. and problems

- Sheet with million of entries
- Multiple users accessing the same sheet
- 1000+ users and 1000+ sheets

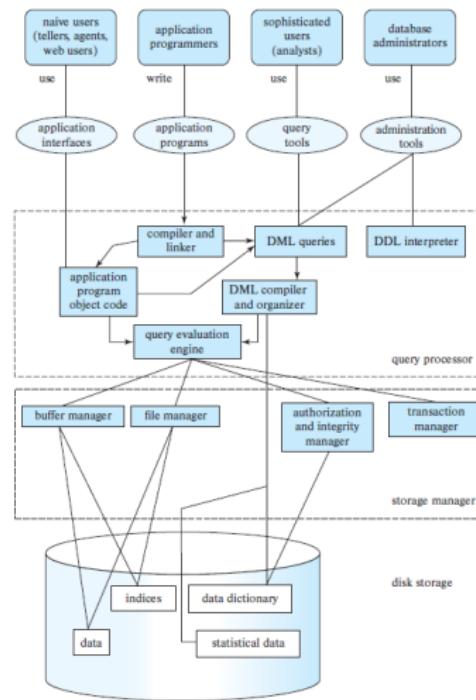
Database systems

A database system provides solution to all these problems

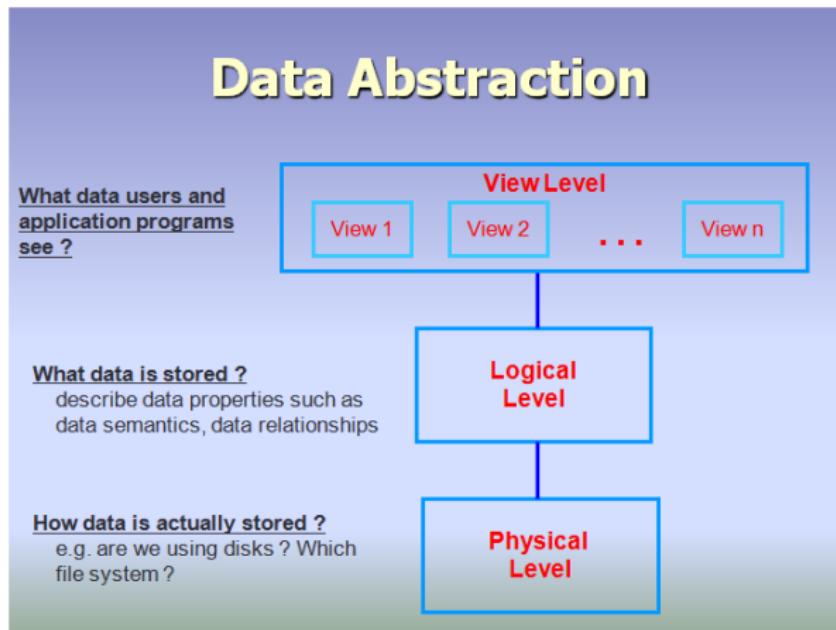
Database systems



Database systems



3 levels of data abstraction



Physical data independence

Changes regarding file structure and access paths (physical layer) have no influence on the conceptual schema (logical layer).

Kind of changes on the physical level

- New hard disk added
- New process added
- Files are split into multiple files

Logical data independence

Changes on the logical layer have no influence on external schemas and applications.

Kind of changes on the logical level

- New attribute or new table added on the conceptual schema
- Changing the name of an attribute or table on the conceptual schema

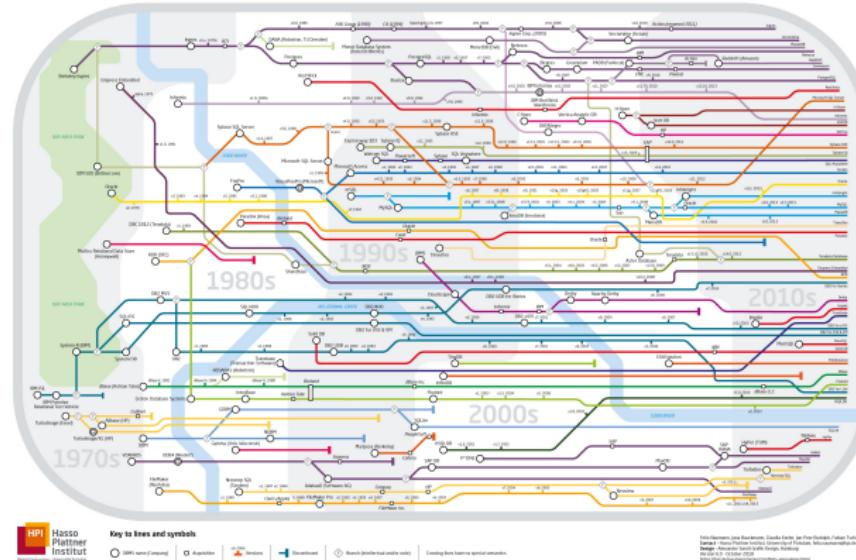
Definitions

- Mini-world
 - Some part of the real world about which information is stored
- Data/Information
 - Known facts about the mini-world that can be recorded and have an implicit meaning
- Database (DB)
 - A collection of related data
- Database Management Systems (DBMS)
 - A software package to facilitate the creation and maintenance of a database
- Database Systems (DBS)
 - A database and a DBMS
- Database instance
 - The content of a DB at a particular time

Genealogy of Relational Database Management Systems

<https://hpi.de/naumann/projects/rdbms-genealogy.html>

Genealogy of Relational Database Management Systems



Popular DBMSs

- Big players
 - Oracle
 - IBM DB2
 - Microsoft SQL server
- Smaller players
 - PostgreSQL
 - MySQL

The End