**ST. XAVIER’S COLLEGE**

**(Affiliated to Tribhuvan University)**

Maitighar, Kathmandu



**DATABASE MANAGEMENT SYSTEM**

**LAB ASSIGNMENT#1**

**Submitted by:**

Dibash Poudel

013BSCCSIT017

**Submitted to:**

|  |  |
| --- | --- |
| **Er. Sanjay Kumar Yadav** |  |

Lecturer

Department of Computer Science

Date of submission: 27th July, 2015

**Simplified database system organization**

Formally, a "database" refers to a set of related data and the way it is organized. Access to this data is usually provided by a "database management system" (DBMS) consisting of an integrated set of computer software that allows [users](https://en.wikipedia.org/wiki/User_(computing)) to interact with one or more databases and provides access to all of the data contained in the database (although restrictions may exist that limit access to particular data). The DBMS provides various functions that allow entry, storage and retrieval of large quantities of information as well as provides ways to manage how that information is organized.

Because of the close relationship between them, the term "database" is often used casually to refer to both a database and the DBMS used to manipulate it.

Outside the world of professional [information technology](https://en.wikipedia.org/wiki/Information_technology), the term *database* is often used to refer to any collection of related data (such as a [spreadsheet](https://en.wikipedia.org/wiki/Spreadsheet) or a card index). This article is concerned only with databases where the size and usage requirements necessitate use of a database management system.[[2]](https://en.wikipedia.org/wiki/Database#cite_note-Ullman-2)

Existing DBMSs provide various functions that allow management of a database and its data which can be classified into four main functional groups:

* **Data definition** – Creation, modification and removal of definitions that define the organization of the data.
* **Update** – Insertion, modification, and deletion of the actual data.[[3]](https://en.wikipedia.org/wiki/Database#cite_note-3)
* **Retrieval** – Providing information in a form directly usable or for further processing by other applications. The retrieved data may be made available in a form basically the same as it is stored in the database or in a new form obtained by altering or combining existing data from the database.[[4]](https://en.wikipedia.org/wiki/Database#cite_note-4)
* **Administration** – Registering and monitoring users, enforcing data security, monitoring performance, maintaining data integrity, dealing with concurrency control, and recovering information that has been corrupted by some event such as an unexpected system failure.[[5]](https://en.wikipedia.org/wiki/Database#cite_note-5)

Both a database and its DBMS conform to the principles of a particular [database model](https://en.wikipedia.org/wiki/Database_model).[[6]](https://en.wikipedia.org/wiki/Database#cite_note-6) "Database system" refers collectively to the database model, database management system, and database.[[7]](https://en.wikipedia.org/wiki/Database#cite_note-7)

Physically, database [servers](https://en.wikipedia.org/wiki/Server_(computing)) are dedicated computers that hold the actual databases and run only the DBMS and related software. Database servers are usually multiprocessor computers, with generous memory and [RAID](https://en.wikipedia.org/wiki/Redundant_array_of_independent_disks) disk arrays used for stable storage. RAID is used for recovery of data if any of the disks fail. Hardware database accelerators, connected to one or more servers via a high-speed channel, are also used in large volume transaction processing environments. DBMSs are found at the heart of most [database applications](https://en.wikipedia.org/wiki/Database_application). DBMSs may be built around a custom [multitasking](https://en.wikipedia.org/wiki/Computer_multitasking) [kernel](https://en.wikipedia.org/wiki/Kernel_(computing)) with built-in [networking](https://en.wikipedia.org/wiki/Computer_network) support, but modern DBMSs typically rely on a standard [operating system](https://en.wikipedia.org/wiki/Operating_system) to provide these functions.[[*citation needed*](https://en.wikipedia.org/wiki/Wikipedia:Citation_needed)] Since DBMSs comprise a significant [economical](https://en.wikipedia.org/wiki/Economy) [market](https://en.wikipedia.org/wiki/Market_(economics)), computer and storage vendors often take into account DBMS requirements in their own development plans.

Keep the data management system as simple as possible and within the talents of the (potential) staff. Simplicity reduces errors by reducing dependency on “key” personnel and by making the system easier to learn and implement. Computers are wonderful tools in data management, but it is easy to complicate things by their use. The use of non-user friendly software packages or uncommon packages breeds complexity. Computerized systems actually increase the cost and technical support of systems. Their benefits are in the realm of increased efficiency and (hopefully) a reduction of errors. A small Project may benefit from a predominantly, manual system when properly designed and Implemented. Databases and DBMSs can be categorized according to the database model(s) that they support (such as relational or XML), the type(s) of computer they run on (from a server cluster to a mobile phone), the [query language](https://en.wikipedia.org/wiki/Query_language)(s) used to access the database (such as SQL or [Query](https://en.wikipedia.org/wiki/XQuery)), and their internal engineering, which affects performance, [scalability](https://en.wikipedia.org/wiki/Scalability), resilience, and security.

DBMS Advantages

The feature list of a DBMS can only be considered an advantage if those features are essential in providing an effective solution to the application data management requirements. A quick summary of the features that address the questions above are:

* **Transaction Support**– Atomic transactions guarantee complete failure or success of an operation. This includes automatic recovery of the database to a transaction consistent point in the event of an abnormal termination of the application (crash, power loss, etc.).
* **Concurrent Access**– The ability to share data by controlling access to data items, many users (process or threads) can access data concurrently.
* **Data Normalization**– A well.­ designed database schema can reduce storage requirements on the target storage media by reducing duplicate data.
* **Expandability, Flexibility, Scalability**– A database system can scale easily to larger data sets.
* **Standards Enforcement**– One example of this advantage would be to use the DBMS for all data storage requirements for the application. Multiple data structures can be manipulated using the same API functions. The can lead to reduced application development times and reduced maintenance costs in the future.
* **Fast Query Access**– Databases allow indexing based on any attribute or data-property (i.e. SQL columns). This helps fast retrieval of data, based on the indexed attribute. This is an importance advantage as data-sets begin to grow large as it provides a more predictable query response time.
* **Interoperability**– Connectivity through industry standard protocols allowing third-party tools to access and analyze data.

**2. Approaches to management of Data**

Data management includes all aspects of data planning, handling, analysis, documentation and storage, and takesplace during all stages of a study. The objective is to create a reliable data base containing high quality data. Data management is a too often neglected part of study design and includes:

• Planning the data needs of the study

• Data collection

• Data entry

• Data validation and checking

• Data manipulation

• Data files backup

• Data documentation

Each of these processes requires thought and time; each requires painstaking attention to detail.

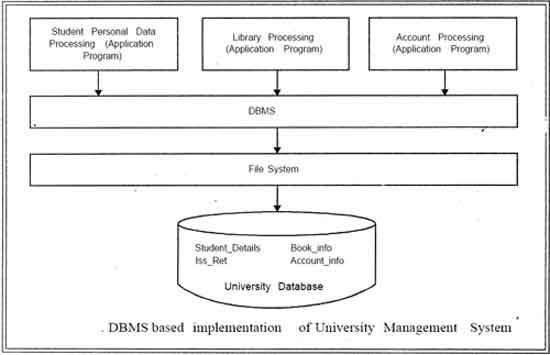
The main element of data management is database files. Database files contain text, numerical, images, and other data in machine readable form. Such files should be viewed as part of a database management systems(DBMs) which allows for a broad range of data functions, including data entry, checking, updating, Documentation and analysis.

There is different approach of data management

1. **Database Approach**

The Database is a shared collection of logically related data, designed to meet the information needs of an organization. A database is a computer based record keeping system whose over all purpose is to record and maintains information. The database is a single, large repository of data, which can be used simultaneously by many departments and users. Instead of disconnected files with redundant data, all data items are integrated with a minimum amount of duplication.

The database is no longer owned by one department but is a shared corporate resource. The database holds not only the organization's operational data but also a description of this data. For this reason, a database is also defined as a self-describing collection of integrated records. The description of the data is known as the Data Dictionary or Meta Data (the 'data about data'). It is the self-describing nature of a database that provides program-data independence.



A database implies separation of physical storage from use of the data by an application program to achieve program/data independence. Using a database system, the user or programmer or application specialist need not know the details of how the data are stored and such details are "transparent to the user". Changes (or updating) can be made to data without affecting other components of the system. These changes include, for example, change of data format or file structure or relocation from one device to another.

In the [DBMS](http://ecomputernotes.com/fundamental/what-is-a-database/advantages-and-disadvantages-of-dbms) approach, application program written in some programming language like [Java](http://ecomputernotes.com/java/what-is-java/what-is-java-explain-basic-features-of-java-language), Visual Basic.Net, and Developer 2000 etc. uses database connectivity to access the database stored in the disk with the help of [operating system](http://ecomputernotes.com/fundamental/disk-operating-system/what-is-operating-system)'s file management system.

1. **File System Approach**

File based systems were an early attempt to computerize the manual

Filing system. File based system is a collection of application programs that perform services for the end-users. Each program defines and manages its data.

However, five types of problem are occurred in using the file based approach:

**1. Separation and isolation of data:** When data is isolated in separate files, it is more difficult for us to access data that should be available. The application programmer is required to synchronize the

Processing of two or more files to ensure the correct data is extracted.

**2. Duplication of data**: When employing the decentralized file based

Approach, the uncontrolled duplication of data is occurred. Uncontrolled duplication of data is undesirable because:

i. Duplication is wasteful ii. Duplication can lead to loss of data integrity

**3. Data dependence**: Using file based system, the physical structure and storage of the data files and records are defined in the application program code. This characteristic is known as programdata dependence. Making changes to an existing structure are rather difficult and will lead to a Modification of program. Such maintenance activities are time consuming and subject to error.

**4. Incompatible file formats**: The structures of the file are dependent on the application programming language. However file structure provided in one programming language such as direct file, indexed sequential file which is available in COBOL programming, may be different from the structure generated by other programming language such as C. The direct incompatibility makes them difficult to process jointly.

**5. Fixed queries / proliferation of application programs:** File based systems are very dependent

Upon the application programmer. Any required queries or reports have to be written by the Application programmer. Normally, a fixed format query or report can only be entertained and no Facility for ahoy queries if offered.

**3. Database vs. File system approach**

A database is generally used for storing related, structured data, with well defined data formats, in an efficient manner for insert, update and/or retrieval (depending on application).  
  
On the other hand, a file system is a more unstructured data store for storing arbitrary, probably unrelated data. The file system is more general, and databases are built on top of the general data storage services provided by file systems.  
  
There are also differences in the expected level of service provided by file systems and databases. While databases must be self consistent at any instant in time (think about banks tracking money!), provide isolated transactions and durable writes, a file system provides much looser guarantees about consistency, isolation and durability. The database uses sophisticated algorithms and protocols to implement reliable storage on top of potentially unreliable file systems. It is these algorithms that make database storage more expensive in terms of processing and storage costs that make general file systems an attractive option for data that does not require the extra guarantees provided by a database.

1. Data Abstraction
2. Reliability
3. Efficiency/ Performance