What the course is about?

- These days, organizations are considering data as one important resource like finance, human resource and time.
- The management of these resources needs special attention and development strategy.
- Due to this reason it is considered that databases are essential to every business.
- They are used to maintain internal records, to present data to customers and clients on the World-Wide-Web, and to support many other commercial processes.
- Databases are likewise found at the core of many modern organization’s information systems.
Databases are at the heart of modern commercial Information systems application development. Their use extends beyond this to many applications and environments where large amounts of data must be stored for efficient update and retrieval.
Cont...

- Database system is about
  - Organization of data
  - Efficient retrieval of data
  - Reliable storage of data
  - Maintaining consistent data
  - Making useful information for decision making
Thus

- This course is designed to equip students with the theoretical and practical skill in
  - understanding,
  - designing,
  - developing,
  - administering and managing a database system in an organization.
Course Content

1. Introductory Concepts
   - Overview of Database System (DBS)
   - File-Based Systems
   - Database Approach
   - Advantages of Database Approach
2. **DBS Architecture**
   - Reality, Data & Metadata
   - The Three Levels of DBS Architecture
     - External View, Conceptual View, Internal View
   - Schemas, Mappings, and Instances
   - The Database Management System (DBMS)
     - Functions of DBMS
     - Requirements of DBMS with Respect to Database Languages
   - Data Definition Language, Data Manipulation Language
   - Roles in Database Environment
3. **Relational DBMS (Data Models)**
   - The Hierarchical Model
   - The Network Model
   - The Relational Model
     - Terminologies
     - Relational Algebra and Relational Calculus
     - Properties of Relation
     - Relational Integrity Rules
4. Database Application Development

4.3. Entity Relationship (E-R) Modeling

- E-R Modeling
- Structural Constraints
- Reducing E-R Model to Table
- Generalization and Specialization
4.4. **Logical Database Design**

- Functional Dependencies
- The Process of Normalization
  - First Normal Form
  - Second Normal Form
  - Third Normal Form
  - Advanced Normal Form
5. **Introduction to Query Language (Lab)**
   - Introduction to Query Language and SQL
   - Data Definition Statements
   - Data Manipulation Statements
   - View Definition
   - Query By Example (QBE)
Cont...

6. DBMS Architecture
   - Client/Server Architecture
   - Distributed Database Architecture
     - Replication
     - Fragmentation
7. Other Topics in Database Systems (Optional)
   - Data Mining
   - Distributed Database
   - Data Warehouse
   - Knowledge Base
References

Text Book
1. A practical Approach to DBS : Thomas M. Connoly

Other Books
1. An Introduction to Database System: C.J. Date
2. Fundamentals of Database System : Elmasri and Navathe
3. Database System Concepts: H.F. Korth & A. Silberschatz
4. Modern Database Management: Jeffrey A. Hoffer
## Course Assessment

<table>
<thead>
<tr>
<th>Assessment Type</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mid Term Exam</td>
<td>25%</td>
</tr>
<tr>
<td>Group Project Work</td>
<td>20%</td>
</tr>
<tr>
<td>Lab Assessment</td>
<td>20%</td>
</tr>
<tr>
<td>Final Exam</td>
<td>35%</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>
Chapter One

Introductory Concepts
Agenda (chapter one)

- Overview of Database System (DBS)
- File-Based Systems
- Database Approach
- Advantages of Database Approach
What is a database?
How it fits into the broader information management picture?
The “big” picture

- System Concepts (system and systems thinking)
- Organizations (as a big system have many subsystems)
- Information Systems- definition, functions and components
- Data/information- are core
Cont...

- Information system development - phases, methodologies
  - Process, data, logic – modeling
  - The focus might be on the process or on objects
  - But still data is important

- All approaches of information system developments, one way or another, have mechanism to design the data, process and business logic in the organization

- The focus here is on the principles of designing and developing the data in the organization
What is a Database?

- A structured collection of related data
- A filing cabinet, an address book, a telephone directory, a timetable, etc.
- Google and your email is a database
- School Student Information System
Cont...

- Database is a collection of information, usually, kept in a list, on a particular subject
- A database is a storage space for content / information (data)
- DB is used
  - To manipulate information so that it can be sorted and/or searched.
  - To make record keeping and tracking fast and efficient.
Data vs. Information

- **Data** – a collection of facts made up of text, numbers and dates:
  - *Murray*  35000  7/18/86

- **Information** - the meaning given to data in the way it is interpreted:
  - *Mr. Murray* is a sales person whose annual salary is **$35,000** and whose hire date is **July 18, 1986**.
But what is data? And where is it now?

- Data is factual information about objects and concepts, such as:
  - measurements
  - statistics
- You can find it in:
  - filing cabinets, spreadsheets, folders, ledgers, lists, colleagues’ memories, piles of papers on your desk
What is the ultimate purpose of a database system?

Is to transform and manage

Data driven decision making
What does “managing data/information” mean?

- Making information *work* for us
- Making information useful
- Avoiding "accidental disorganisation"
- Making information easily accessible and integrated with the rest of our work
Basic Definitions

- **Database:**
  - A collection of related data.

- **Data:**
  - Known facts that can be recorded and have an implicit meaning.

- **Mini-world:**
  - Some part of the real world about which data is stored in a database.

- **Database Management System (DBMS):**
  - A software package/system to facilitate the creation and maintenance of a computerized database.

- **Database System:**
  - The DBMS software together with the data itself. Sometimes, the applications are also included.
Basic Database Concepts

- **Table**
  - A set of related records

- **Record**
  - A collection of data about an individual item

- **Field**
  - A single item of data common to all records
An Example of a Table

<table>
<thead>
<tr>
<th>Name</th>
<th>GatorLink</th>
<th>Phone</th>
<th>College</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abebe</td>
<td>Hailu</td>
<td>392-3900</td>
<td>Pharmacy</td>
</tr>
<tr>
<td>Hanna</td>
<td>Ambelu</td>
<td>392-5555</td>
<td>Medicine</td>
</tr>
<tr>
<td>Hailu</td>
<td>Zemecha</td>
<td>846-5656</td>
<td>Informatics</td>
</tr>
</tbody>
</table>
A RECORD

- Is a single electronic index card.
- One row in Excel
A FIELD

- (category) is the place where one item of information is recorded; the smallest part of the database.
Where are DB Systems used?

- Backend for CBIS applications
  - Student record System
- Backend for large Websites
  - Google
- Backend for Web services
  - Amazon
Databases everywhere!

Because of the versatility of databases, we find them powering all sorts of projects:

- A web site that is capturing registered users
- A client tracking application for social service organisations
- A medical record system for a health care facility
- A collection of word processed documents
- A system that issues airline reservations
WHO USES DATABASES?

- Almost everyone:
  - Business
  - Doctors
  - Teachers
  - Students
  - ..............
Database Applications

• Banking: all transactions
• Airlines: reservations, schedules
• Universities: registration, grades
• Sales: customers, products, purchases
• Manufacturing: production, inventory, orders, supply chain
• Human resources: employee records, salaries, tax deductions
Data Handling Approaches
Data Handling Approaches

- Data management passes through the different levels of development along with the development in technology and services.
- These levels could best be described by categorizing the levels into three levels of development.
- Even though there is an advantage and a problem overcome at each new level, all methods of data handling are in use to some extent.
- The major three levels are;
  - Manual Approach
  - Traditional File Based Approach
  - Database Approach
Manual Approach

- data storage and retrieval follows the primitive and traditional way of information handling
  - cards and paper are used for the purpose.
  - The data storage and retrieval will be performed using human labour.
  - Files for as many event and objects as the organization has are used to store information.
Cont...

- Each of the files containing various kinds of information is labelled and stored in one or more cabinets.
- The cabinets could be kept in safe places for security purpose based on the sensitivity of the information contained in it.
- Insertion and retrieval is done by searching first for the right cabinet then for the right the file then the information.
- One could have an indexing system to facilitate access to the data.
Limitations of the Manual approach
- Prone to error
- Difficult to update, retrieve, integrate
- You have the data but it is difficult to compile the information
- Limited to small size information
- Cross referencing is difficult

An alternative approach of data handling is a computerized way of dealing with the information.

The computerized approach could also be either decentralized or centralized based on where the data resides in the system.
Traditional File Based Approach

- After the introduction of Computer for data processing to the business community, the need to use the device for data storage and processing increase.
- There were, and still are, several computer applications with file based processing used for the purpose of data handling.
- Even though the approach evolved over time, the basic structure is still similar if not identical.
  - File based systems were an early attempt to computerize the manual filing system.
  - This approach is the decentralized computerized data handling method.
A collection of application programs perform services for the end-users.

In such systems, every application program that provides service to end users define and manage its own data.

Such systems have number of programs for each of the different applications in the organization.

Since every application defines and manages its own data, the system is subjected to serious data duplication problem.

File, in traditional file based approach, is a collection of records which contains logically related data.

Cont...
Cont...

- As business application become more complex
  - demands more flexible and
  - reliable data handling methods

- Limitations of the Traditional File Based approach
  - Separation or Isolation of Data: Available information in one application may not be known.
  - Limited data sharing
  - Lengthy development and maintenance time
Cont...

- Duplication or redundancy of data
- Data dependency on the application
- Incompatible file formats between different applications and programs creating inconsistency.
- Fixed query processing which is defined during application development
Graphical Representation of File Based System

Sales

- Data entry and reports
- File handling routines
- File definition

Sales application programs

Sales files

Contracts

- Data entry and reports
- File handling routines
- File definition

Contracts application programs

Contracts files

Sales Files
- **Property_for_Rent** (Property Number, Street, Area, City, Post Code, Property Type, Number of Rooms, Monthly Rent, Owner Number)
- **Owner** (Owner Number, First Name, Last Name, Address, Telephone Number)
- **Renter** (Renter Number, First Name, Last Name, Address, Telephone Number, Preferred Type, Maximum Rent)

Contracts Files
- **Lease** (Lease Number, Property Number, Renter Number, Monthly Rent, Payment Method, Deposit, Paid, Rent Start Date, Rent Finish Date, Duration)
- **Property_for_Rent** (Property Number, Street, Area, City, Post Code, Monthly Rent)
- **Renter** (Renter Number, First Name, Last Name, Address, Telephone Number)
The limitations for the traditional file based data handling approach arise from two basic reasons.

Definition of the data is embedded in the application program which makes it difficult to modify the database definition easily.

No control over the access and manipulation of the data beyond that imposed by the application programs.
The most significant problem experienced by the traditional file based approach of data handling is the “update anomalies”. We have three types of update anomalies;

- **Modification Anomalies**: a problem experienced when one or more data value is modified on one application program but not on others containing the same data set.

- **Deletion Anomalies**: a problem encountered where one record set is deleted from one application but remain untouched in other application programs.

- **Insertion Anomalies**: a problem experienced when ever there is new data item to be recorded, and the recording is not made in all the applications. And when same data item is inserted at different applications, there could be errors in encoding which makes the new data item to be considered as a totally different object.
Database Approach

- Following a famous paper written by Ted Codd in 1970, concepts and development approach of database systems changed significantly.
- Codd proposed that database systems should present the user with a view of data organized as tables called relations.
- Behind the scenes, there might be a complex data structure that allowed rapid response to a variety of queries.
But, unlike the user of earlier database systems, the user of a relational system would not be concerned with the storage structure.

Queries could be expressed in a very high-level language, which greatly increased the efficiency of database programmers.

The database approach emphasizes the integration and sharing of data throughout the organization.
• Thus in Database Approach
  • Database is just a computerized record keeping system or a kind of electronic filing cabinet.
  • Database is a repository for collection of computerized data files.
  • Database is a shared collection of logically related data designed to meet the information needs of an organization.
  • Since it is a shared corporate resource, the database is integrated with minimum amount of or no duplication.
  • Database is a collection of logically related data where these logically related data comprises entities, attributes, relationships, and business rules of an organization's information.
In addition to containing data required by an organization, database also contains:

- a description of the data which called as “Metadata” or
- “Data Dictionary” or
- “Systems Catalogue” or
- “Data about Data”.

Since a database contains information about the data (metadata), it is called a self descriptive collection on integrated records.

The purpose of a database is to store information and to allow users to retrieve and update that information on demand.

Database is designed once and used simultaneously by many users.
Cont...

- Unlike the traditional file based approach in database approach there is program data independence.
- That is the separation of the data definition from the application.
- Thus the application is not affected by changes made in the data structure and file organization.
- Each database application will perform the combination of: Creating database, Reading, Updating and Deleting data.
Benefits of the database approach

- **Data can be shared**: two or more users can access and use same data instead of storing data in redundant manner for each user.

- **Improved accessibility of data**: by using structured query languages, the users can easily access data without programming experience.

- **Redundancy can be reduced**: isolated data is integrated in database to decrease the redundant data stored at different applications.
Cont...

- **Quality data can be maintained**: the different integrity constraints in the database approach will maintain the quality leading to better decision making.

- **Inconsistency can be avoided**: controlled data redundancy will avoid inconsistency of the data in the database to some extent.

- **Transaction support can be provided**: basic demands of any transaction support systems are implanted in a full scale DBMS.
Cont...

• **Integrity can be maintained**: data at different applications will be integrated together with additional constraints to facilitate shared data resource.

• **Security majors can be enforced**: the shared data can be secured by having different levels of clearance and other data security mechanisms.

• **Improved decision support**: the database will provide information useful for decision making.
Standards can be enforced: the different ways of using and dealing with data by different units of an organization can be balanced and standardized by using a database approach.

Compactness: since it is an electronic data handling method, the data is stored compactly (no voluminous papers).

Speed: data storage and retrieval is fast as it will be using the modern fast computer systems.

Less labour: unlike the other data handling methods, data maintenance will not demand much resource.

Centralized information control: since relevant data in the organization will be stored at one repository, it can be controlled and managed at the central level.
Graphical Representation of Database System

- **Sales**
  - Sales application programs
  - Data entry and reports

- **Contracts**
  - Contracts application programs
  - Data entry and reports

- **DBMS**
  - Property, Owner, Renter and Lease details + File definitions
  - Database

**Property_for_Rent** *(Property Number, Street, Area, City, Post Code, Property Type, Number of Rooms, Monthly Rent, Owner Number)*

**Owner** *(Owner Number, First Name, Last Name, Address, Telephone Number)*

**Renter** *(Renter Number, First Name, Last Name, Address, Telephone Number), Preferred Type, Maximum Rent)*

**Lease** *(Lease Number, Property Number, Renter Number, Payment Method, Deposit, Paid, Rent Start Date, Rent Finish Date)*
Cont...

- Does this mean that DB approach is a solution with out any limitation?
- What do you think about its limitation?…….
Limitations and risk of Database Approach

- Introduction of new professional and specialized personnel.
- Complexity in designing and managing data
- The cost and risk during conversion from the old to the new system
- High cost to be incurred to develop and maintain the system
- Complex backup and recover services from the users perspective
- Reduced performance due to centralization and data independency
- High impact on the system when failure occurs to the central system.
Chapter Summary

- Data are at the center of organizational functions
- Information systems are important in data handling
- Data handling - manual, file based, database approach
- Databases are important components of an information system
  - In Computer based information systems
Next on

DBMS Architecture