Introduction to databases

References

Book:
1. Elmasri, Ramez, and Shamkant Navathe. *Fundamentals of database systems*. Addison-Wesley Publishing Company, 2010.
2. Garcia-Molina, Hector. *Database systems: the complete book*. Pearson Education India, 2008.
3. Ramakrishnan, Raghu, and Johannes Gehrke. *Database management systems*. McGraw Hill, 2000.

Journal:
1. https://dev.mysql.com/doc/refman/8.0/en/tutorial.html
2. https://www.w3schools.com/sql/

Learning Maps

Sequence	Title
1	Introduction to databases
2	Relational Databases
3	Relational Algebra
4	Structured Query Language – Part 1
5	Structured Query Language – Part 2
6	Constraints and Triggers
7	Entity Relationship Model
8	Functional Dependency
9	Normalization
10	Storage - Indexing
11	Query Processing
12	Transaction Management – Part 1
13	Transaction Management – Part 2

Intro > Overview

☐ A : Voice and PPT Overview
☐ B : Text-based Overview
■ C : Video and PPT Overview

Opening Message	→ In this lesson, we will study the course overview, some concepts on database and we will identify some key characteristics of database and file system approach in data management
Lesson topic	1. Course overview 2. Basic concepts on database 3. Data management: database vs. file approach
Learning Goals	Upon completion of this lesson, students will be able to: 1. Recall the concepts of database, DBMS, data model, file system. 2. Identify the characteristics of database and file system approach in data management

Intro > Keywords

Keyword	Description
Database	A shared collection of related data designed to meet the information needs of an organization
DBMS	A software to facilitate the creation and maintenance of a database
Data model	A set of concepts used to describe the structure of a database: data types, relationships, constraints, semantics, …
Schema	A data structure fulfilled all features of the parts of the real world which is of interest to the users
Instance	The data itself (in the context of data model and database schema)

Lesson > Topic 1: Course overview

- Course objective
- Motivation for studying databases

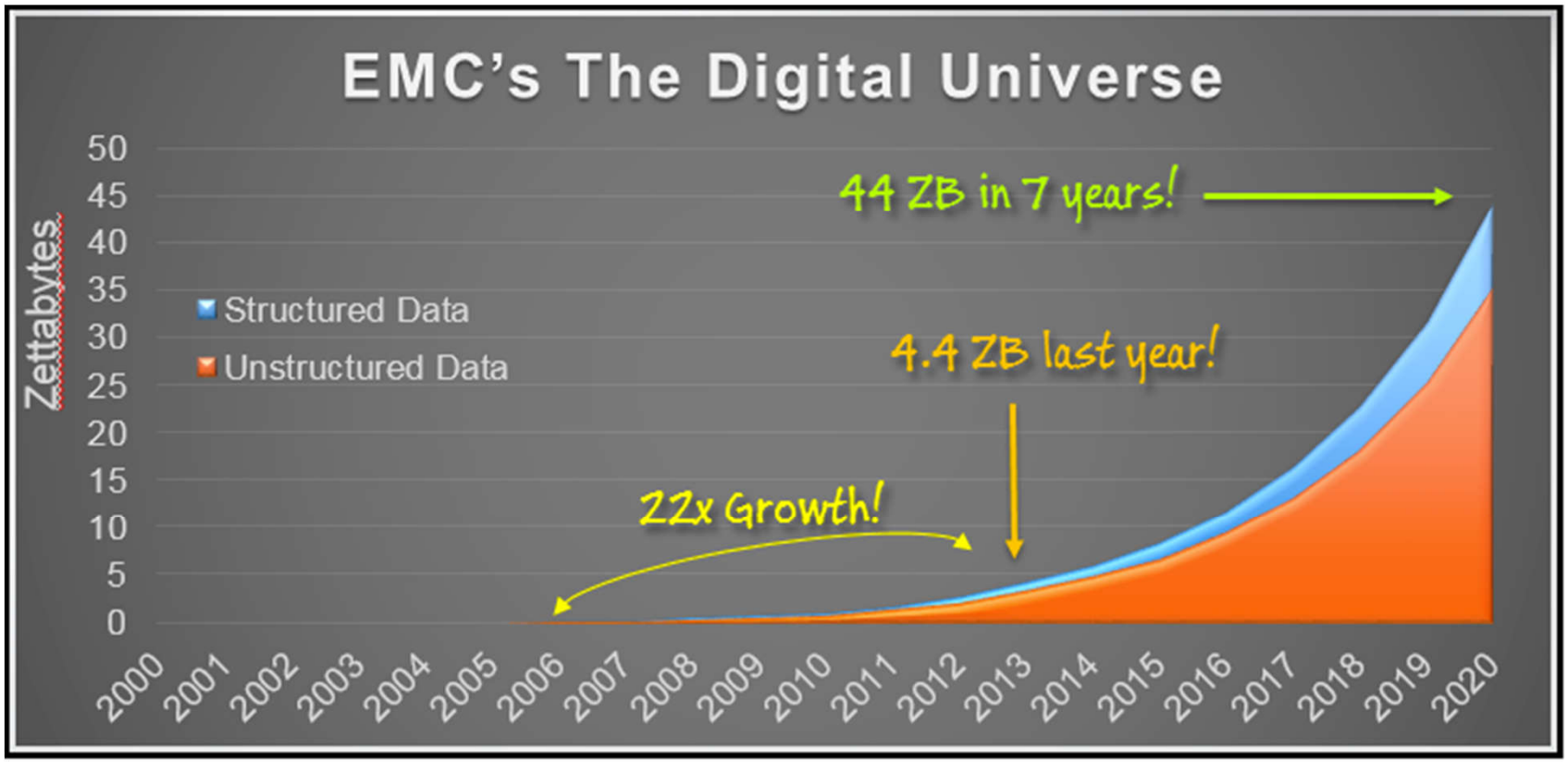
1. Course objective

- Understanding
  - purposes and roles of databases in Information systems
  - the basic structure of DBMS, such as index management and query processing
  - basic methods for transaction management, such as concurrency control and fault recovery
- Experiencing
  - with storing and querying data in RDBMS
  - database design such as normalization and integrity constraints

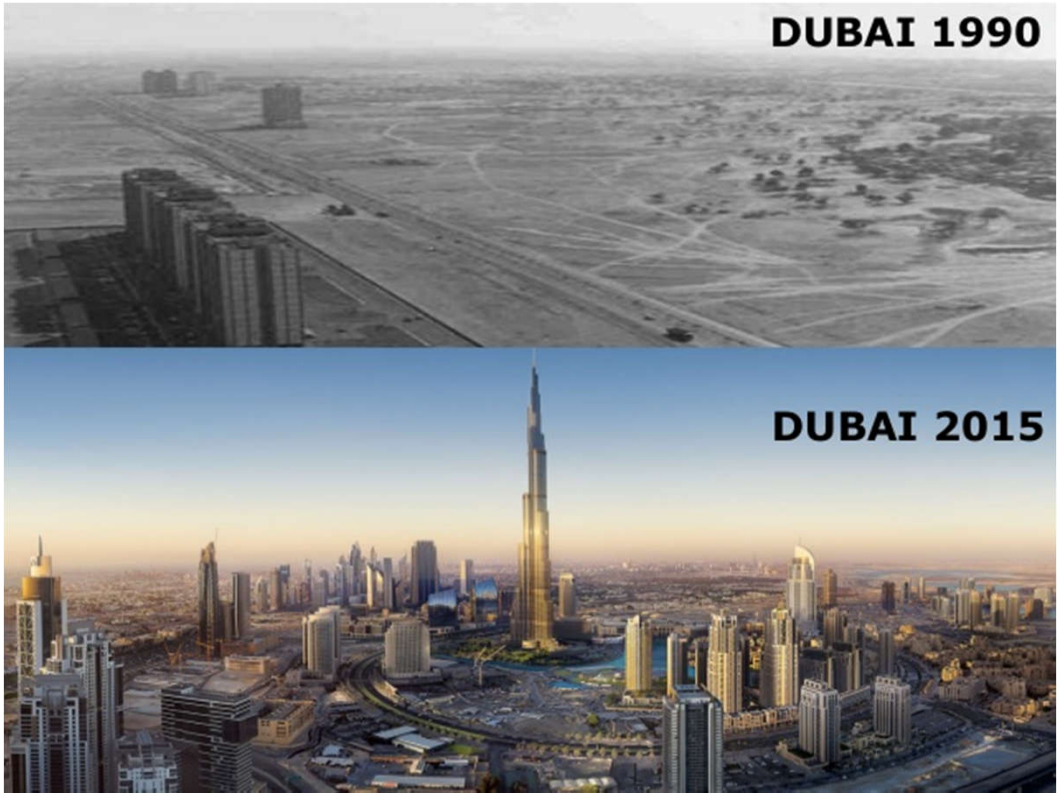
2. Motivation for studying databases

- Major research field with long history (since the begining of computer)
- 90% applications use databases
- Hot jobs in startups, big coorporates
- Massive Industry: Oracle, IBM, Microsoft, Google, AWS

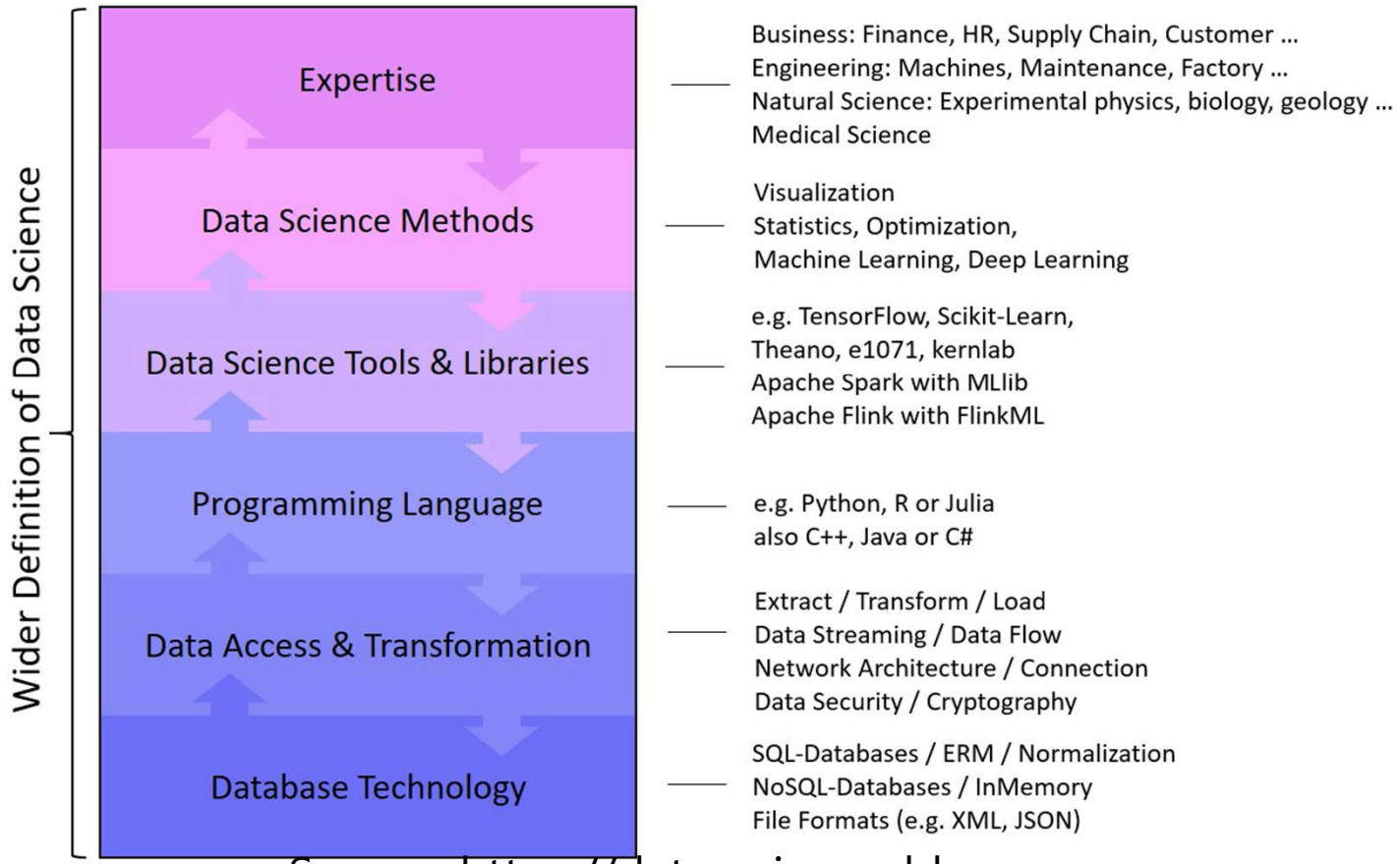
2.1. How big is our digital universe?



2.2 Data is the new oil of the 21st century



Data Science Knowledge Stack



Source: https://data-science-blog.com

Lesson > Topic 2: Basic concepts

- Data
- Database
- Data model vs. schema vs. instance
- Database management system (DBMS)
- Database environment
- Database users

2.1. Data

- Data has many definitions
  - Data is any sequence of one or more symbols given meaning by specific act(s) of interpretation. (wikipedia)
  - Information in raw or unorganized form (such as alphabets, numbers, or symbols) that refer to, or represent, conditions, ideas, or objects. Data is limitless and present everywhere in the universe. (www.businessdictionary.com)
  - E.g. a specific student data: ID, Name, Age, Gender, Address

2.2. Database

A shared collection of related data designed to meet the information needs of an organization

- Logically coherent
- Internally consistent
- Specific purpose
- Representation of the real world
  - Entities (e.g., Students, Courses)
  - Relationships (e.g., Tam is enrolled in C++)

2.2. Database (cont'd)

- Example: a course management system
  - Entities
    - Students
    - Courses
    - Teachers
  - Relationships
    - Students take in some courses
    - Course are given by some teachers

2.3. Model vs. Schema vs. Instance

- Data Model
  - set of concepts used to describe the structure of a database: data types, relationships, constraints, semantics, …
  - tool for data abstraction
- Schema
  - data structure fulfilled all features of the parts of the real world which is of interest to the users
- Instance
  - Data itself

2.4. Example

- Data Model

```
type <type_name> =
    record
        <field_name> : <data_type>;
        <field_name> : <data_type>;
        …
end;
```

- Schema

```
type student = record
        ID : string;
        fullName: string;
        Birthday: date;
        Address: string ;
        Class: string;
end;
```

- Instance

(« Stud001 », « Nguyen », 1/4/1983, «1 Dai Co Viet », « 1F VN K50 »)

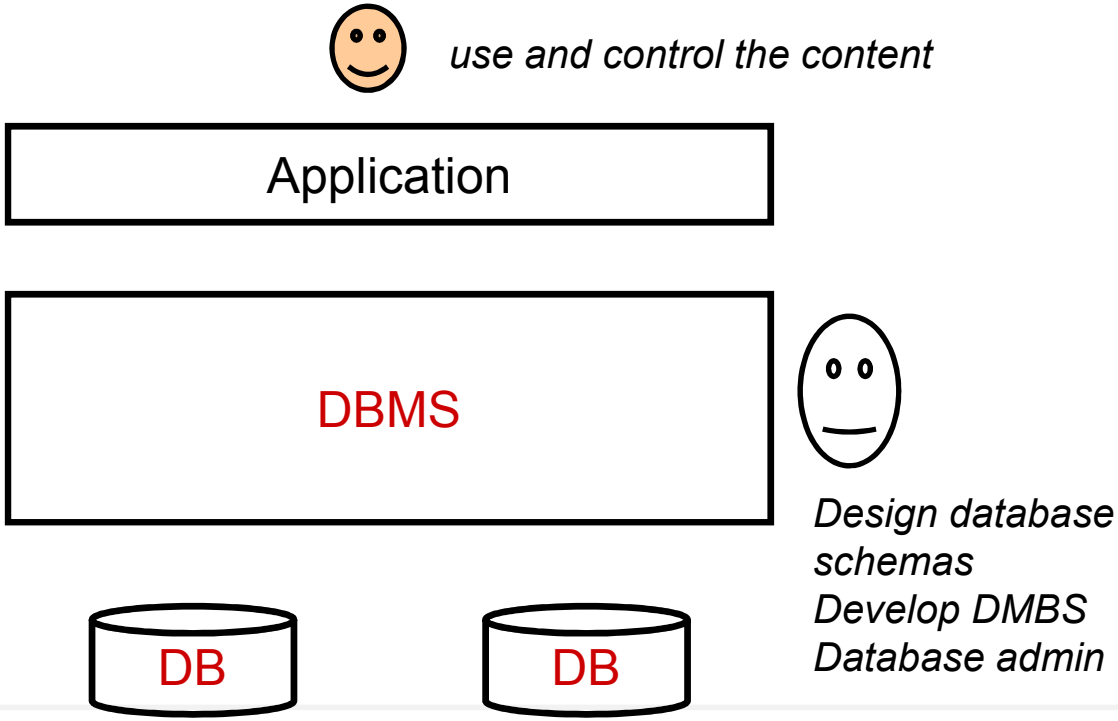
2.5. Database Management System (DBMS)

A software to facilitate the creation and maintenance of a database

- Defining ~ specifying types of data
- Constructing ~ storing & populating
- Manipulating ~ querying, updating, reporting

2.6. Database Environment

- A database environment is a collective system of components that regulates the management, the use of data, and the data itself
- Hardware
- Software
- Data
- Users
- Procedures/Manuals



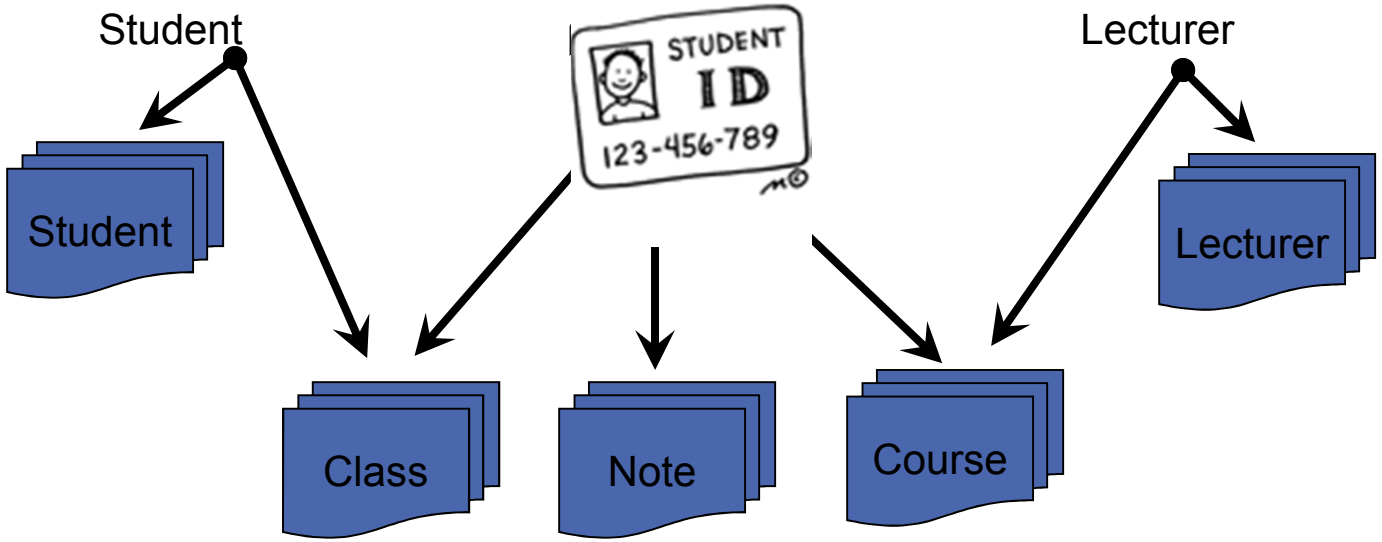
2.7. Database Users

- **Database administrators**
  - authorize access to the database
  - co-ordinate and monitoring its use
  - acquire software, and hardware resources, controlling its use and monitoring efficiency of operations.
- **Database Designers**
  - define the content, the structure, the constraints, and functions or transactions against the database.
  - communicate with the end-users and understand their needs.
- **End-users**
  - use the data for queries, reports and some of them actually update the database content.
  - Casual end users
  - Naive users
  - Sophisticated end users

Lesson > Topic 3: Data management

- File management system approach
- Database management system approach

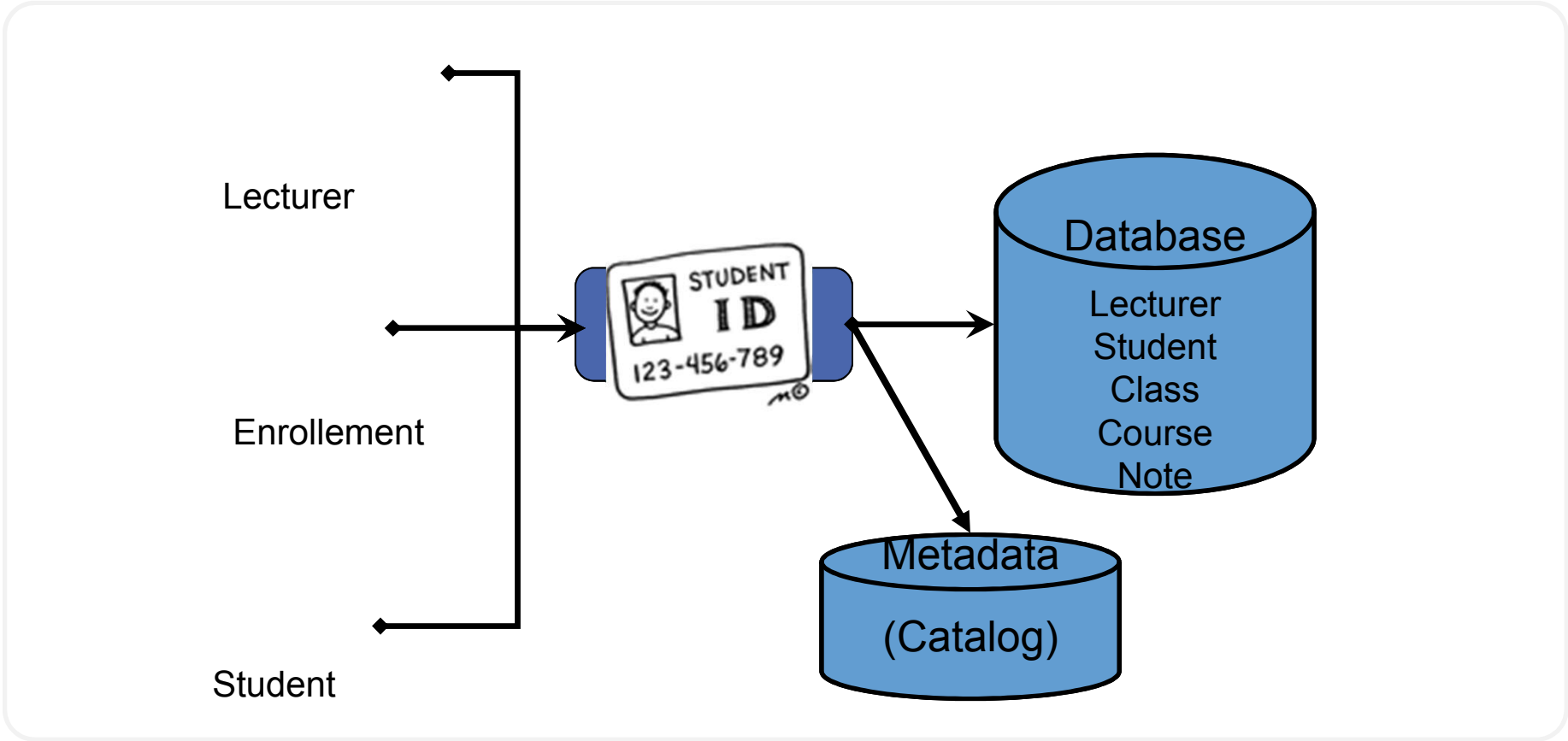
3.1. File System approach



3.1.1. Limitations of file-based approach

- Uncontrolled redundancy
- Inconsistent data
- Inflexibility
- Limited data sharing
- Poor enforcement of standards
- Low programmer productivity
- Excessive program maintenance
- Excessive data maintenance

3.2. Database approach



3.2.1. Advantages of database approach

- Controlled redundancy
  - consistency of data & integrity constraints
- Integration of data
  - self-contained & represents semantics of application
- Data and operation sharing
  - multiple interfaces
- Flexibility
  - data independence
  - data accessibility
  - reduced program maintenance
- Services & Controls
  - security & privacy controls
  - backup & recovery
  - enforcement of standards
- Ease of application development

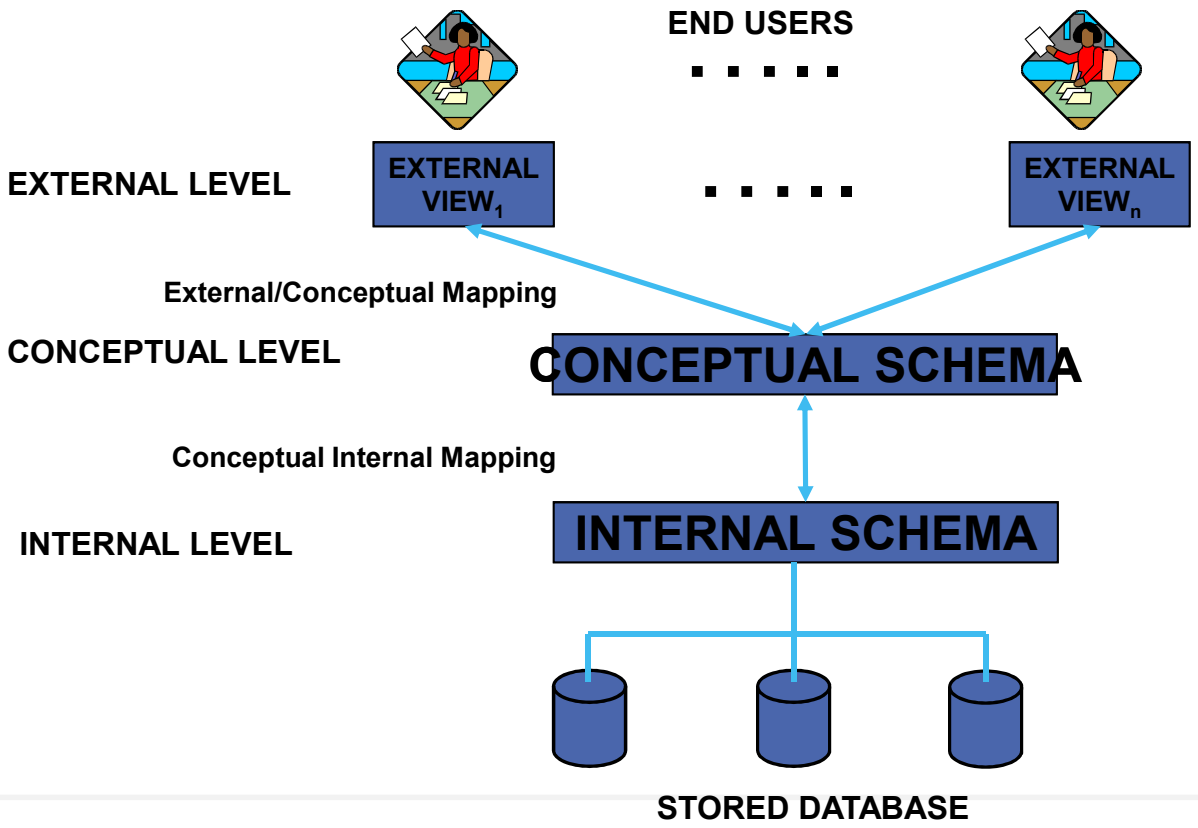
3.2.2. Characteristics of Database Approach

- Self-describing
  - **catalog** (or **meta-data**) stores the *description* of the database
  - allow the DBMS software to work with different DBs
- Data Abstraction:
  - **data model** used to hide storage details
  - present the users with a *conceptual view* of the DB
- Sharing of data
  - support multiple view of a DB
  - allow concurrent access on a DB

3.2.3. Characteristics … (cont'd)

- Persistence
  - store data on secondary storage
- Retrieval
  - a declarative query language
  - a procedural database programming language
- Performance
  - retrieve and store data quickly
  - deal with large volume of data

3.2.4. 3-tier Schema Model (ANSI-SPARC Architecture)



Remarks

- DBMS is
  - more expensive
  - more complex
  - general

When to use DBMS ?

"More than 80 % of real world computer applications are associated with databases"*

* Korth & Silberschatz. Database System Concepts.

Quiz

No	Question (Multiple Choice)	Answer (1,2,3,4)	Commentary
1	What is a shared collection of related data designed to meet the information needs of an organization? A. Data B. Database C. Database Management System D. Data model	B	
2	Which are roles of the Database Management System? A. Defining ~ specifying types of data B. Constructing ~ storing & populating C. Manipulating ~ querying, updating, reporting D. Generating ~ issuing data automatically	A, B, C	
3	Database is good for all kind of data? A. Yes B. No	B	

Outro > Summary

No	Topic	Summary
1	Course overview	- Course objective - Motivation for studying databases
2	Basic concepts	- Data - Database - Data model vs. schema vs. instance - Database management system (DBMS) - Database environment - Database users
3	Data management	- File management system approach - Database management system approach