Master of Business Administration

(Open and Distance Learning Mode)

Semester – I

IT Applications for Business

Centre for Distance and Online Education (CDOE)

DEVI AHILYA VISHWAVIDYALAYA, INDORE

"A+" Grade Accredited by NAAC

IET Campus, Khandwa Road, Indore - 452001

www.cdoedavv.ac.in

www.dde.dauniv.ac.in

Program Coordinator

Dr. Manishkant Arya
Centre for Distance and Online Education (CDOE)
Devi Ahilya Vishwavidyalaya, Indore – 452001

Content Design Committee

Dr. Sangita Jain
Institute of Management Studies
Devi Ahilya Vishwavidyalaya, Indore – 452001

Dr. Yamini Karmarkar
Institute of Management Studies
Devi Ahilya Vishwavidyalaya, Indore – 452001

Dr. Geeta Neema
International Institute of Professional Studies
Devi Ahilya Vishwavidyalaya, Indore – 452001

Dr. Manishkant Arya
Centre for Distance and Online Education (CDOE)
Devi Ahilya Vishwavidyalaya, Indore - 452001

Language Editors

Dr. Arti Sharan
Institute of Engineering & Technology
Devi Ahilya Vishwavidyalaya, Indore – 452001

Dr. Ruchi Singh
Institute of Engineering & Technology
Devi Ahilya Vishwavidyalaya, Indore – 452001

SLM Author(s)

Ms. Shikha Jain
MBA
EMRC, Devi Ahilya Vishwavidyalaya, Indore – 452001

Mr.Vinay Mishra
MBA
EMRC, Devi Ahilya Vishwavidyalaya, Indore – 452001

Edition : 2022 (Restricted Circulation)
Published by : Centre for Distance and Online Education (CDOE), Devi Ahilya Vishwavidyalaya
Printed at : University Press, Devi Ahilya Vishwavidyalaya, Indore – 452001

IT Applications for Business

Table of Content

Chapter 3: Information for IS

3.0 Objectives

3.1Information Technology for IS

3.2 Database Management: Basic concepts

3.3Data Models

3.4 Advantages of Database Approach

3.5 Overview of E-R Modeling

3.6 Entity

3.7 Attributes

3.8 Relationships: Key Concepts

3.9 Normalization

3.10 Basic Normal Forms

3.11 overview of SQL

3.12 Summary

3.13 Glossary

3.14 References

3.15 Further Readings

3.16 Model Questions

Chapter 4: Developing Access Database

4.0 Objectives

4.1Developing Access Database:

4.2 Defining fields for a table

4.3 Choosing required data types

4.4 Defining fields properties

4.5 setting primary key

4.6 defining indexes

4.7 Saving the table

4.8 Modifying table structure

4.9 Establishing relationship among tables

4.10 entering and viewing data from tables

4.11creating queries by example

4.12 running a query

4.13 Summary

4.14 Glossary

4.15 Model Questions

4.16 References

Chapter 5: Building IS

5.0 Objectives

5.1 Basics of MS-Access Forms and Reports

5.2 Developing Database Applications and DSS

5.3 Excel Data Analysis Tools like

5.4 Goal seeking

5.5 Sensitivity Analysis

5.6 Filtering

5.7 Solver

5.8 Summary

5.9 Glossary

5.10 References

5.11 Further Readings

5.12 Model Questions

Chapter 6: Information Systems in Business

6.0 Objectives

6.1 Decision making

6.2 Types of Decision

6.3 Decision-Making Levels

6.4 Transaction Processing System (TPS)

6.5 Management Information System (MIS)

6.6 Executive Information System (EIS)

6.7 Decision Support System (DSS)

6.8 Expert System (ES)

6.9 Office Automation System (OAS)

Chapter 9: NETWORK TOPOLOGIES

Chapter 10: WEB TECHNOLOGIES

IT Application for Managers

UNIT-I Information System (IS) Concepts

Structure

1.0 OBJECTIVES

After studying this lesson you will be able to:

1. Understand the term data and information
2. Understand the system and different types of system
3. Understand the information system
4. Understand the importance of quality of information

1.1Understanding Data and Information

Data: Data is raw facts. Data is like raw material. Data does not interrelate and also it does not help in decision making. Data is defined as groups of nonrandom symbols in the form of text, images, voice representing quantities, action and objects. It can also be defined as information in raw or unorganized forms (such as alphabets, numbers or symbols) that refer to, or represent, conditions, ideas or objects. The data is limitless in nature and present everywhere in the universe.

Information: Information is the product of data processing. Information is interrelated data. Information is equivalent to finished goods produced after processing the raw material. The information has a value in

decision making. Information brings clarity and creates an intelligent human response in the mind. **Information** is that which "informs". In other words, it answers the question of some kind. It is thus related to data and knowledge, as data represents values attributed to parameters, and knowledge signifies understanding of real things or abstract concepts. As it regards data, the information's existence is not necessarily coupled to an observer (it exists beyond an event horizon, for example), while in the case of knowledge, the information requires a cognitive observer.

Harsh and colleague define information as one of four types and all these types are important component of a management information system. Furthermore, the various types build upon and interact with each other. A common starting level is Descriptive information. (See Figure below).

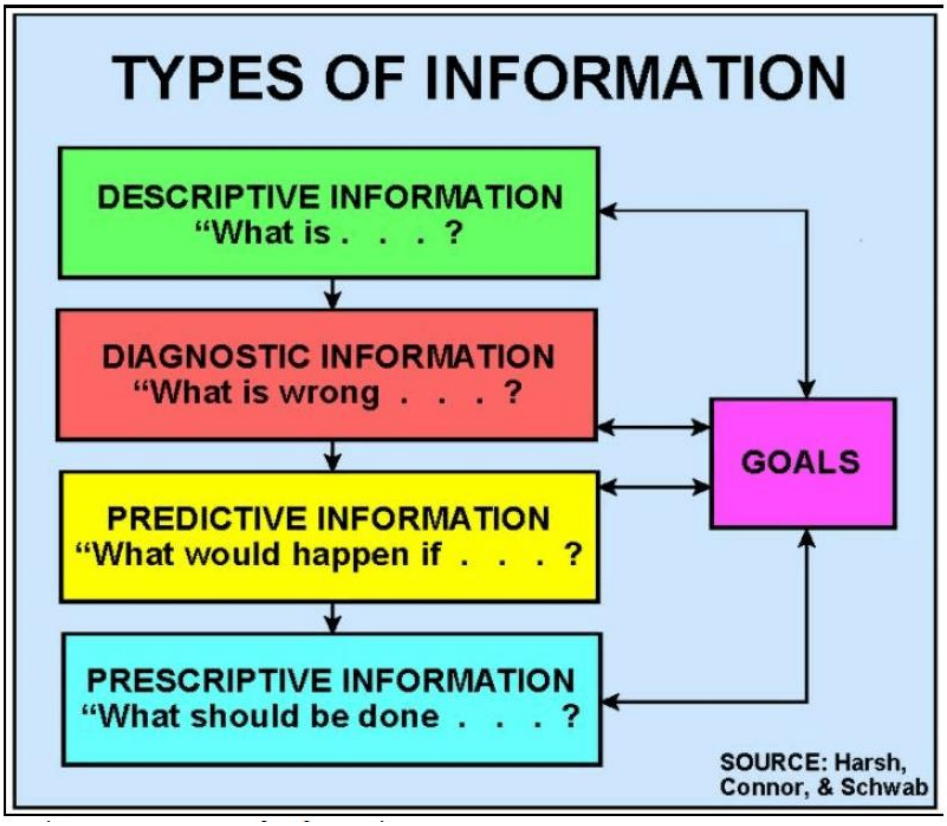


Fig: types of Information

This information portrays the "what is" condition of a business, and it describes the state of the business at a specified point in time. Descriptive information is very important to the business manager, because without it, many problems would not be identified. Descriptive information includes a variety of types of information including financial results, production records, test results, product marketing, and maintenance records. Information can be encoded into various forms for transmission and interpretation (for example, information may be encoded into a sequence of signs, or transmitted via a sequence of signals). It can also be encrypted for safe storage and communication.

Descriptive information can also be used as inputs to secure other needed types of information. For example, “what is” information is needed for supplying restraints in analyzing farm adjustment alternatives. It can also be used to identify problems other than the “what is” condition. Descriptive information is necessary but not completely sufficient in identifying and addressing farm management problems. The second type of information is diagnostic information. This information portrays this “what is wrong” condition, where “what is wrong” is measured as the disparity between “what is” and “what ought to be.” This assessment of how things are versus how they should be (a fact-value conflict) is probably our most common management problem. Diagnostic information has two major uses. It can first be used to define problems that develop in the business. Are production levels too low? Is the rate earned on investment too low? These types of question cannot be answered with descriptive information alone (such as with financial and production records). A manager may often be well supplied with facts about his business, yet be unable to recognize this type of problem.

The manager must provide norms or standards which, when compared with the facts for a particular business, will reveal an area of concern. Once a problem has been identified, a manager may choose an appropriate course of action for dealing with the problem (including doing nothing).

Corrective measures may be taken so as to better achieve the manager’s goals. Several pitfalls are involved for managers in obtaining diagnostic information. Adequate, reliable, descriptive information must be available along with appropriate norms or standards for particular business situations. Information is inadequate for problem solving if it does not fully describe both “what is” and “what ought to be.”As description is concerned with “what is” and diagnostics with ”what is wrong,” prediction is concerned with “what if...?”

Predictive information is generated from an analysis of possible future events and is exceedingly valuable with “desirable” outcomes. With predictive information, one either defines problems or avoids problems in advance. Prediction also assists in analysis. When a problem is recognized, a manager will have to analyze the situation and specify at least one alternative (including doing nothing) to deal with it.

Predictive information is needed by managers to reduce the risk and uncertainty concerning technology, prices, climate, institutions, and human relationships affecting the business. Such information is vital in formulating production plans and examining related financial impacts. Predictive information takes many forms. What are the expected prices next year? What yields are anticipated? How much capital will be required to upgrade production technologies? What would be the difference in expected returns in switching from a livestock farm to a cropping farm? Management has long used various budgeting techniques, simulation models, and other tools to evaluate expected changes in the business.

Without detracting from the importance of problem identification and analysis in management, the crux of management tasks is decision making. For every problem a manager faces, there is a “right” course of action. However, the rightness of a decision can seldom, if ever, be measured in absolute terms. The choice is conditionally right, depending upon a farm manager’s knowledge, assumptions, and conditions he wishes to impose on the decision.

Prescriptive information is directed toward answering the “what should be done” question. Provision of this information requires the utilization of the predictive information. Predictive information by itself is not adequate for decision making. An evaluation of the predicted outcomes together with the goals and values of the manger provides that basis for making a decision.

For example, suppose that a manager is considering a new changing marketing alternative. The new alternative being considered has higher “predicted” returns but also has higher risks and requires more management monitoring. The decision as to whether to change plans depends upon the managers evaluation of the worth of additional income versus the commitment of additional time and higher risk. Thus, the goals and values of a farm manager will ultimately enter into any decision.

Types of Information

Managers in an organization require information to perform various activities for achieving various targets. The information is mostly differentiated on the behalf of its source of origin, source of generating information and types of decision taken at different levels.

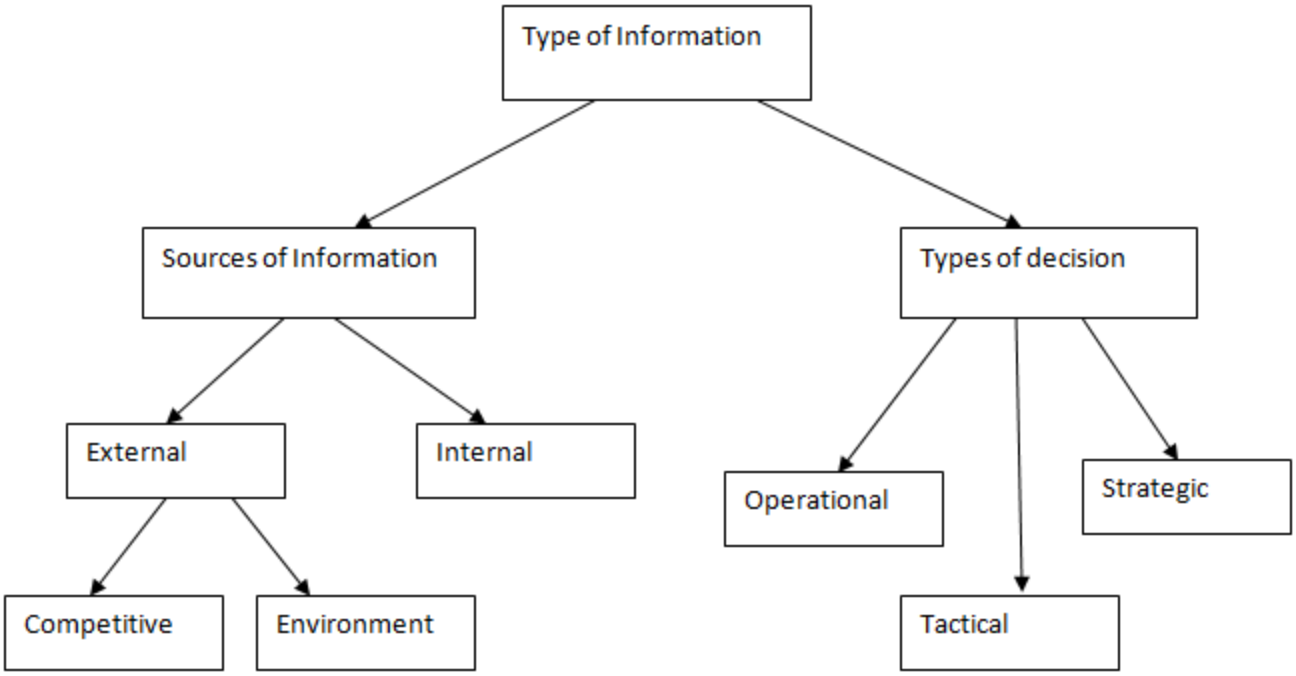


Fig: 3.1 Types of information

Much of the information used by management concerns the **internal** operations of the company. However, **external** information about the environment in which the organization exists is crucial to all organizations. This may include :

- Intelligence gathering about competitors' activities
- Information about population shifts
- Economic and social factors
- Government legislation

This type of information is of great importance to managers who are trying to shave production costs, find new markets, develop new products, or have strategic decisions to make about the future direction of the company.

Information is collected in many ways – through conversations and interpersonal 'networking', reading newspapers, trade reviews and magazines, attending conferences and meetings, browsing the Internet.

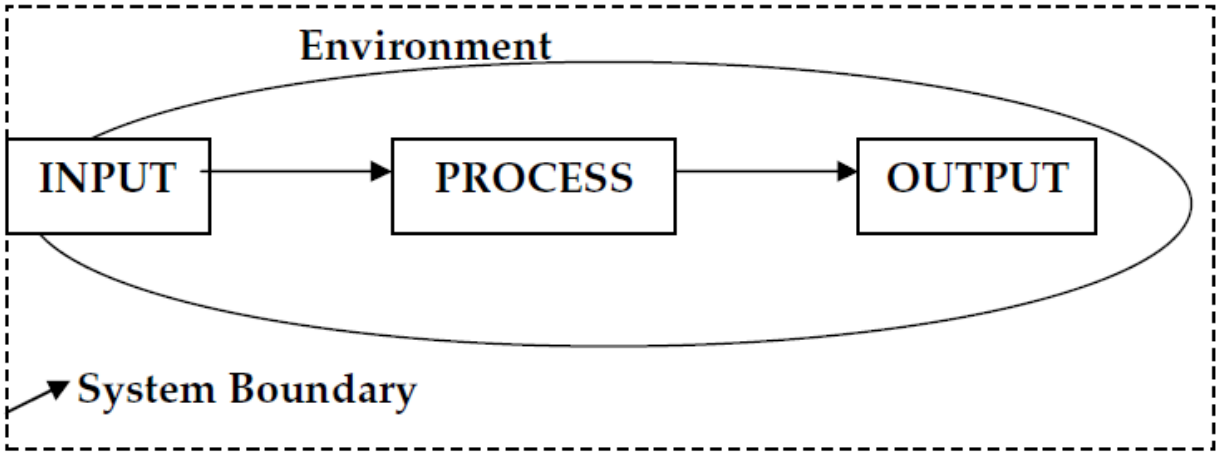
A **formal information system** relies on procedures for the collecting, storing, processing and accessing of data in order to obtain information.

1.2 System Concepts

System is defined as "A system is an orderly grouping of interdependent components linked together according to a plan to achieve a specific goal."

The word component may refer to physical parts (engine, wheels of cars), management steps (planning, organizing, controlling) or a sub subsystem in a multi level structure. It is to be noted that a system is not a randomly arranged set. It is arranged with some logic governed by rules, regulation, principles and policies.

the man-made system involving input, process and output, as represented in figure. A system may have multiple inputs and multiple outputs. All systems operate in an environment. The environment may influence the system in its design and performance. When a system is designed to achieve certain objective, it automatically sets the boundaries for itself. The understanding of boundaries of the system is essential to bring clarity in explaining the system components and their arrangement.



System Definition

"A system is an orderly grouping of interdependent components linked together according to a plan to achieve a specific goal.". it is set of detailed methods, procedures and routines created to carry out a specific activity, perform a duty or solve a problem. It is organizes, purposeful structure that consists of interrelated and interdependent elements (components, entities, factors, members etc).

These elements continually influence one another (directly or indirectly) to maintain their activity and the existence of the system, in order to achieve the goal of the system.
All systems have:

(a) Inputs, outputs and feedback mechanisms,

(b) Maintain an internal steady-state (called homeostasis) despite a changing external environment,

(c) Display properties that are different than the whole (called emergent properties) but are not possessed by any of the individual elements, and

(d) Have boundaries that are usually defined by the system observer.

A system can be defined as a collection of components that work together towards a common goal. The objective of a system is to receive inputs and transform these into outputs. Not every system has a single goal and often a system contains several subsystems with sub-goals, all contributing to meeting the overall system goal. In the systems data are used as the input for a process that creates information as an output. In order to monitor the performance of the system, some kind of feedback mechanism is required. The control must be exerted to correct any problems that occur and ensure that the system is fulfilling its purpose. There are thus five components of a generic system in terms of input, process, output, feedback and control.

Overview of the systems life cycle

Large systems development projects may involve dozens of people working over several months or even years, so they cannot be allowed to proceed in a haphazard fashion. The goals of an information system must be thoroughly understood, and formal procedures and methods applied to ensure that the project is delivered on time and to the required specification.

The systems life cycle methodology approaches the development of information systems in a very methodical and sequential manner. Each stage is composed of certain well-defined activities and responsibilities, and is completed before the next stage begins. This approach was popular in the 1960s and 70s, when systems were largely transaction-processing systems and had a much heavier reliance on programming than most modern information systems, which are database-oriented.

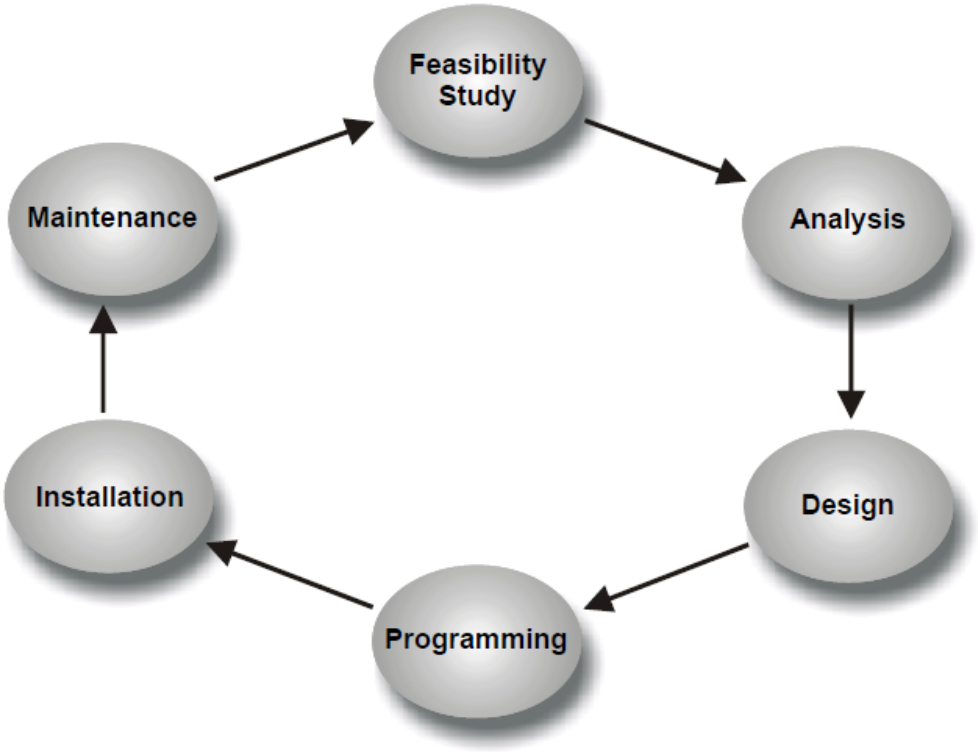


Fig: the system life cycle

1.3 System Types

At the most basic level, systems are divided into two categories: (1) Closed systems: theoretical systems that do not interact with the environment and are not influenced by its surroundings. Only the components within the system are significant. Example: a sealed jar--nothing enters or exits the jar, but whatever is inside can interact. (2) Open systems: real-world systems whose boundaries allow exchanges of energy, material and information with the larger external environment or system in which they exist.

Physical or Abstract: Physical system is tangible entities that may be static or dynamic in nature. Abstract system is conceptual or non-physical. The abstract is conceptualization of physical situations.

Open and Closed: An open system continually interacts with its environment. It receives input from the outside and delivers output to outside. A closed system is isolated from environment influences

Sub System and Super System: Each system is part of a large system. The business firm is viewed as the system or total system when focus is on production, distribution of goal and sources of profit and income. The total system consists of all the objects, attributes and relationship necessary to accomplish an objective given a number of constraints. Sub systems are the smaller systems within a system. Super system denotes extremely large and complex system

Permanent and Temporary System: A permanent system is a system enduring for a time span that is long relative to the operation of human. Temporary system is one having a short time span.

Natural and Man Made System: System which is made by man is called man made system. Systems which are in the environment made by nature are called natural system.

Deterministic and Probabilistic: A Deterministic system is one in which the occurrence of all events is perfectly predictable. If we get the description of the system state at a particular time, the next state can be easily predicted.

Probabilistic system is one in which the occurrence of events cannot be perfectly predicted.

Man-made Information System: It is generally believed that the information reduces uncertainty about a state or event. An information system is the basis for interaction between the user and the analyst. It determines the nature of relationship among decision makers. An information system may be defined as a set of devices, procedures and operating system designed around user-base criteria to produce information and communicating it to the user for planning control and performance.

System Characteristics

Following characteristics are present in all system:

a) Organization

b) Interaction

c) Interdependence

d) Integration

e) Central Objective

Organization : Organization implies structure and order. It is the arrangement of components that helps to achieve objectives. Hierarchical relationship starting with the president on top and leading down ward to the blue collar worker represent the organization structure

Interaction : Interaction refers to the procedure in which each component interact with other components of the system. In an organization, for example purchasing must interact with product, advertising with sales and payroll with personnel.

Interdependence: Independence is concerned with how a system is tied together; it is more than sharing a physical part or location. It means that parts of the system part or location within the system, even through each part performance. A unique function successful integration will typically produce a better request as whole rather than if each component works independently.

Central Objective: Objective may be real or stated. Objective is determined by higher management and user must be aware about the central objective well in advance.

1.4 Business as a system

With the previous definitions of information and systems we can now define a business information systems as a group of interrelated components that work collectively to carry out input, processing, output, storage and control actions in order to convert data into information products that can be used to support forecasting, planning, control, coordination, decision making operational activities in an organization. In terms of the components that undertake this activity, they can be classified into five basic resources of people, hardware, software, communications and data. People resources include the users and developers of an information system and those who help maintain and operate the system such as IS managers and technical support staff. Hardware resources include computers and other items such as printers. Software resources refer to computer programs known as software and associated instructions manuals. Communications resources include networks and the hardware and software needed to support them. Data resources cover the data that an organization has access to such as computer databases and paper files.

In most organizations Business Information Systems(BIS) make extensive use of information technology such as personal computers. The reasons why computerized BIS have become widespread are evident in their advantages such as speed, accuracy and dependability. They also have a high degree of flexibility due to their ability to be programmed to carry out a wide variety of tasks. There are however some disadvantages to BIS such as lack of creativity that human possess and the difficulty of incorporating other factors into their decision making such as innovation and intuition.

Information systems may be divided into two categories of systems that support an organization's day to day business activities and systems that support managerial decision making. Operations Information Systems (OIS) are generally concerned with process control, transaction processing and communications. Management Information Systems (MIS) are concerned with providing support to managerial decision making. Recently this division of BIS into operational and management systems, although useful for managers reviewing the types of BIS in use, does not now accurately reflect the reality of systems used within an organizations, particularly with the increased use of inter-organizational e-commerce and electronic data interchange(EDI). For example e-business systems and enterprise resource planning systems cut across both operational and management systems to provide business with more integrated information systems.

Business Information Systems (BIS) contains within it not just the technology that drives business. It also relates to the processes, data and the people that use these technologies, processes and data every day. Today, BIS is involved in almost every aspect of doing business and it is about using IT in useful and effective ways to improve how business operate. Studying BIS gives students the opportunity to understand how businesses operate and benefit from using new technologies. Learning how to implement, use, and manage various types of information systems equips students with the advantage of knowing how to creatively solve business problems by employing current technologies in new and innovative ways. That being said, we should keep a couple of things in mind when we think of BIS:

1. BIS ≠ IT

   BIS is not about building technology, but about knowing how to use it! BIS is a highly diverse, extremely valuable, but unfortunately unfamiliar discipline to most students, as it is often equated with computing. However, employees will find that BIS and IT are not only operating in completely different teams, they tend to sit on different floors in the same organization as well. A good way of looking at BIS is to consider the two most important aspects of businesses these days: employees get the business side that runs day to day activities, and also get the technology side that facilitates these activities. People working within a BIS-type team are like the "translators" who sit in between the two components - translating business needs into IT requirements

2. BIS = variety

For the BIS professional, no two days are ever exactly the same. BIS students learn how BIS concepts and practices support many aspects of today's businesses, especially solving complex business problems in accounting, economics, finance, marketing, human resources, to name a few. Core areas include information management, business innovation, business process improvement, enterprise systems, and project management. Because BIS is everywhere and in everything we do, user can choose to combine it with whatever their heart desires - nothing functions without BIS anymore!

3. BIS = creativity

In BIS, there is not always one solution to a problem. Replicating the current solution is not an option if are attempting to IMPROVE a process and if user want to stand out from the competition, therefore user needs to think about meeting business requirements in an innovative way.

4. BIS = challenging

BIS is an ever changing field. New technologies emerge, new processes are developed, and no two clients are alike. This means that throughout will always be challenged with new situations and new solutions will have to be developed. There is no such thing as a "standard fit" in BIS. Some situations will call on to understand a company's processes before you can make recommendations.

1.5 what is information System?

An **information system** (**IS**) is a system composed of people and computers that processes or interprets information. The plural term ***information systems*** (construed as singular) is also used for the actual academic study of the field, in other words for the study of complementary networks of hardware and software that people and organizations use to collect, filter, process, create and distribute data. Any specific information system aims to support operations, management and decision making. In a broad sense, the term is used to refer not only to the information and communication technology (ICT) that an organization uses, but also to the way in which people interact with this technology in support of business processes.

In an information system, input data consist of facts and figures, which form the systems raw material. Information is data that has been usefully processed. However, an information system does not only contain

data and information. There are also other elements inside the system, which are related and are in support of one another. The presence of these related elements makes information more useful whereby, it can be made available, can be processed, distributed, manipulated, saved, and so on. This combination gives rise to a system, which is orderly and as such it is called an "Information System".

The activity of converting data into information is called a process. An information system contains FIVE main components the hardware, software, data, process and human

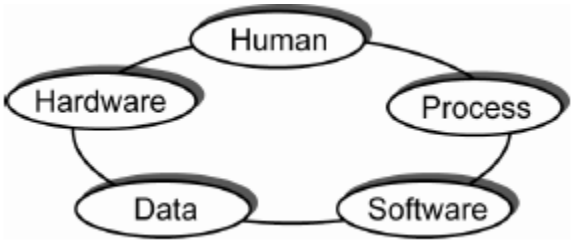


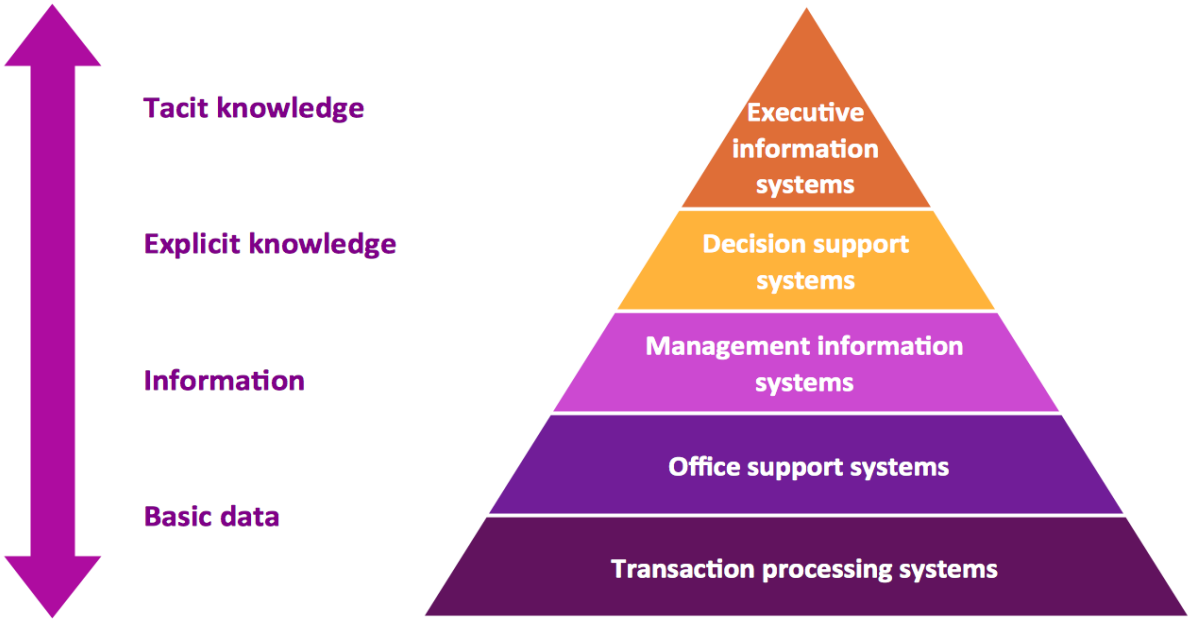
Figure: Components of Information System

A business has several information systems **:**

(A) Formal Information System

(B) Informal Information System

(C) Computer Based Information System



Formal Information System :

It is based on organizational chart represented by the organization.

Informal Information System :

It is an employee based system designed to meet personal and vocational needs and to help in the solution of work-related problems. It also funnels information upward through indirect channels. It works within the framework of the business and its stated policies.

Computer Based Information System (CBIS) :

Definition: A computer-based information system, or CBIS, uses computers to collect, process, store, analyze and distribute information for a specific purpose, such as meeting a business objective. The main components of a CBIS include hardware, software, data, procedures and people. In a CBIS, the hardware is the physical machinery, such as a computer, printer, display screen and cables. The hardware devices work together to accept data, or raw facts, as input before processing the data into useful information and then displaying the information as output. Software refers to computer programs that provide instructions for processing the data. The procedures are what people perform when working with a CBIS to process data and produce information. Traditionally, a business uses four types of computer-based information systems. The one used by the most people in the business is a transaction processing system, which processes events or activities that affect the entire organization, such as customer orders. A management information system provides information to managers so that they can plan and run the business. Managers also use decision support systems when making major decisions that affect the entire business. Senior executives use a special type of decision support system known as an executive information system, which provides an overview of the entire business.

This category of information system depends mainly on the computer for handling business application. System analysis develops different types of information system to meet variety of business needs. There is class of system collectively known as computer based information system. They can be classified as:

· Transaction Processing System (TPS)
· Management Information System (MIS)
· Decision Making System (DSS)
· Office Automation System (OAS)

- *Management information systems*, produce fixed, regularly scheduled reports based on data extracted and summarized from the firm's underlying transaction processing systems to middle and operational level managers to identify and inform structured and semi-structured decision problems. **Management information system** (**MIS**) refers to the processing of information through computers and other intelligent devices to manage and support managerial decisions within an organization.

- *Decision support systems (DSS)* are computer program applications used by middle and higher management to compile information from a wide range of sources to support problem solving and decision making. A DSS is used mostly for semi-structured and unstructured decision problems. A decision support system (DSS) is a **computerized** information system used to support decision-making in an organization or a business. A DSS lets users sift through and analyze massive reams of data and compile information that can be used to solve problems and make better decisions.
- *Executive information systems (EIS)* is a reporting tool that provides quick access to summarized reports coming from all company levels and departments such as accounting, human resources and operations. It is basically software that allows users to transform enterprise data into quickly accessible and **executive**-level reports, such as those used by billing, accounting and staffing departments. An ESS enhances decision making for **executives**.
- *Marketing Information Systems* are Management Information Systems designed specifically for managing the marketing aspects of the business. An organized approach to the study of the information needs of an organization's management at every level in making operational, tactical, and strategic decisions. Its objective is to design and implement procedures, processes, and routines that provide suitably detailed reports in an accurate, consistent, and timely manner.
  In a management information system, modern, computerized systems continuously gather relevant data, both from inside and outside an organization. This data is then processed, integrated, and stored in a centralized database (or data warehouse) where it is constantly updated and made available to all who have the authority to access it, in a form that suits their purpose.
- *Office automation systems (OAS)* support communication and productivity in the enterprise by automating workflow and eliminating bottlenecks. **Office automation** refers to the varied computer machinery and software used to digitally create, collect, store, manipulate, and relay office information needed for accomplishing basic tasks. Raw data storage, electronic transfer, and the management of electronic business information comprise the basic activities of an office automation system. Office automation helps in optimizing or automating existing office procedures. OAS may be implemented at any and all levels of management. It has configurations of networked computer hardware and software. A variety of office automation systems are now applied to business and communication functions that used to be performed manually or in multiple locations of a company, such as preparing written communications and strategic planning. In addition, functions that once required coordinating the expertise of outside specialists in typesetting, printing, or electronic recording can now be integrated into the everyday work of an organization, saving both time and money. *School Information management systems* (SIMS) cover school administration, and often including teaching and learning materials.

- *Enterprise resource planning* facilitates the flow of information between all business functions inside the boundaries of the organization and manage the connections to outside stakeholders. Enterprise resource planning (**ERP**) is business process management software that allows an organization to use a system of integrated applications to manage the business and automate many back office functions related to technology, services and human resources.

The following are some of the benefits that can be attained using MISs

- Companies are able to identify their strengths and weaknesses due to the presence of revenue reports, employees' performance record etc. Identifying these aspects can help a company improve its business processes and operations.
- Giving an overall picture of the company.
- Acting as a communication and planning tool.
- The availability of customer data and feedback can help the company to align its business processes according to the needs of its customers. The effective management of customer data can help the company to perform direct marketing and promotion activities.
- MISs can help a company gain a competitive advantage. Competitive advantage is a firm's ability to do something better, faster, cheaper, or uniquely, when compared with rival firms in the market.

ACTIVITIES

1. Take an organization of your knowledge and give a conceptual view of MIS and physical view of MIS.
2. Identify the nature of impact of MIS on people, organization and the management style.

The 5 components that must come together in orders to produce a computer-based information system are:

1. Hardware: The term hardware refers to machinery. This category includes the computer itself, which is often referred to as the central processing unit (CPU), and all of its support equipments. Among the support equipments are input and output devices, storage devices and communications devices. Information system's hardware refers to all types of hardware and the media used for input, processing, managing, distributing and saving information that are being used in an organisation. Examples of the hardware are the physical computers, networks, communication equipment, scanners, digital drives, and so on. Basic hardware for a certain computer consists of four main elements as shown in Figure

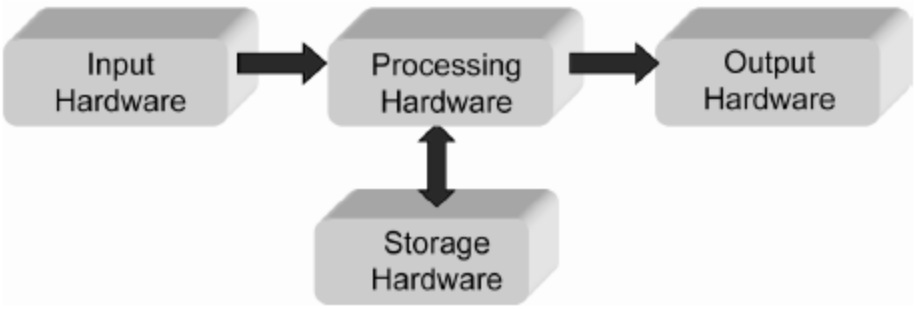


Figure: Basic hardware of a computer

2. Software: The term software refers to computer programs and the manuals (if any) that support them. Computer programs are machine-readable instructions that direct the circuitry within the hardware parts of the system to function in ways that produce useful information from data.

Programs are generally stored on some input / output medium, often a disk or tape. Software consists of two categories the system software and the application software.

- System Software controls the computer and contains the operating system and device drivers, which can communicate with the hardware. It can also modify data into a new form, prevent viruses and make copies.
- Application Software contains programs which can help users and enable companies to perform business functions. Users can increase productivity with the presence of application software such as spreadsheets, word processing, ordering systems, and accounts receivable.

3. Data: Data are facts that are used by programs to produce useful information. Like programs, data are generally stored in machine-readable form on disk or tape until the computer needs them. Data refers to the raw facts on any thing or entities like student names, courses and marks. The raw data that has not yet been provided can be processed to become more useful information. Information is an organized, meaningful and useful interpretation of data such as a company’s performances or a student's academic performance. Information systems change data into information, which is useful and capable of giving a certain meaning to its users.

4. Procedures: Procedures are the policies that govern the operation of a computer system. "Procedures are to people what software is to hardware" is a common analogy that is used to illustrate the role of procedures in a system. Process or procedure explains the activities carried out by users, managers and staff. Process is important for supporting a certain business model available as written documents or as reference materials on-line. The procedure for using a certain matter is very wide and very important to ensure that it can be implemented with success. All the information system components contain management and implementation procedures on their own, and they are different from each other.

5. People: Every system needs people if it is to be useful. Often the most over-looked element of the system are the people, probably the component that most influence the success or failure of information systems. The main objective of an information system is to provide invaluable information to managers and users, whether inside or outside the company. Users can be broken up into three categories, which are:

- End-Users, consisting of the staff, customers, suppliers and others who communicate with the information system.
- Internal Users, including the managers, technicians, sales representatives and corporate officers.
- External Users, consisting of the customers who use the company's system for performing transactions, suppliers who use the system for planning sales, and the staff who use the system outside office hours.

The success or failure of an information system depends on whether the system that has been developed can fulfill the user's requirements, and the users feel satisfied with the results and the system's operation. A successful system requires integrated efforts from information technology experts such as the system analysts, programmers and the information technology managers so as to fulfill business needs and to support company's objectives.

Developing Information Systems

The actions that are taken to create an information system that solves an organizational problem are called system development". These include:

system analysis, system design, computer programming/implementation, testing, conversion, production and finally maintenance.

Conversion is the process of changing or converting the old system into the new. This can be done in three basic ways, though newer methods (prototyping, Extreme Programming, JAD, etc.) are replacing these traditional conversion methods in many cases:

- Direct cut – The new system replaces the old at an appointed time.
- Pilot study -– Introducing the new system to a small portion of the operation to see how it fares. If good then the new system expands to the rest of the company.
- A formal information system is based on the organization represented by the organization chart. The chart is a map of position and their authority relationship, indicated by boxes and connected by straight lines. it is concerned with the pattern of authority, communication and work flow.
- The informal information system is employee based system design to meet personnel and vocational needs and to help in the solution of work-related problems. it also funnels information upward through indirect channels.
- Gaining competitive advantage is critical for organizations. Baltzan and Phillips defined competitive advantage as 'a product or service that an organization's customers value more highly than similar offerings from its competitors' (in other words, you have something useful (i.e. products, services, capabilities) that your competitors do not have). Competitive advantages are typically temporary as competitors often seek ways to duplicate the competitive advantage. In order to stay ahead of competition, organizations have to continually develop new competitive advantages.
- Michael Porter's Five Forces Model is a useful tool to assist in assessing the competition in an industry and determining the relative attractiveness of that industry. Porter states that in order to do an industry analysis a firm must analyze five competitive forces:
- Rivalry of competitors within its industry
- Threat of new entrants into an industry and its markets
- Threat posed by substitute products which might capture market share
- Bargaining power of customers
- Bargaining power of suppliers.

Table: Competitive Strategies & Roles of Information Systems

Competitive Strategy	Roles of Information Systems
Cost Leadership	Organizations can use information systems to fundamentally shift the cost of doing business or reduce the costs of business processes or/and to lower the costs of customers or suppliers, i.e., using online business to consumer & business to business models, e-procurement systems to

	reduce operating costs
Differentiation	Organizations can use information systems to develop differentiated features or/and to reduce competitors' differentiation advantages, i.e., using online live chatting systems and social networks to better understand and serve customers; using technology to create inform diaries to offer value-added service and improve customers' stickiness to your web site/business; applying advanced and established measures for online operations to offline practices (i.e., more accurate and systematic ways of measuring efficiency and effectiveness of advertising).
Innovation	Organizations can use information systems to identify and create (or assist in creating) new products and services or/and to develop new/niche markets or/and to radically change business processes via automation (i.e., using digital modeling and simulation of product design to reduce the time and cost to the market. They also can work on new initiatives of establishing pure online businesses/operations. At the same time, the Internet and telecommunications networks provide better capabilities and opportunities for innovation. "Combinational innovation" and Open innovation are two good examples. There are a large number of component parts on the networks that are very expensive or extremely different before the establishment of the networks, and organizations could combine or recombine components/parts on the networks to create new innovations. Meanwhile everyone is connected via personal computers, laptops and other mobile devices through cabled Internet or wireless networks or mobile networks, there are plenty of opportunities to co-create with customers, external partners and internal people.
Growth (including mergers and acquisitions)	Organizations can use information systems to expand domestic and international operations or/and to diversify and integrate into other products and services, i.e., establishing global intranet and global operation platform; establishing omni-channel strategy to gain growth(omni-channel strategy looks at leveraging advantages of both online (or digital) and offline (or non-digital) channels).
Strategic Alliance	Organizations can use information systems to create and enhance relations with partners via applications, such as developing virtual organizations

	and inter-organizational information systems.

SELF ASSESSMENT QUESTIONS

Please mark the correct option

A. Which one of the following is not a business driver for an information system?
   1. business process redesign
   2. knowledge asset management
   3. proliferation of networks and the Internet
   4. security and privacy
   5. collaboration and partnership

B. A task of developing a technical blueprint and specifications for a solution that fulfills the business requirements is undertaken in the following phase of the system development process
   1. system initiation
   2. system implementation
   3. system analysis
   4. system design
   5. feasibility analysis

C. If a university sets up a web-based information system that faculty could access to record student grades and to advise students, that would be an example of a/an:
   1. CRM
   2. intranet
   3. ERP
   4. extranet
   5. none of the above

D. Which of the following is not a technology driver for an information system?
   1. enterprise applications
   2. object technologies
   3. knowledge asset management
   4. collaborative technologies
   5. networks and the Internet

1.6 Summary

Information System is a business application in the computer. It is made up of the database, application programs and manual and machine procedures. It also encompasses the computer systems that do the processing.

Processing the Data: The database stores the subjects of the business (master files) and its activities (transaction files). The application programs provide the data entry, updating, query and report processing.

The Procedures: The manual procedures document how data are obtained for input and how the system's

output is distributed. Machine procedures instruct the computer how to perform scheduled activities, in which the output of one program is automatically fed into another.

Transaction Processing: The daily work is the online, interactive processing of the business transactions and updating of customer, inventory and vendor files (master files).

Batch Processing: At the end of the day or other period, programs print reports and update files that were not updated on a daily basis. Periodically, files must be updated for routine maintenance such as adding and deleting employees and making changes to product descriptions.

Advantages Of Information System

Management information systems have changed the dynamics of running businesses efficiently. Decentralization is one of the biggest advantages; it allows monitoring of operations at low levels and frees up resources for departmental managers to devote time to strategic activities. Coordination of specialized projects and activities is much better and decision makers in the organization are aware of issues and problems in all departments. Another advantage of MIS is that it minimizes information overload, which can be quite common with conventional businesses in the modern era.

Better Planning and Control

MIS has to be designed and managed in such way that it aggregates information, monitors the company's activities and operations and enhances communication and collaboration among employees. This ensures better planning for all activities and better ways to measure performance, manage resources and facilitate compliance with industry and government regulations. Control helps in forecasting, preparing accurate budgets and providing the tools and vital information to employees, top management and business partners.

Aid Decision Making

The purpose of MIS is to generate synthesized and processed information from computerized/automated and certain manual systems. Information distribution to all levels of corporate managers, professionals and key executives becomes quite seamless with streamlined MIS. Managers are able to make quick, timely and informed decisions. Top management and board members can take strategic decisions, plan future growth and business expansion activities based on the data and information generated by MIS.

ERP system – sometimes also referred to as “basic system” or “operations system”, the “ERP” abbreviation comes from English and means “Enterprise Resource Planning”. This term refers to systems used for

planning and management of enterprise resources. Such systems are commonly used in everyday operations of the company (therefore are sometimes referred to as “operations systems”). Typical areas these systems focus on include accounting, stock details, human resources and others.

MIS, management information system – while ERP system serves for “normal daily operations”, MIS is usually system for purposes of management. Its power lies mainly in a different approach to data and different options of data presentation. It also offers some highly sophisticated tools for planning. The MIS almost always works in collaboration with ERP system, which provides some of the data for MIS.

Disadvantages of Information System

Depending on organization deployment, usage and extraneous factors, some disadvantages related to Management Information Systems can come to the fore. Allocation of budgets for MIS upgrades, modifications and other revisions can be quite tricky at times. If budgets are not allocated uniformly or as per immediate requirements, key functionalities might get affected and benefits might not be realized consistently. Integration issues with legacy systems can affect the quality of output and vital business intelligence reports.

Constant Monitoring Issues

Change in management, exits or departures of department managers and other senior executives has a broad effect on the working and monitoring of certain organization practices including MIS systems. Since MIS is a critical component of an organization's risk management strategy and allied systems, constant monitoring is necessary to ensure its effectiveness. Quality of inputs into MIS needs to be monitored; otherwise consistency in the quality of data and information generated gets affected. Managers are not able to direct business, operational and decision-making activities with the requisite flexibility.

1.7 Glossary

Information system: (**IS**) is a system composed of people and computers that processes or interprets information. The plural term ***information systems*** (construed as singular) is also used for the actual academic study of the field, in other words for the study of complementary networks of hardware and software that people and organizations use to collect, filter, process, create and distribute data.

Management Information System: A management information system (MIS) is a computerized database of financial information organized and programmed in such a way that it produces regular reports on operations for every level of management in a company. It is usually also possible to obtain special reports from the system easily. The main purpose of the MIS is to give managers feedback about their own performance; top management can monitor the company as a whole. Information displayed by the MIS typically shows "actual" data over against "planned" results and results from a year before; thus it measures progress against goals.

Marketing Research: **Marketing Research** is systematic problem analysis, model building and fact finding for the purpose of important decision making and control in the **marketing** of goods and services.

Management records: Records management, also known as records and information management, is an organizational function devoted to the management of information in an organization throughout its life cycle, from the time of creation or inscription to its eventual disposition. This includes identifying, classifying, storing, securing, retrieving, tracking and destroying or permanently preserving records.

1.8 References

1. O'Brien, J (1999). *Management Information Systems – Managing Information Technology in the Internetworked Enterprise*. Boston: Irwin McGraw-Hill. ISBN 0-07-112373-3.
2. Laudon, Kenneth C.; Laudon, Jane P. (2009). *Management Information Systems: Managing the Digital Firm* (11 ed.). Prentice Hall/CourseSmart. p. 164.
3. *Transaction processing systems (TPS)* collect and record the routine transactions of an organization. Examples of such systems are sales order entry, hotel reservations, payroll, employee record keeping, and shipping.
4. Bidgoli, Hossein, (2004). The Internet Encyclopedia, Volume 1, John Wiley & Sons, Inc. p. 707.
5. Pant, S., Hsu, C., (1995), Strategic Information Systems Planning: A Review, Information Resources Management Association International Conference, May 21–24, Atlanta.

1.9 Further Readings

6. Taylor, Victoria. "Supply Chain Management: The Next Big Thing?". *Sept. 12, 2011*. Business Week. Retrieved 5 March 2014.
7. Lynn, Samara. "What is CRM?". PC Mag. Retrieved 5 March 2014.
8. Joshi, Girdhar (2013). *Management Information Systems*. New Delhi: Oxford University Press. p. 328. ISBN 9780198080992.
9. Laudon, K.,&Laudon, J. (2010). Management information systems: Managing the digital firm. (11th ed.). Upper Saddle River, NJ: Pearson Prentice Hall

1.10 Model Questions

1. Differentiate between MIS, TPS, OAS.

2.What are the challenges of Information Systems?

3. What is common in DSS and ESS?

4. State the difference between MIS and a computer system?

ANSWERS TO SELF ASSESSMENT QUESTIONS

A. 3
B. 4
C. 2
D. 3

CHAPTER 2

Structure

2.0 Objectives

After studying this chapter you will be able :

1. Understand different types of Information System
2. Understand about the Attributes of Information Quality
3. Understand about the information system development

2.1 Classification of IS in organization

Information system – in general context, refer to any system capable of providing information to its' users.. Information system is, for example, accounting in the form of "paper" account books, as it used to be done in the past. Information system is used more specifically and, unless expressly stated to represent something else or to represent a software product for supporting information systems.

The different categories of information system are:

Formal Information System:

A formal information system is based on the organization represented by the organization chart. The chart is a map of position and their authority relationship, indicated by boxes and connected by straight lines. it is concerned with the pattern of authority, communication and work flow.

Informal Information System:

The informal information system is employee based system design to meet personnel and vocational needs and to help in the solution of work-related problems. it also funnels information upward through indirect channels. In this way, it is considered to be a useful system because it works within the framework of the business and its stated policies.

Public and Private Information System

Public information system helps to provide information to public agencies like census, hospitals etc to take decisions whereas private information system provide information to private agencies to understand the customer feedback, competitors analysis etc helpful for the companies to understand the market trend.

Computer Based Information System (CBIS) : This category of information system depends mainly on the computer for handling business applications. System analyst develops different types of information systems to meet variety of business needs. There is a class of system collectively known as computer based information system.

They can be classified as

· Transaction Processing System (TPS)

· Management Information System(MIS)

· Decision Support System (DSS)

· Office Automation System (OAS)

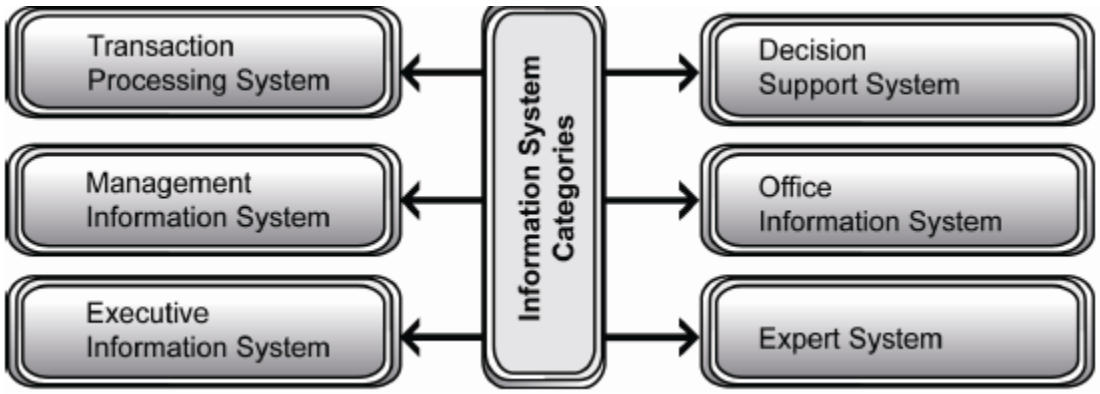


Fig: Categories of Information System

Transaction Processing System (TPS) : The most fundamental computer based system in an organization pertains to the processing of business transactions. A transaction processing system can be defined as a system that captures, classifies, stores, maintains, updates and retrieves transaction data for record keeping and input to the other types of CBIS. Transaction Processing System is aimed at improving the routine business activities. A transaction is any event or activity that affects the whole organization. Placing order, billing customers, hiring of employees and depositing cheaques are some of the common transactions. Types of transactions that occur vary from organization to organization but this is true that all organizations process transaction as a major part of their daily business activities. Transaction Processing System provides speed and accuracy and can be programmed to follow routines without any variance.

Management Information System (MIS) : Data processing by computers has been extremely effective because of several reasons. The main reason is that huge amount of data relating to accounts and other transactions can be processed very quickly. MIS are more concerned with levels of management with information essential to the running of smooth business. This Information must be as relevant, timely, accurate, complete and concise as is economically feasible.

Decision Support System (DSS) : It is an information system that offers the kind of information that may not be predictable. Business professionals may need such information only once. These systems do not produce regularly scheduled management reports. Instead, they are designed to respond to wide range of requests. It is true that all the decisions in an organization are not of a recurring nature. Decision support systems assist managers, who make decisions that are not highly structured, often called unstructured or semi structured decision. The decision support systems support, but do not replace, judgments of managers.

Office Automation System (OAS) : Office Automation Systems are among thenewest and most rapidly expanding computer based information systems. They are being developed with the hope and expectation that they will increase the efficiency and productivity of office workers, typists, secretaries, administrative assistants, staff professionals, managers and others.

Benefits if IS for organizations:

Information systems gain their importance by processing the data from company inputs to generate information that is useful for managing your operations. To increase the information system's effectiveness, you can either add more data to make the information more accurate or use the information in new ways.

Communication

Part of management is gathering and distributing information, and information systems can make this process more efficient by allowing managers to communicate rapidly. Email is quick and effective, but managers can use information systems even more efficiently by storing documents in folders that they share with the employees who need the information. This type of communication lets employees collaborate in a systematic way. Each employee can communicate additional information by making changes that the system tracks. The manager collects the inputs and sends the newly revised document to his target audience.

Operations

How to manage the company's operations depends on the information you have. Information systems can offer more complete and more recent information, allowing you to operate your company more efficiently. The information systems can use this to gain a cost advantage over competitors or to differentiate it by offering better customer service. Sales data give you insights about what customers are buying and let you stock or produce items that are selling well. With guidance from the information system, you can streamline your operations.

Decisions

The company information system can help to make better decisions by delivering all the information to managers need and by modeling the results of the decisions. A decision involves choosing a course of action from several alternatives and carrying out the corresponding tasks. When there is available have accurate, up-to-date information, manager can make the choice with confidence. If more than one choice looks appealing, manager can use the information system to run different scenarios. For each possibility, the system can calculate key indicators such as sales, costs and profits to help you determine which alternative gives the most beneficial result.

Records

The company needs records of its activities for financial and regulatory purposes as well as for finding the causes of problems and taking corrective action. The information system stores documents and revision histories, communication records and operational data. The trick to exploiting this recording capability is organizing the data and using the system to process and present it as useful historical information. Manger can use such information to prepare cost estimates and forecasts and to analyze how the actions affected the key company indicators.

ACTIVITIES

1. Analyze the business as a system and identify the components of a business system.
2. Discuss the characteristics of information systems in various levels of your organization.

2.2 Attributes of Information Quality.

Information quality" is a measure of the value which the information provides to the user of that information. "Quality" is often perceived as subjective and the quality of information can then vary among users and among uses of the information. Nevertheless, a high degree of quality increases its objectivity or at least the inter subjectivity. Accuracy can be seen as just one element of Information Quality but, depending upon how it is defined, can also be seen as encompassing many other dimensions of quality.

A list of dimensions or elements used in assessing subjective Information Quality is:

- Intrinsic Information Quality: Accuracy, Objectivity, Believability, Reputation
- Contextual Information Quality: Relevancy, Value-Added, Timeliness, Completeness, Amount of information
- Representational Information Quality: Interpretability, Format, Coherence, Compatibility
- Accessibility Information Quality: Accessibility, Access security

While information as a distinct term has various ambiguous definitions, there's one which is more general, such as "description of events". While the occurrences being described cannot be subjectively evaluated for quality, since they're very much autonomous events in space and time, their description can—since it possesses a garnishment attribute, unavoidably attached by the medium which carried the information, from the initial moment of the occurrences being described.

2.3 Information System Development: concept

The information system development includes the Problem Solving and System Development. There are four steps to build an information system.

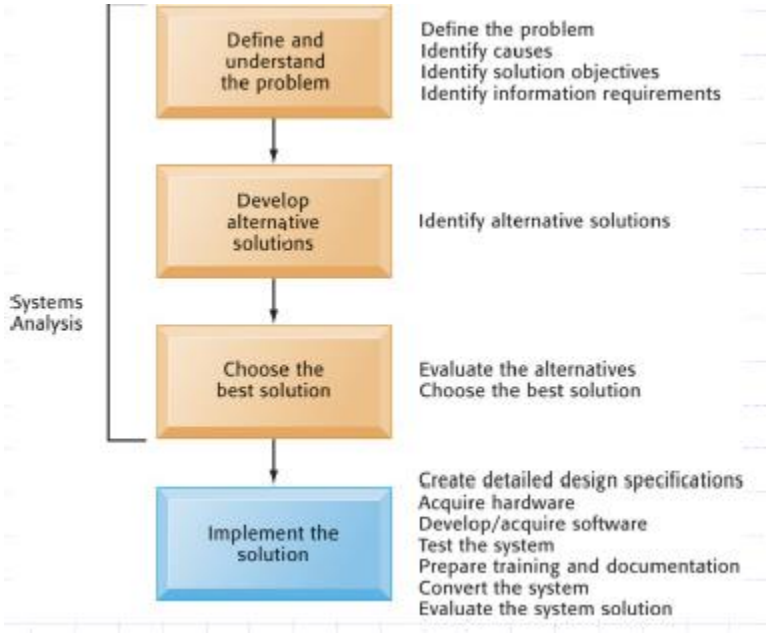


Figure: Problem Solving Process

Four steps to building an information system

1. Define and understand the problem
2. Develop alternative solutions
3. Evaluate and choose a solution
4. Implement the solution

1. ☒ First Three Steps → System Analysis
2. Defining and Understanding Problem

   What caused the problem?

   Why does it persist?

   Why hasn't it been solved?

   What are the objectives of a solution?

   Information requirements

   Who needs what Information, where, when, how

   Ability to rapidly total and organize order trans., Track orders by…

3. Developing Alternative Solutions
   - Paths to a solution determined by systems analysis
   - Some solutions do not require an information system
   - Some solutions require modification of existing systems
   - Some solutions require new systems

4. Evaluating and Choosing Solutions
   - Feasibility issues {financial, technical, organizational}
   - Costs and benefits
   - Advantages and disadvantages {Patriot trait Girls Scout had 3 alternatives}
   - Business value of systems
   - Change management

5. Implementing the Solution
   - Systems design {Create detailed design spec}
   - Completing implementation
   - Hardware selection and acquisition
   - Software development and programming
   - Testing – *Unit, System, Acceptance*
   - Training, documentation – *End-user and Technical* {Online practice, step-by-step ins}

- Conversion – *Changing from Old to New System* {Parallel, Direct Cutover, Phased}
- Production & maintenance – *Completing Conversion* {Review, Objectives, Modification}
- Managing the change {Introduce in orderly and effective manner, training, answering questions}

SELF ASSESSMENT QUESTIONS

Please mark the correct option

A. Which of the following is a deliverable of the system implementation phase in a formal system development process?
   1. technical hardware and software solution for the business problem
   2. business problem statement
   3. statement of the system users' business requirements
   4. technical blueprint and specifications for a solution that fulfills the business requirements
   5. none of the above

B. An information system that supports the planning and assessment needs of executive management is:
   1. DSS
   2. TPS
   3. ERP
   4. MIS
   5. none of the above

C. Decision makers who are concerned with tactical (short-term) operational problems and decision making are:
   1. middle managers
   2. executive managers
   3. supervisors
   4. mobile managers
   5. none of the above

D. The application of information to scan an organisation's environment is:
   1. external communication.
   2. information overload.
   3. sensing.
   4. internal communication.
   5. none of the above.

E. When a bank uses information to launch a personalised credit card product this: 1. manages risks.
2. creates a new opportunity.
3. adds value.
4. reduces costs.
5. none of the above.

2.4 Approaches to IS Development

1. System Development Life Cycle (SDLC)

System Development Life Cycle: System Development life cycle (SDLC) is used to plan and manage the system development process. Although it is primarily identified with structured analysis, the SDLC describes activities and functions that systems developers typically perform, regardless of how those activities and functions fit into a particular methodology.

The system development life cycle have following steps of development **:**

i) Systems Planning

ii) Systems Analysis

iii) Systems Design

iv) Systems Implementation

v) Systems Operation and Support (System Maintenance)

The systems life cycle approach to development is also known as the 'waterfall model', and a variation on the basic diagram as shown in Figure. Note that the arrows go up and down the 'waterfall', reflecting the fact that developers often have to rework earlier stages in the light of experience gained as development progresses.

A project milestone terminates each stage of a life-cycle-oriented approach. At this stage, the 'deliverable' resulting from that stage – such as the documentation for the analysis or the design, or the program code or finished database application, is *signed off* by all concerned parties and approval is given to proceed. The 'concerned parties' usually include the end-users, the management and the developers, as well as other experts such as database administration personnel. This sequence continues until the evaluation stage has been completed and the finished system is delivered to the end-users.

In this model, the end-user has very little say in the development process, which is carried out by technical specialists such as systems analysts and programmers. He or she is presented with the finished system at the end of the development cycle and if it is not quite what was wanted, it is generally too late to make changes. Therefore, it is extremely important that the system requirements are very clearly specified and understood by all parties before being signed off.

Such levels of certainty are difficult to achieve and this is one of the major drawbacks of the 'waterfall model'.

Advantages of Waterfall Iterative Model

1) Waterfall model is simple to implement and also the amount of resources required for it are minimal.

2) In this model, output is generated after each stage (as seen before), therefore it has high visibility. The client and project manager gets a feel that there is considerable progress. Here it is important to note that in any project psychological factors also play an important role.

3) Project management, both at internal level and client's level, is easy again because of visible outputs after each phase. Deadlines can be set for the completion of each phase and evaluation can be done from time to time, to check if project is going as per milestones.

4) This methodology is significantly better than the haphazard approach to develop software. It provides a template into which methods of analysis, design, coding, testing and maintenance can be placed.

5) This methodology is preferred in projects where quality is more important as compared to schedule or cost.

Disadvantages of Waterfall Iterative Model of SDLC

1) Real projects rarely follow the sequential flow and iterations in this model are handled indirectly. These changes can cause confusion as the project proceeds.

2) It is often difficult to get customer requirements explicitly. Thus specifications can't be free zed. If that case arises baseline approach is followed, wherein output of one phase is carried forward to next phase. For example, even if SRS is not well defined and requirements can't be free zed, still design starts. Now if any changes are made in SRS then formal procedure is followed to put those changes in baseline document.

3) In this model we freeze software and hardware. But as technology changes at a rapid pace, such freezing is not advisable especially in long-term projects.

4) This method is especially bad in case client is not IT-literate as getting specifications from such a person is tough.

5) Even a small change in any previous stage can cause big problem for subsequent phases as all phases are dependent on each-other.

6) Going back a phase or two can be a costly affair.

Projects where Waterfall Method is suitable for SDLC:-

1) In development of database-related software, eg commercial projects.

2) In development of E-commerce website or portal.
3) In Development of network protocol software

2. Prototyping

Prototyping is the process of creating an incomplete model of the future full-featured system, which can be used to let the users have a first idea of the completed program or allow the clients to evaluate the program.

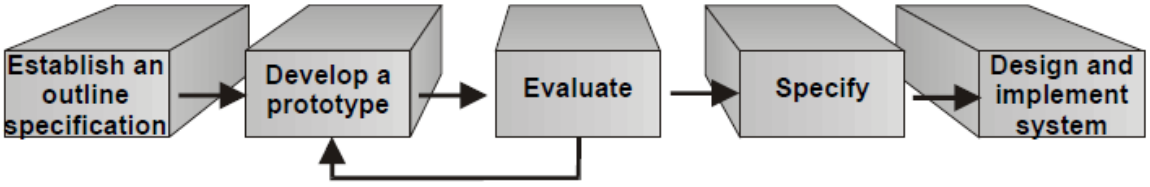
The waterfall model of the system life cycle has major shortcomings and often bears little relation to what happens in practice.

One reason for this is that it doesn't allow for modifications to the design as the project proceeds, with both user and developer learning as they go along. Users frequently have difficulty in explaining their requirements at the start of a proposed system since they do not know what is possible and cannot visualize how the final system will work. This can result in a system which does not really
match their requirements. (See Figure)

Using the **prototyping** approach, a model of a new system is built in order to evaluate it or have it approved before building the production model. Applied to software projects, this means, for example, using special software to quickly design input screens and create a program to input and validate data.

This gives the user a chance to experience the 'look and feel' of the input process and suggest alterations before going any further. The earlier a user is involved, the easier it will be to make changes.

Fig: Prototyping Model



Benefits of prototyping

The benefits of prototyping are:

- Misunderstandings between software developers and users can be identified when the prototype is demonstrated;
- Missing functions may be detected;
- Incomplete or inconsistent user requirements may be detected and can be completed or corrected;

- A prototype version will be quickly available to demonstrate the feasibility and usefulness of the proposed system to management;
- The prototype can sometimes be used for training before the final system is delivered.

Prototyping may be used in a number of different ways, and various terms have been coined to describe them:

- **Piloting** – using a prototype to test the feasibility of a design proposal;
- **Modeling** – building to develop an understanding of the user's requirements;
- **Throw-away prototyping** – both piloting and modeling are 'throw-away prototypes': once they have achieved their purpose the real system is built;

Evolutionary prototyping – each prototype built represents a step closer to the final solution.

Advantages :

i) The designer and implementer can obtain feedback from the users early in the project development.

ii) The client and the contractor can compare that the developing system matches with the system specification, according to which the system is built.

iii) It also gives the engineer some idea about the accuracy of initial project estimates and whether the deadlines can be successfully met.

Disadvantages of Prototyping :

i) **Insufficient Analysis:** Since a model has to be created, developers will not properly analyze the complete project. This may lead to a poor prototype and a final project that will not satisfy the users.

ii) **User Confusion of Prototype and Finished System:** Users can begin to think that a prototype, intended to be thrown away, is actually a final system that merely needs to be finished or polished. Users can also become attached to features that were included in a prototype for consideration and then removed from the specification for a final system.

iii) **Excessive Development Time of the Prototype:** A key property to prototyping is the fact that it is supposed to be done quickly. If the developers forget about this fact, they will develop a prototype that is too complex.

iv) **Expense of Implementing Prototyping:** The start up costs for building a development team focused on prototyping may be high. Many companies have to train the team for this purpose which needs extra expenses.

3. RAD

Rapid application development (**RAD**) is both a general term used to refer to alternatives to the conventional waterfall model of software development as well as the name for James Martin's approach to rapid development. In general, RAD approaches to software development put less emphasis on planning tasks and more emphasis on development. In contrast to the waterfall model, which emphasizes rigorous specification and planning, RAD approaches emphasize the necessity of adjusting requirements in reaction to knowledge gained as the project progresses. This causes RAD to use prototypes in addition to or even sometimes in place of design specifications. RAD approaches also emphasize a flexible process that can adapt as the project evolves rather than rigorously defining specifications and plans correctly from the start. In addition to James Martin's RAD methodology, other approaches to rapid development include Agile methods and the spiral model. RAD is especially well suited (although not limited to) developing software that is driven by user interface requirements. Graphical user interface builders are often called rapid application development tools.

The James Martin approach to RAD divides the process into four distinct phases:

1. **Requirements Planning phase** – combines elements of the system planning and systems analysis phases of the Systems Development Life Cycle (SDLC). Users, managers, and IT staff members discuss and agree on business needs, project scope, constraints, and system requirements. It ends when the team agrees on the key issues and obtains management authorization to continue.
2. **User design phase** – during this phase, users interact with systems analysts and develop models and prototypes that represent all system processes, inputs, and outputs. The RAD groups or subgroups typically use a combination of Joint Application Development (JAD) techniques and CASE tools to translate user needs into working models. *User Design* is a continuous interactive process that allows users to understand, modify, and eventually approve a working model of the system that meets their needs.
3. **Construction phase** – focuses on program and application development task similar to the SDLC. In RAD, however, users continue to participate and can still suggest changes or improvements as actual screens or reports are developed. Its tasks are programming and application development, coding, unit-integration and system testing.
4. **Cutover phase** – resembles the final tasks in the SDLC implementation phase, including data conversion, testing, changeover to the new system, and user training. Compared with traditional methods, the entire process is compressed. As a result, the new system is built, delivered, and placed in operation much sooner.

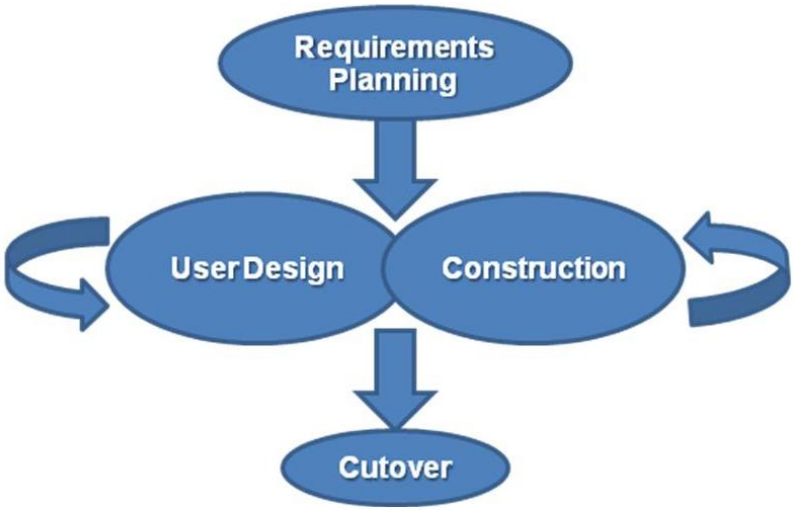


Figure: Phases in RAD

Pros and Cons of Rapid Application Development

In the modern Information Technology environment many systems are now built using some degree of Rapid Application Development. Not necessarily the James Martin approach. In addition to Martin's methodology Agile methods and the Rational Unified Process are often used for RAD development.

The advantages of RAD include:

- Better Quality. By having users interact with evolving prototypes the business functionality from a RAD project can often be much higher than that achieved via a waterfall model. The software can be more usable and has a better chance to focus on business problems that are critical to end users rather than technical problems of interest to developers.
- Risk Control. Although much of the literature on RAD focuses on speed and user involvement a critical feature of RAD done correctly is risk mitigation. It's worth remembering that Boehm initially characterized the spiral model as a risk based approach. A RAD approach can focus in early on the key risk factors and adjust to them based on empirical evidence collected in the early part of the process. E.g., the complexity of prototyping some of the most complex parts of the system.
- More projects completed on time and within budget. By focusing on the development of incremental units the chances for catastrophic failures that have dogged large waterfall projects is reduced. In the Waterfall model it was common to come to a realization after six months or more of analysis and development that required a radical rethinking of the entire system. With RAD this kind of information can be discovered and acted upon earlier in the process.

The disadvantages of RAD include:

- The risk of a new approach. For most IT shops RAD was a new approach that required experienced professionals to rethink the way they worked. Humans are virtually always averse to change and any project undertaken with new tools or methods will be more likely to fail the first time simply due to the requirement for the team to learn.
- Requires time of scarce resources. One thing virtually all approaches to RAD have in common is that there is much more interaction throughout the entire life-cycle between users and developers. In the waterfall model, users would define requirements and then mostly go away as developers created the system. In RAD users are involved from the beginning and through virtually the entire project. This requires that the business is willing to invest the time of application domain experts. The paradox is that the better the expert, the more they are familiar with their domain, the more they are required to actually run the business and it may be difficult to convince their supervisors to invest their time. Without such commitments RAD projects will not succeed.
- Less control. One of the advantages of RAD is that it provides a flexible adaptable process. The idea is to be able to adapt quickly to both problems and opportunities. There is an inevitable trade-off between flexibility and control, more of one means less of the other. If a project (e.g. life-critical software) values control more than agility RAD is not appropriate.
- Poor design. The focus on prototypes can be taken too far in some cases resulting in a "hack and test" methodology where developers are constantly making minor changes to individual components and ignoring system architecture issues that could result in a better overall design. This can especially be an issue for methodologies such as Martin's that focus so heavily on the User Interface of the system.
- Very large systems. RAD typically focuses on small to medium sized project teams. The other issues cited above (less design and control) present special challenges when using a RAD approach for very large scale systems

2.5 Phases in SDLC

Phases in SDLC,

The system development life cycle have following steps of development **:**

i) Systems Planning

ii) Systems Analysis

iii) Systems Design

iv) Systems Implementation

v) Systems Operation and Support (System Maintenance)

Systems Planning: A system's planning usually begins with a formal request to the IT department, called a system's request that describes problems or desired changes in an information system or a business process. A system's request can come from a top manager, a planning team, a department head, or the IT department itself. The request can be very significant or relatively minor. A major request might involve a new information system or the replacement of an existing system that cannot handle current requirements. In contrast, a minor request might ask for a new feature or a change to the user interface in current system. The purpose of the planning phase is to identify clearly the nature and scope of the business opportunity or problem by performing a preliminary investigation, often called a feasibility study. The preliminary investigation is a critical step because the outcome will affect the entire development process. The end product, or deliverable, is a report that describes business considerations, reviews anticipated benefits and costs, and recommends a course of action based on economic, technical, and operational factors.

Systems Analysis: The purpose of the systems analysis phase is to understand business requirement and build a logical model of the new system. The first step is requirement modeling, where you define and describe business process. Requirement modeling continues the investigation that began during systems planning and involves various fact-finding techniques, such as interview, surveys, observation, and sampling. During the next tasks, data modeling, process modeling, and object modeling, you develop a logical model of business process the system must support. The model consists of various types of diagrams, depending on the methodology being used.

The end product for the systems analysis phase is the System Requirements Document. The systems requirements document describes management and user requirements, alternative plans and costs, and analysis your recommendation. Looking ahead to design and implementation, several possibilities exist: develop a new system in-house, purchase a commercial package, or modify an existing system.

Systems Design: The purpose of systems design is to create a blueprint for the new system that will satisfy all documented requirements, whether the system is being developed in-house or purchased as a package. During systems design, identify all necessary outputs, inputs, interfaces, and processes. In addition, design internal and external controls, including maintainable, and secure. The design is documented in the System Design Specification and presented to management and users for their review and approval. Management and user involvement is critical to avoid any misunderstandings about what the new system will do, how it will do it, and what it will cost.

Systems Implementation: During systems implementation, the new system is constructed. Programs are written, tested, and documented, and the system is installed. If the system was purchased as a package, systems analysts perform any necessary modifications and configurations. The objective of the

implementation phase is to deliver a completely functioning and documented information system. At the Summary of this phase, the system is ready for use. Final preparations include converting data to the new system's files, training of users, and performing the actual transition to the new system. The systems implementation phase also includes and assessment, called a systems evaluation, to determine whether the system operates properly and its costs and benefits are within expectations.

Systems Operation and Support (Maintenance): During systems operation and support, the IT staff maintains and enhances the system. Maintenance change correct errors and adapt to changes in the environment, such as new tax rates. Enhancements provide new features and benefits. The objective during this phase is to maximize return on the IT investment. A well-designed system will be reliable, maintainable, and scalable. A scalable design can expand to meet new business requirements and volumes. Information systems development is always a work in progress. Business process change rapidly, and most information systems need to be replaced or significantly updated after several years of operation.

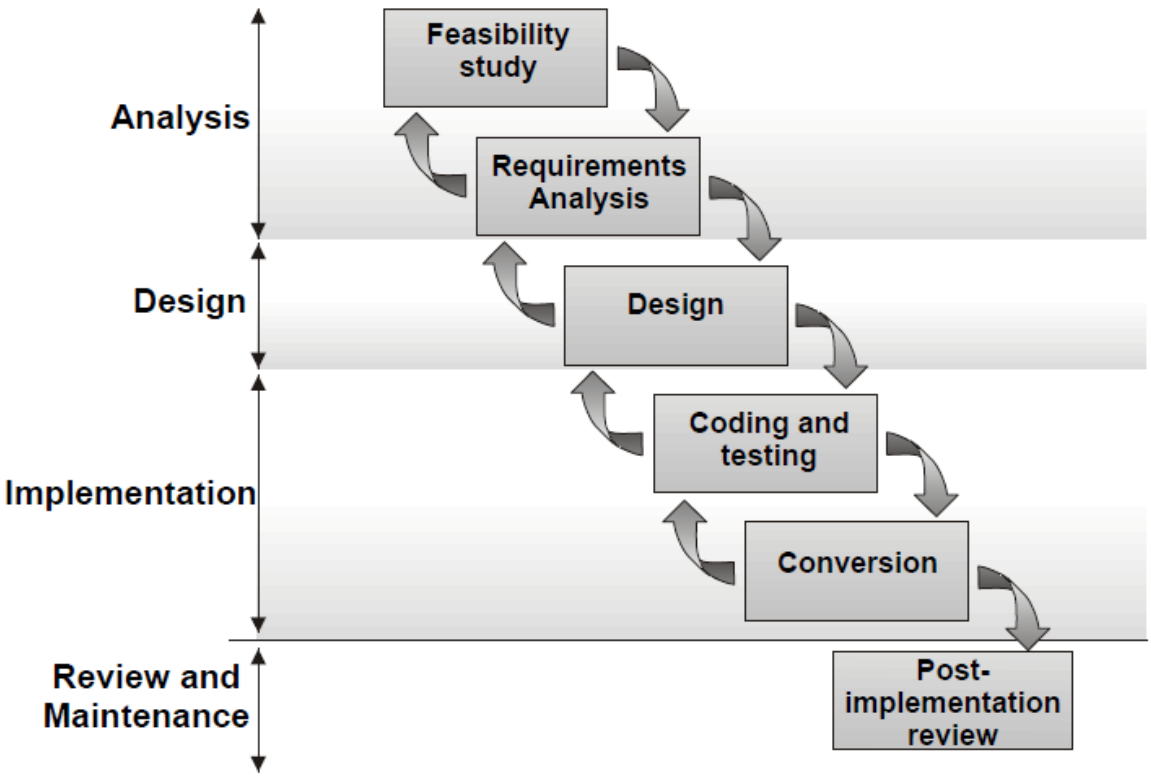


FIG: The SDLC(Waterfall Model)

2.6 Make or Buy Decision for IS Development

When you are supposed to make a make-or-buy decision for ID development, there are four numbers need to be aware of. The decision will be based on the values of these four numbers:

- The volume
- The fixed cost of making
- Per-unit direct cost when making
- Per-unit cost when buying

Now, there are two formulas that use the above numbers. They are 'Cost to Buy' and 'Cost to Make'. The higher value loses and the decision maker can go ahead with the less costly solution.

> Cost to Buy (CTB) = Volume x Per-unit cost when buying
>
> Cost to Make (CTM) = Fixed costs + (Per-unit direct cost x volume)

Reasons for Making

There are number of reasons a company would consider when it comes to making in-house. Following are a few:

- Cost concerns
- Desire to expand the manufacturing focus
- Need of direct control over the product
- Intellectual property concerns
- Quality control concerns
- Supplier unreliability
- Lack of competent suppliers
- Volume too small to get a supplier attracted
- Reduction of logistic costs (shipping etc.)
- To maintain a backup source
- Political and environment reasons
- Organizational pride

Reasons for Buying

Following are some of the reasons companies may consider when it comes to buying from a supplier:

- Lack of technical experience
- Supplier's expertise on the technical areas and the domain
- Cost considerations

- Need of small volume
- Insufficient capacity to produce in-house
- Brand preferences
- Strategic partnerships

2.7 Summary

A transaction process system (TPS) is an information processing system for business transactions involving the collection, modification and retrieval of all transaction data. Characteristics of a TPS include performance, reliability and consistency. TPS is also known as transaction processing or real-time processing.

Management information systems (**MIS**) is the study of people, technology, organizations, and the relationships among them. MIS professionals help firms realize maximum benefit from investment in personnel, equipment, and business processes. MIS is a people-oriented field with an emphasis on service through technology. Management information systems are typically computer systems used for data managing to make searching, analyzing data, and spring information easier. Management information systems are distinct from other information systems in that they are used to analyze and facilitate strategic and operational activities.

- *Management information systems*, produce fixed, regularly scheduled reports based on data extracted and summarized from the firm's underlying transaction processing systems to middle and operational level managers to identify and inform structured and semi-structured decision problems.
- *Decision support systems (DSS)* are computer program applications used by middle and higher management to compile information from a wide range of sources to support problem solving and decision making. A DSS is used mostly for semi-structured and unstructured decision problems.
- *Executive information systems (EIS)* is a reporting tool that provides quick access to summarized reports coming from all company levels and departments such as accounting, human resources and operations.
- *Marketing Information Systems* are Management Information Systems designed specifically for managing the marketing aspects of the business.
- *Office automation systems (OAS)* support communication and productivity in the enterprise by automating workflow and eliminating bottlenecks. OAS may be implemented at any and all levels of management.

- *School Information management systems* (SIMS) cover school administration, and often including teaching and learning materials.
- *Enterprise resource planning* facilitates the flow of information between all business functions inside the boundaries of the organization and manage the connections to outside stakeholders

The systems development life cycle (SDLC) is a conceptual model used in project management that describes the stages involved in an information system development project, from an initial feasibility study through maintenance of the completed application.

Various SDLC methodologies have been developed to guide the processes involved, including the waterfall model (which was the original SDLC method); rapid application development (RAD); joint application development (JAD); the fountain model; the spiral model; build and fix; and synchronize-and-stabilize. Frequently, several models are combined into some sort of hybrid methodology. Documentation is crucial regardless of the type of model chosen or devised for any application, and is usually done in parallel with the development process. Some methods work better for specific types of projects, but in the final analysis, the most important factor for the success of a project may be how closely the particular plan was followed.

In general, an SDLC methodology follows the following steps:

1. The existing system is evaluated. Deficiencies are identified. This can be done by interviewing users of the system and consulting with support personnel.
2. The new system requirements are defined. In particular, the deficiencies in the existing system must be addressed with specific proposals for improvement.
3. The proposed system is designed. Plans are laid out concerning the physical construction, hardware, operating systems, programming, communications, and security issues.
4. The new system is developed. The new components and programs must be obtained and installed. Users of the system must be trained in its use, and all aspects of performance must be tested. If necessary, adjustments must be made at this stage.
5. The system is put into use. This can be done in various ways. The new system can phased in, according to application or location, and the old system gradually replaced. In some cases, it may be more cost-effective to shut down the old system and implement the new system all at once.

6. Once the new system is up and running for a while, it should be exhaustively evaluated. Maintenance must be kept up rigorously at all times. Users of the system should be kept up-to-date concerning the latest modifications and procedures

Advantages of Spiral Model

1. Avoidance of Risk is enhanced.
2. Strong approval and documentation control.
3. Implementation has priority over functionality.
4. Additional Functionality can be added at a later date.

Disadvantages of Spiral Model

1. Highly customized limiting re-usability
2. Applied differently for each application
3. Risk of not meeting budget or schedule
4. Possibility to end up implemented as the Waterfall framework

Advantages of Prototype model

1. Strong Dialogue between users and developers
2. Missing functionality can be identified easily
3. Confusing or difficult functions can be identified
4. Requirements validation, Quick implementation of, incomplete, but functional, application
5. May generate specifications for a production application
6. Environment to resolve unclear objectives
7. Encourages innovation and flexible designs

Disadvantages of Prototype model

1. Contract may be awarded without rigorous evaluation of Prototype
2. Identifying non-functional elements difficult to document
3. Incomplete application may cause application not to be used as the full system was designed
4. Incomplete or inadequate problem analysis
5. Client may be unknowledgeable
6. Approval process and requirement is not strict
7. Requirements may frequently change significantly

2.8 Glossary

Formal Information System

Informal Information System

Public and Private Information System

Computer Based Information System (CBIS)

Transaction Processing System (TPS)

Management Information System(MIS)

Decision Support System (DSS)

Office Automation System (OAS)

System Development life cycle (SDLC)

Systems Planning

Systems Analysis

Systems Design

Systems Implementation

Systems Operation and Support

Prototyping approach

2.9 References

O'Brien, J (1999). *Management Information Systems – Managing Information Technology in the Internetworked Enterprise*. Boston: Irwin McGraw-Hill. ISBN 0-07-112373-3.

Laudon, Kenneth C.; Laudon, Jane P. (2009). *Management Information Systems: Managing the Digital Firm* (11 ed.). Prentice Hall/CourseSmart. p. 164.

Transaction processing systems (TPS) collect and record the routine transactions of an organization. Examples of such systems are sales order entry, hotel reservations, payroll, employee record keeping, and shipping.

Bidgoli, Hossein, (2004). The Internet Encyclopedia, Volume 1, John Wiley & Sons, Inc. p. 707.

Pant, S., Hsu, C., (1995), Strategic Information Systems Planning: A Review, Information Resources Management Association International Conference, May 21–24, Atlanta.

2.10Further Readings

1. Taylor, Victoria. "Supply Chain Management: The Next Big Thing?". *Sept. 12, 2011*. Business Week. Retrieved 5 March 2014.
2. Lynn, Samara. "What is CRM?". PC Mag. Retrieved 5 March 2014.
3. Joshi, Girdhar (2013). *Management Information Systems*. New Delhi: Oxford University Press. p. 328. ISBN 9780198080992.
4. Laudon, K.,&Laudon, J. (2010). Management information systems: Managing the digital firm. (11th ed.). Upper Saddle River, NJ: Pearson Prentice Hall

2.11 Model Questions

1. Explain in detail System Development life cycle (SDLC).
2. What do you understand by Systems Implementation and Systems Operation and Support?
3. Explain in detail Prototyping approach?

ANSWERS TO SELF ASSESSMENT QUESTIONS

A. 1
B. 5
C. 1
D. 3
E. 3

UNIT –II

CHAPTER 3

Structure

3.0Objectives

After studying this chapter you will be able to:

1. **Understand Database management**
2. **Understand different data models**
3. **Understand the advantage of Database Approach**
4. **Understand the concepts of E-R Modeling**
5. **Answer about the Normalization**

3.1Information Technology for IS

Information technology (**IT**) is the application of computers and telecommunications equipment to store, retrieve, transmit and manipulate data, often in the context of a business or other enterprise.

The term is commonly used as a synonym for computers and computer networks, but it also encompasses other information distribution technologies such as television and telephones. Several industries are associated with information technology, including computer hardware, software, electronics, semiconductors, internet, telecom equipment, e-commerce and computer services.

It is often observed that term information system and information technology are used interchangeably. In a literal sense, information technology is a subset of information systems. Information systems consist of people, processes, machines and information technology. The great advancement in information systems is due to development in information technology and introduction of computers.

Information System

An information system can be defined as set of coordinated network of components, which act together towards producing, distributing and or processing information. An important characteristic of computer-based information systems information is precision, which may not apply to other types.

In any given organization information system can be classified based on the usage of the information. Therefore, information systems in business can be divided into operations support system and management support system.

Information Technology

Everyday knowingly or unknowingly, everyone is utilizing information technology. It has grown rapidly and covers many areas of our day to day life like movies, mobile phones, the internet, etc.

Information technology can be broadly defined as integration of computer with telecommunication equipment for storing, retrieving, manipulating and storage of data. According to Information Technology Association of America, information technology is defined as “the study, design, development, application, implementation, support or management of computer-based information systems.”

Information technology greatly enhances the performance of economy; it provides edge in solving social issues as well as making information system affordable and user friendly.

Information technology has brought big change in our daily life be it education, life at home, work place, communication and even in function of government.

Comparison of Information System and Information Technology

Information system and information technology are similar in many ways but at the same time they are different. Following are some aspects about information system as well as information technology.

- **Origin:** Information systems have been in existence since pre-mechanical era in form of books, drawings, etc. However, the origin of information technology is mostly associated with invention of computers.
- **Development:** Information systems have undergone great deal of evolution, i.e. from manual record keeping to the current cloud storage system. Similarly, information technology is seeing constant changes with evermore faster processor and constantly shrinking size of storage devices.
- **Business Application:** Businesses have been using information systems for example in form of manual books of accounts to modern TALLY. The mode of communication has also gone under big change, for example, from a letter to email. Information technology has helped drive efficiency across organization with improved productivity and precision manufacturing.

3.2 Database Management: Basic concepts

Database: an organized collection of data. Outside the world of professional information technology, the term *database* is often used to refer to any collection of related data (such as a spreadsheet or a card index). A Database is a collection of related data organized in a way that data can be easily accessed, managed and updated. Any piece of information can be a data, for example name of your school. Database is actually a place where related piece of information is stored and various operations can be performed on it.

Database management system (DBMS): group of programs to manage database,

- Manipulates database
- Provides an interface between database and the user of the database and other application programs

A **DBMS** is a software that allows creation, definition and manipulation of database. DBMS is actually a tool used to perform any kind of operation on data in database. DBMS also provides protection and security to database. It maintains data consistency in case of multiple users. Here are some examples of popular dbms, MySql, Oracle, Sybase, Microsoft Access and IBM DB2 etc. The DBMS provides users and programmers with a systematic way to create, retrieve, update and manage data. A DBMS makes it possible for end users to create, read, update and delete data in a database. The DBMS essentially serves as an interface between the database and end users or application programs, ensuring that data is consistently organized and remains easily accessible.

Database management systems (**DBMS**s) are computer software applications that interact with the user, other applications, and the database itself to capture and analyze data. A general-purpose DBMS is designed to allow the definition, creation, querying, update, and administration of databases. Well-known DBMSs include MySQL, PostgreSQL, Microsoft SQL Server, Oracle, SAP and IBM DB2. A database is not generally portable across different DBMSs, but different DBMSs can interoperate by using standards such as SQL and ODBC or JDBC to allow a single application to work with more than one DBMS. Database management systems are often classified according to the database model that they support; the most popular database systems since the 1980s have all supported the relational model as represented by the SQL language. Sometimes a DBMS is loosely referred to as a 'database'.

A "database" refers to a set of related data and the way it is structured or 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 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 provide ways to manage how that information is organized.

The DBMS manages three important things: the data, the database engine that allows data to be accessed, locked and modified -- and the database schema, which defines the database's logical structure. These three foundational elements help provide concurrency, security, data integrity and uniform administration procedures. Typical database administration tasks supported by the DBMS include change management, performance monitoring/tuning and backup and recovery. Many database management systems are also responsible for automated rollbacks, restarts and recovery as well as the logging and auditing of activity.

The DBMS is perhaps most useful for providing a centralized view of data that can be accessed by multiple users, from multiple locations, in a controlled manner. A DBMS can limit what data the end user sees, as well as how that end user can view the data, providing many views of a single database schema. End users and software programs are free from having to understand where the data is physically located or on what type of storage media it resides because the DBMS handles all requests.

The DBMS can offer both logical and physical data independence. That means it can protect users and applications from needing to know where data is stored or having to be concerned about changes to the physical structure of data (storage and hardware). As long as programs use the application programming interface (API) for the database that is provided by the DBMS, developers won't have to modify programs just because changes have been made to the database.

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.

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.
- **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.
- **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.

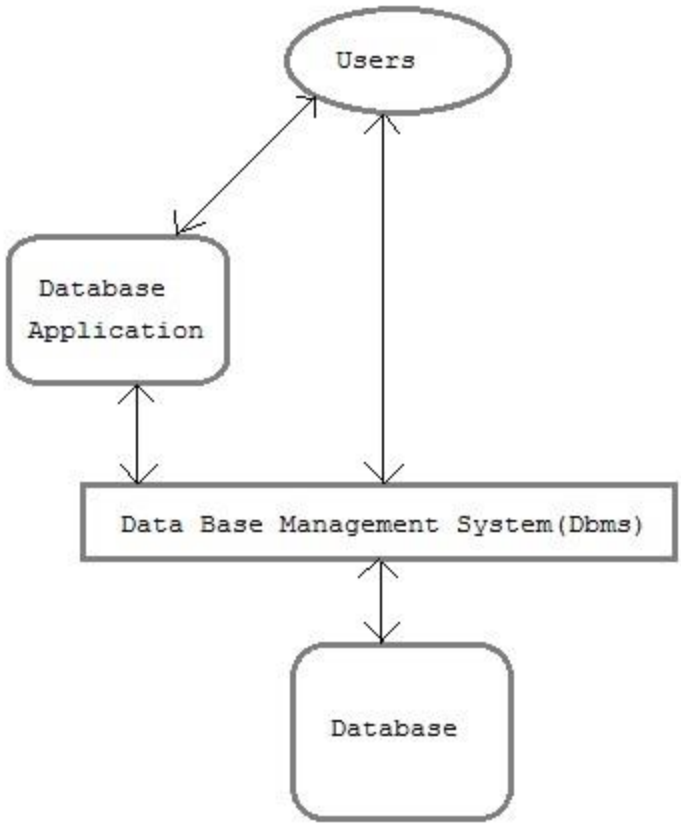
Both a database and its DBMS conform to the principles of a particular database model. "Database system" refers collectively to the database model, database management system, and database.

Physically, database servers 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 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. DBMSs may be built around a custom multitasking kernel with built-in networking support, but modern DBMSs typically rely on a standard operating system to provide these functions. Since DBMSs comprise a significant economical market computer and storage vendors often take into account DBMS requirements in their own development plans. 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(s) used to access the database (such as SQL or XQuery), and their internal engineering, which affects performance, scalability, resilience, and security.

Database administrator (DBA): skilled IS professional who directs all activities related to an organization’s database

DBMS Components

The database system can be divided into four components.



- **Users:** Users may be of various type such as DB administrator, System developer and End users.
- **Database application :** Database application may be Personal, Departmental, Enterprise and Internal
- **DBMS:** Software that allow users to define, create and manages database access, Ex: MySql, Oracle etc.
- **Database:** Collection of logical data.

ACTIVITIES

1. As an entrepreneur, you decide to open a fast food restaurant. You can purchase a franchise from one of the established corporations or create your own restaurant. Compare the choices by identifying the decisions you will face with each approach. What data will you need to collect?
2. Can value of information overload? How does it occur? And how would you control it?

Data Management:

Without data and the ability to process it, an organization could not successfully complete most business activities

 Data consists of raw facts

 To transform data into useful information, it must first be organized in a meaningful way

Database management involves the monitoring, administration, and maintenance of the databases and database groups in your enterprise.

Functions of DBMS

- Provides data Independence
- Concurrency Control
- Provides Recovery services
- Provides Utility services
- Provides a clear and logical view of the process that manipulates data

Advantages of DBMS

Using a DBMS to store and manage data comes with advantages, but also overhead. One of the biggest advantages of using a DBMS is that it lets end users and application programmers access and use the same data while managing data integrity. Data is better protected and maintained when it can be shared using a DBMS instead of creating new iterations of the same data stored in new files for every new application. The DBMS provides a central store of data that can be accessed by multiple users in a controlled manner.

Central storage and management of data within the DBMS provides:

- Data abstraction and independence
- Data security
- A locking mechanism for concurrent access
- An efficient handler to balance the needs of multiple applications using the same data
- The ability to swiftly recover from crashes and errors, including restartability and recoverability
- Robust data integrity capabilities
- Logging and auditing of activity
- Simple access using a standard application programming interface (API)
- Uniform administration procedures for data

Another advantage of a DBMS is that it can be used to impose a logical, structured organization on the data. A DBMS delivers economy of scale for processing large amounts of data because it is optimized for such operations.

A DBMS can also provide many views of a single database schema. A view defines what data the user sees and how that user sees the data. The DBMS provides a level of abstraction between the conceptual schema that defines the logical structure of the database and the physical schema that describes the files, indexes and other physical mechanisms used by the database. When a DBMS is used, systems can be modified much more easily when business requirements change. New categories of data can be added to the database

without disrupting the existing system and applications can be insulated from how data is structured and stored.

Of course, a DBMS must perform additional work to provide these advantages, thereby bringing with it the overhead. A DBMS will use more memory and CPU than a simple file storage system. And, of course, different types of DBMSes will require different types and levels of system resources.

The advantages can be listed as:

- Segregation of application program.
- Minimal data duplicacy
- Easy retrieval of data.
- Reduced development time and maintenance need.

Disadvantages of DBMS

- Complexity
- Costly
- Large in size

Database Approach

- **Traditional approach to data management:** separate data files are created and stored for each application program
- **Database approach to data management:** a pool of related data is shared by multiple application programs. Offers significant advantages over the traditional file-based approach

3.3Data Models

A **data model** organizes data elements and standardizes how the data elements relate to one another. Since data elements document real life people, places and things and the events between them, the data model represents reality, for example a house has many windows or a cat has two eyes. Computers are used for the accounting of these real life things and events and therefore the data model is a necessary standard to ensure exact communication between human beings.

Data models are often used as an aid to communication between the business people defining the requirements for a computer system and the technical people defining the design in response to those requirements. They are used to show the data needed and created by business processes.

Precise accounting and communication is a large expense and organizations traditionally paid the cost by having employees translate between themselves on an ad hoc basis. In critical situations such as air travel, healthcare and finance, it is becoming commonplace that the accounting and communication must be precise and therefore requires the use of common data models to obviate risk.

Hoberman defined "A data model is a way finding tool for both business and IT professionals, which uses a set of symbols and text to precisely explain a subset of real information to improve communication within the organization and thereby lead to a more flexible and stable application environment."

A data model explicitly determines the structure of data. Data models are specified in a data modeling notation, which is often graphical in form.

A data model can be sometimes referred to as a data structure, especially in the context of programming languages. Data models are often complemented by function models, especially in the context of enterprise models. A data model provides the details of information to be stored, and is of primary use when the final product is the generation of computer software code for an application or the preparation of a functional specification to aid a computer software make-or-buy decision.

The role of data models

The main aim of data models is to support the development of information systems by providing the definition and format of data. According to West and Fowler "if this is done consistently across systems then compatibility of data can be achieved. If the same data structures are used to store and access data then different applications can share data. The results of this are indicated above. However, systems and interfaces often cost more than they should, to build, operate, and maintain. They may also constrain the business rather than support it. A major cause is that the quality of the data models implemented in systems and interfaces is poor".

- "Business rules, specific to how things are done in a particular place, are often fixed in the structure of a data model. This means that small changes in the way business is conducted lead to large changes in computer systems and interfaces".
- "Entity types are often not identified, or incorrectly identified. This can lead to replication of data, data structure, and functionality, together with the attendant costs of that duplication in development and maintenance".
- "Data models for different systems are arbitrarily different. The result of this is that complex interfaces are required between systems that share data. These interfaces can account for between 25-70% of the cost of current systems".

- "Data cannot be shared electronically with customers and suppliers, because the structure and meaning of data has not been standardized. For example, engineering design data and drawings for process plant are still sometimes exchanged on paper".

The reason for these problems is a lack of standards that will ensure that data models will both meet business needs and be consistent. According to Hoberman "A data model is a way finding tool for both business and IT professionals, which uses a set of symbols and text to precisely explain a subset of real information to improve communication within the organization and thereby lead to a more flexible and stable application environment."

A data model explicitly determines the structure of data or structured data. Typical applications of data models include database models, design of information systems, and enabling exchange of data. Usually data models are specified in a data modeling language.

Communication and precision are the two key benefits that make a data model important to applications that use and exchange data. A data model is the medium which project team members from different backgrounds and with different levels of experience can communicate with one another. Precision means that the terms and rules on a data model can be interpreted only one way and are not ambiguous.

A data model can be sometimes referred to as a data structure, especially in the context of programming languages. Data models are often complemented by function models, especially in the context of enterprise models

3.4 Advantages of Database Approach

There are many advantages of database approach. The organizations today are focusing on the database approach for data handling. The major advantage database approach provides are:

· Increase data share ability
· Increase data integrity
· Increase the speed in implementing applications
· Ease data access by programmers and users
· Increase data independence
· Reduce program maintenance
· Provide a management view of the organization
· Improve the standards of the systems developers.

Increase data share ability: Large organizations, such as insurance companies, banks, local councils and manufacturing companies, have for some time been putting large amounts of data onto their computer systems. Frequently, the same data was being collected, validated, stored and accessed separately for a number of purposes. For example, there could be a file of customer details for sales order processing and another for sales ledger. This 'data redundancy' is costly and can be avoided by following a database approach. In fact some data duplication is reasonable in a database environment, but it should be known, controlled and be there for a purpose, such as efficient response to some database queries. However, the underlying data should be collected only once, and verified only once, so that there is little chance of inconsistency. With conventional files, the data is often collected at different times and validated by different validation routines, and therefore the output produced by different systems could well be inconsistent. In such situations the data resource is not easily managed and this leads to a number of problems. With reduced redundancy, data can be managed and shared, but it is essential that good integrity and security features operate in such systems. In other words, there needs to be control of the data resource. Furthermore, each application should run 'unaware' of the existence of others using the database. Good share ability implies a ready availability of the data to all users. The computer system must therefore be powerful enough so that performance is good even when there are a large number of users concurrently accessing the database.

Increase data integrity: In a shared environment, it is crucial for the success of the database system to control the creation, deletion and update of data and to ensure its correctness and its 'up-to-dateness' in general, ensure the quality of the data. Furthermore, with so many users accessing the database, there must be some control to prevent failed transactions leaving the database in an inconsistent state. However, this should be easier to effect in a database environment, because of the possibilities of central management of the data resource, than in an environment where each application sets up its own files. Standards need only be agreed and implemented once for all users.

Increase speed of implementing applications: Applications ought to be implemented in less time, since systems development staff can largely concentrate on the processes involved in the application itself rather than on the collection, validation, sorting and storage of data. Much of the data required for a new application may already be held on the database, put there for another purpose. Accessing the data will also be easier because the data manipulation features of the database management system will handle this.

Ease data access by programmers and users: Early database management systems used well-known programming languages such as Cobol to access the database. Cobol, for example, was extended to include new instructions, which were used when it was necessary to access data on the database. However we now have specific database query languages, such as SQL, and other software tools that ease the process of

applications development in a database environment. Once the database had been set up, applications development time should be greatly reduced.

Increase data independence: There are many aspects to data independence. It is the ability to change the format of the data, the medium on which the data is held or the data structures, without having to change the programs that use the data. Conversely, it also means that it is possible to change the logic of the programs without having to change the file definitions, so that programmer productivity is increased. It also means that there can be different user views of the data even though it is stored only once. This separation of the issues concerning processes from the issues concerning data is a key reason for opting for the database solution. It provides far greater flexibility when compared to conventional file-based applications.

Reduce program maintenance: Stored data will need to be changed frequently as the real world, that it represents, changes. New data types need to be added, formats changed or new access methods introduced. The data independence of a database environment, discussed above, circumvents the necessity of changing each program with each change to the data structure or type.

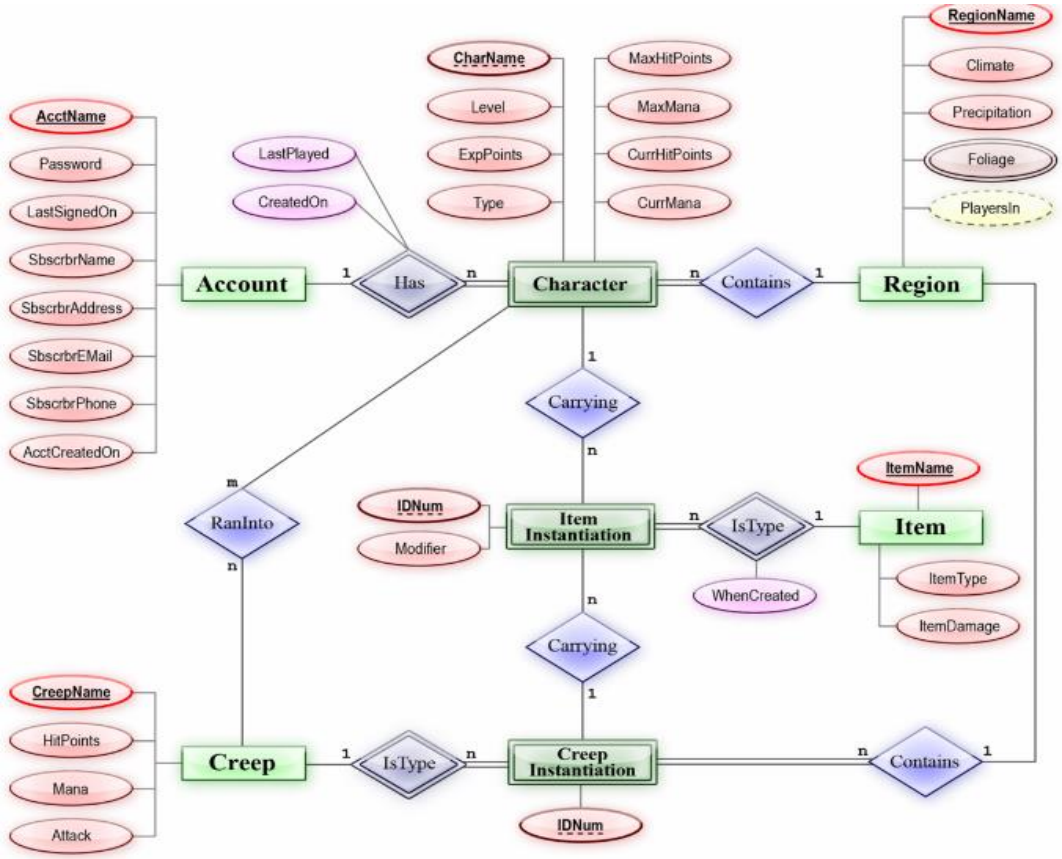
Improve standards: Applications tend to be implemented by different project teams of systems analysts and programmers and it has been difficult to apply standards and conventions for all applications. Computer people are reputed to dislike adopting the norms of the firm, and it is difficult to impose standards where applications are developed piecemeal. With a central database, it is possible to impose standards for file creation, access and update, and to impose good controls, enabling un-authorized access to be restricted and providing adequate back-up and security features.

Provide a management view: This is a very important aspect in the context of this book. Managers have frequently complained that they do not get the benefits from the expensive computing resource that they have sanctioned. However, managers have become aware of the need for a corporate view of their organization. Such a view requires data from a number of sections, departments, divisions and sometimes companies in a larger organization. This corporate view cannot be gained if files are established on an application-by application basis and not integrated as in a database. With decision support systems using the database, it becomes possible for problems previously considered solvable only by intuition and judgment to be solved with an added ingredient, that of information, which is timely, accurate and presented at the required level of detail. Some of this information could be provided on a regular basis whilst some will be of a ‘one-off’ nature. Database systems should respond to both types of request.

3.5 Overview of E-R Modeling

Entity–relationship modeling was developed by Peter Chen. An entity-relationship model (ERM) is an abstract conceptual data model (or semantic data model) used in software engineering to represent structured data. There are several notations used for ERMs. An **entity relationship model**, also called an **entity-relationship** (**ER**) **diagram**, is a graphical representation of entities and their relationships to each other, typically used in computing in regard to the organization of data within databases or information systems.

An **entity–relationship model** (**ER model**) is a data model for describing the data or information aspects of a business domain or its process requirements, in an abstract way that lends itself to ultimately being implemented in a database such as a relational database. The main components of ER models are entities (things) and the relationships that can exist among them, and databases.



An entity–relationship model is a systematic way of describing and defining a business process. The process is modeled as components (*entities*) that are linked with each other by *relationships* that express the dependencies and requirements between them, such as: *one building may be divided into zero or more apartments, but one apartment can only be located in one building.* Entities may have various properties (*attributes*) that characterize them. Diagrams created to represent these entities, attributes, and relationships graphically are called entity–relationship diagrams.

An ER model is typically implemented as a database. In the case of a relational database, which stores data in tables, every row of each table represents one instance of an entity. Some data fields in these tables point to indexes in other tables; such pointers represent the relationships.

The three schema approach to software engineering uses three levels of ER models that may be developed.

Conceptual data model

This is the highest level ER model in that it contains the least granular detail but establishes the overall scope of what is to be included within the model set. The conceptual ER model normally defines master reference data entities that are commonly used by the organization. Developing an enterprise-wide conceptual ER model is useful to support documenting the data architecture for an organization.

A conceptual ER model may be used as the foundation for one or more *logical data models* (see below). The purpose of the conceptual ER model is then to establish structural metadata commonality for the master data entities between the set of logical ER models. The conceptual data model may be used to form commonality relationships between ER models as a basis for data model integration.

Logical data model

A logical ER model does not require a conceptual ER model, especially if the scope of the logical ER model includes only the development of a distinct information system. The logical ER model contains more detail than the conceptual ER model. In addition to master data entities, operational and transactional data entities are now defined. The details of each data entity are developed and the entity relationships between these data entities are established. The logical ER model is however developed independent of technology into which it is implemented.

Physical data model

One or more physical ER models may be developed from each logical ER model. The physical ER model is normally developed to be instantiated as a database. Therefore, each physical ER model must contain enough detail to produce a database and each physical ER model is technology dependent since each database management system is somewhat different.

The physical model is normally instantiated in the structural metadata of a database management system as relational database objects such as database tables, database indexes such as unique key indexes, and database constraints such as a foreign key constraint or a commonality constraint. The ER model is also normally used to design modifications to the relational database objects and to maintain the structural metadata of the database.

The first stage of information system design uses these models during the requirements analysis to describe information needs or the type of information that is to be stored in a database. The data modeling technique can be used to describe any ontology (i.e. an overview and classifications of used terms and their relationships) for a certain area of interest. In the case of the design of an information system that is based on a database, the conceptual data model is, at a later stage (usually called logical design), mapped to a logical data model, such as the relational model; this in turn is mapped to a physical model during physical design. It is also used in database management system.

Entity – Relationship Diagrams:

The object-relationship pair can be represented graphically using an ER diagram. An entity represents an object.

Examples: a computer, an employee, a song, a mathematical theorem. Entities are represented as rectangles. A relationship captures how two or more entities are related to one another. Examples: an *owns* relationship between a company and a computer, a *supervises* relationship between an employee and a department, a *performs* relationship between an artist and a song. Relationships are represented as diamonds, connected by lines to each of the entities in the relationship. Entities and relationships can both have attributes.

Examples: an employee entity might have an employee ID number attribute; the *proved* relationship may have a *date* attribute. Attributes are represented as ellipses connected to their entity by a line.

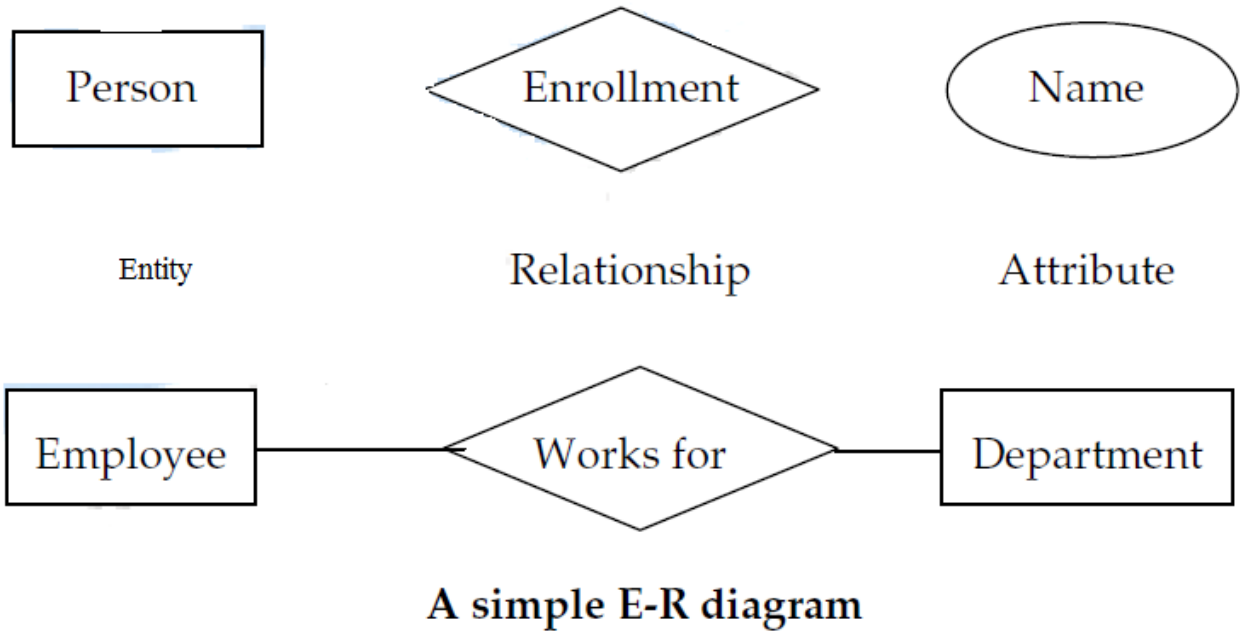


Fig: Simple ER Diagram

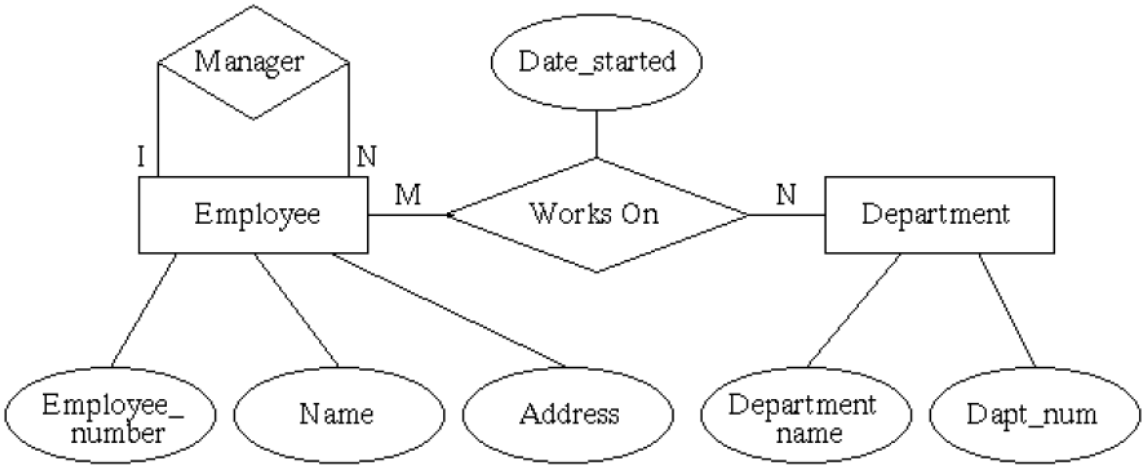


Fig: An E- R Diagram with attributes

3.6 Entity

An entity may be defined as a thing capable of an independent existence that can be uniquely identified. An entity is an abstraction from the complexities of a domain. When we speak of an entity, we normally speak of some aspect of the real world that can be distinguished from other aspects of the real world. An entity can be a real-world object, either animate or inanimate, that can be easily identifiable. For example, in a school database, students, teachers, classes, and courses offered can be considered as entities. All these entities have some attributes or properties that give them their identity.

An entity set is a collection of similar types of entities. An entity set may contain entities with attribute sharing similar values. For example, a Students set may contain all the students of a school; likewise a Teachers set may contain all the teachers of a school from all faculties. Entity sets need not be disjoint.

An entity is a thing that exists either physically or logically. An entity may be a physical object such as a house or a car(they exist physically), an event such as a house sale or a car service, or a concept such as a customer transaction or order(they exist logically—as a concept). Although the term entity is the one most commonly used, following Chen we should really distinguish between an entity and an entity-type. An entity-type is a category. An entity, strictly speaking, is an instance of a given entity-type. There are usually many instances of an entity-type. Because the term entity-type is somewhat cumbersome, most people tend to use the term entity as a synonym for this term. Entities can be thought of as nouns. Examples: a computer, an employee, a song, a mathematical theorem.

3.7 Attributes

An **attribute** is a specification that defines a property of an object, element, or file. It may also refer to or set the specific value for a given instance of such. For clarity, attributes should more correctly be considered

metadata. An attribute is frequently and generally a property of a property. However, in actual usage, the term attribute can and is often treated as equivalent to a property depending on the technology being discussed. An attribute of an object usually consists of a name and a value; of an element, a type or class name; of a file, a name and extension.

- Each named attribute has an associated set of rules called operations: one doesn't sum characters or manipulate and process an integer array as an image object— one doesn't process text as type floating point (decimal numbers).
- It follows that an object definition can be extended by imposing data typing: a representation format, a default value, and legal operations (rules) and restrictions ("Division by zero is not to be tolerated!") are all potentially involved in defining an attribute, or conversely, may be spoken of as attributes of that object's type. A JPEG file is not decoded by the same operations (however similar they may be— these are all graphics data formats) as a PNG or BMP file, nor is a floating point typed number operated upon by the rules applied to typed long integers.

For example, in computer graphics, line objects can have attributes such as thickness (with real values), color (with descriptive values such as brown or green or values defined in a certain color model, such as RGB), dashing attributes, etc. A circle object can be defined in similar attributes plus an origin and radius.

Types of Attributes

- **Simple attribute** − Simple attributes are atomic values, which cannot be divided further. For example, a student's phone number is an atomic value of 10 digits.
- **Composite attribute** − Composite attributes are made of more than one simple attribute. For example, a student's complete name may have first_name and last_name.
- **Derived attribute** − Derived attributes are the attributes that do not exist in the physical database, but their values are derived from other attributes present in the database. For example, average_salary in a department should not be saved directly in the database, instead it can be derived. For another example, age can be derived from data_of_birth.
- **Single-value attribute** − Single-value attributes contain single value. For example − Social_Security_Number.
- **Multi-value attribute** − Multi-value attributes may contain more than one values. For example, a person can have more than one phone number, email_address, etc.

These attribute types can come together in a way like −

- simple single-valued attributes

- simple multi-valued attributes
- composite single-valued attributes
- composite multi-valued attributes

Entity-Set and Keys

Key is an attribute or collection of attributes that uniquely identifies an entity among entity set. For example, the roll_number of a student makes him/her identifiable among students.

- **Super Key** − A set of attributes (one or more) that collectively identifies an entity in an entity set.
- **Candidate Key** − A minimal super key is called a candidate key. An entity set may have more than one candidate key.
- **Primary Key** − A primary key is one of the candidate keys chosen by the database designer to uniquely identify the entity set.

3.8 Relationships: Key Concepts

Characteristics of a suitable set of relations include:

–the *minimal* number of attributes necessary to support the data requirements of the enterprise;

–attributes with a close logical relationship are found in the same relation;

–*minimal* redundancy with each attribute represented only once with the important exception of attributes that form all or part of foreign keys

The association among entities is called a relationship. For example, an employee **works_at** a department, a student **enrolls** in a course. Here, Works_at and Enrolls are called relationships.

Relationship Set

A set of relationships of similar type is called a relationship set. Like entities, a relationship too can have attributes. These attributes are called **descriptive attributes**.

Degree of Relationship

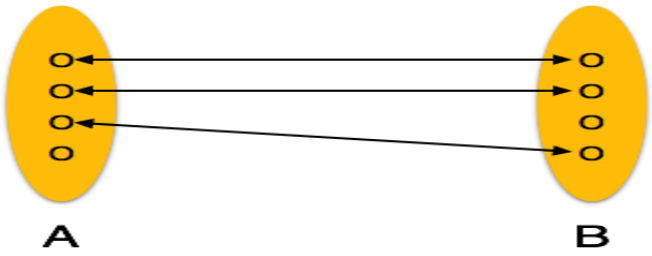
The number of participating entities in a relationship defines the degree of the relationship.

- Binary = degree 2
- Ternary = degree 3
- n-ary = degree

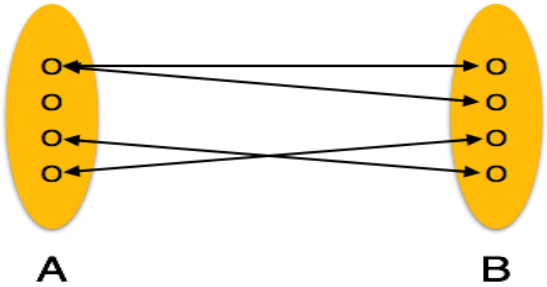
Mapping Cardinalities

Cardinality defines the number of entities in one entity set, which can be associated with the number of entities of other set via relationship set.

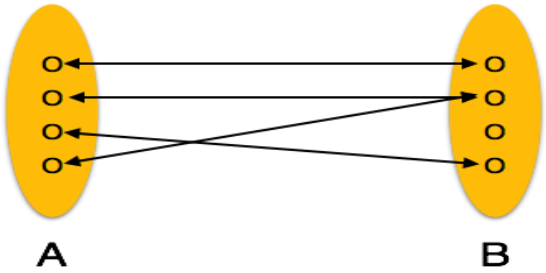
- **One-to-one** − One entity from entity set A can be associated with at most one entity of entity set B and vice versa.



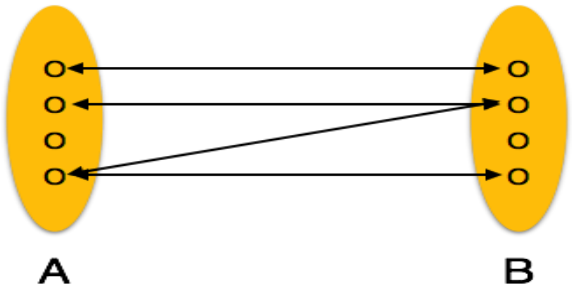
- **One-to-many** − One entity from entity set A can be associated with more than one entities of entity set B however an entity from entity set B, can be associated with at most one entity.



- **Many-to-one** − More than one entities from entity set A can be associated with at most one entity of entity set B, however an entity from entity set B can be associated with more than one entity from entity set A.



- **Many-to-many** − One entity from A can be associated with more than one entity from B and vice versa.



LIMITATIONS OF ER Modeling

- ER assume information content that can readily be represented in a relational database. They describe only a relational structure for this information.
- They are inadequate for systems in which the information cannot readily be represented in relational form, such as with semi-structured data.
- For many systems, possible changes to information contained are nontrivial and important enough to warrant explicit specification.
- Even where it is suitable in principle, ER modeling is rarely used as a separate activity. One reason for this is today's abundance of tools to support diagramming and other design support directly on relational database management systems. These tools can readily extract database diagrams that are very close to ER diagrams from existing databases, and they provide alternative views on the information contained in such diagrams.
- For modeling temporal databases, numerous ER extensions have been considered. Similarly, the ER model was found unsuitable for multidimensional databases (used in OLAP applications); no dominant conceptual model has emerged in this field yet, although they generally revolve around the concept of OLAP cube (also known as *data cube* within the field).

3.9 Normalization

Normalization is a technique for producing a set of suitable relations that support the data requirements of an enterprise. Normalization is the process of organizing the fields and tables of a relational database to minimize redundancy. Normalization usually involves dividing large tables into smaller (and less redundant) tables and defining relationships between them. The objective is to isolate data so that additions, deletions, and modifications of a field can be made in just one table and then propagated through the rest of the database using the defined relationships.

Normalization is used for mainly two purpose,

- Eliminating redundant (useless) data.
- Ensuring data dependencies make sense i.e. data is logically stored.

Edgar F. Codd, the inventor of the relational model, introduced the concept of normalization and what we now know as the First Normal Form (1NF) in 1970. Codd went on to define the Second Normal Form (2NF) and Third Normal Form (3NF) in 1971, and Codd and Raymond F. Boyce defined the Boyce-Codd Normal Form (BCNF) in 1974. Informally, a relational database table is often described as "normalized" if it is in the Third Normal Form. Most 3NF tables are free of insertion, update, and deletion anomalies.

A standard piece of database design guidance is that the designer should first create a fully normalized design; then selective de-normalization can be performed for performance reasons.

A typical example of normalization is that a unique ID is stored everywhere in the system but its name is held in only one table. The name can be updated more easily in one row of one table. A typical update in such an example would be the RIM company changing its name to BlackBerry. That update would be done in one place and immediately the correct "BlackBerry" name would be displayed throughout the system.

3.10 Basic Normal Forms

The **normal forms** (abbrev. **NF**) of relational database theory provide criteria for determining a table's degree of immunity against logical inconsistencies and anomalies. The higher the normal form applicable to a table, the less vulnerable it is. Each table has a "**highest normal form**" (**HNF**): by definition, a table always meets the requirements of its HNF and of all normal forms lower than its HNF; also by definition, a table fails to meet the requirements of any normal form higher than its HNF.

The normal forms are applicable to individual tables; to say that an entire database is in normal form *n* is to say that all of its tables are in normal form *n*.

Newcomers to database design sometimes suppose that a normalization proceeds in an iterative fashion, i.e. a 1NF design is first normalized to 2NF, then to 3NF, and so on. This is not an accurate description of how normalization typically works. A sensibly designed table is likely to be in 3NF on the first attempt; furthermore, if it is 3NF, it is overwhelmingly likely to have an HNF of 5NF. Achieving the "higher" normal forms (above 3NF) does not usually require an extra expenditure of effort on the part of the designer, because 3NF tables usually need no modification to meet the requirements of these higher normal forms.

	Normal form	Defined by	Year	Definition
1NF	First normal form	E.F. Codd (1970), C.J. Date (2003)	1970 & 2003	The domain of each attribute contains only atomic values, and the value of each attribute contains only a single value from that domain.

2NF	Second normal form	E.F. Codd	1971	No non-prime attribute in the table is functionally dependent on a proper subset of any candidate key
3NF	Third normal form	E.F. Codd(1971), C. Zaniolo (1982)	1971 & 1982	Every non-prime attribute is non-transitively dependent on every candidate key in the table. The attributes that do not contribute to the description of the primary key are removed from the table. In other words, no transitive dependency is allowed.
EKNF	Elementary Key Normal Form	C. Zaniolo	1982	Every non-trivial functional dependency in the table is either the dependency of an elementary key attribute or a dependency on a super key
BCNF	Boyce–Codd normal form	Raymond F. Boyce and E.F. Codd	1974	Every non-trivial functional dependency in the table is a dependency on a super key
4NF	Fourth normal form	Ronald Fagin	1977	Every non-trivial multi valued dependency in the table is a dependency on a super key
5NF	Fifth normal form	Ronald Fagin	1979	Every non-trivial join dependency in the table is implied by the super keys of the table
DKNF	Domain/key normal form	Ronald Fagin	1981	Every constraint on the table is a logical consequence of the table's domain constraints and key constraints
6NF	Sixth normal form	C.J. Date, Hugh Darwen, & Nikos Lorentzos	2002	Table features no non-trivial join dependencies at all (with reference to generalized join operator)

SELF ASSESSMENT QUESTIONS

Please mark the correct option

A. When a bank uses business performance management software to monitor its performance in differences regions this:
   1. reduces costs.
   2. manages risks.
   3. adds value.
   4. creates a new opportunity.
   5. none of the above.

B. When a bank offers web self-service for customers to answer their questions, the primary outcome is:
   1. adds value.
   2. manages risks.
   3. reduces costs.
   4. creates a new opportunity.
   5. none of the above.

C. The general transformation cycle for information is:
   1. information to data to knowledge.
   2. knowledge to data to information.
   3. data to knowledge to information.
   4. data to information to knowledge.
   5. none of the above.

D. The most important attribute of information quality that a manager requires is:
   1. relevance.
   2. media.
   3. presentation.
   4. timeliness.
   5. none of the above.

3.11 overview of SQL

SQL (**Structured Query Language**) is a special-purpose programming language designed for managing data held in a relational database management system (RDBMS), or for stream processing in a relational data stream management system (RDSMS).

Originally based upon relational algebra and tuple relational calculus, SQL consists of a data definition language and a data manipulation language. The scope of SQL includes data insert, query, update and delete, schema creation and modification, and data access control. Although SQL is often described as, and to a great extent is, a declarative language (4GL), it also includes procedural elements.

SQL was one of the first commercial languages for Edgar F. Codd's relational model, as described in his influential 1970 paper, "A Relational Model of Data for Large Shared Data Banks." Despite not entirely adhering to the relational model as described by Codd, it became the most widely used database language.

SQL became a standard of the American National Standards Institute (ANSI) in 1986, and of the International Organization for Standardization (ISO) in 1987. Since then, the standard has been revised to include a larger set of features. Despite the existence of such standards, though, most SQL code is not completely portable among different database systems without adjustments.

The SQL language is subdivided into several language elements, including:

- *Clauses*, which are constituent components of statements and queries. (In some cases, these are optional.)
- *Expressions*, which can produce either scalar values, or tables consisting of columns and rows of data
- *Predicates*, which specify conditions that can be evaluated to SQL three-valued logic (3VL) (true/false/unknown) or Boolean truth values and are used to limit the effects of statements and queries, or to change program flow.
- *Queries*, which retrieve the data based on specific criteria. This is an important element of *SQL*.
- *Statements*, which may have a persistent effect on schemata and data, or may control transactions, program flow, connections, sessions, or diagnostics.
    - SQL statements also include the semicolon (";") statement terminator. Though not required on every platform, it is defined as a standard part of the SQL grammar.
- *Insignificant whitespace* is generally ignored in SQL statements and queries, making it easier to format SQL code for readability

3.12 Summary

An information system can be defined as set of coordinated network of components, which act together towards producing, distributing and or processing information. An important characteristic of computer-based information systems information is precision, which may not apply to other types.

Database management involves the monitoring, administration, and maintenance of the databases and database groups in your enterprise.

The major advantage database approach provides are:

· Increase data share ability

· Increase data integrity

· Increase the speed in implementing applications

· Ease data access by programmers and users

· Increase data independence

· Reduce program maintenance

· Provide a management view of the organization

· Improve the standards of the systems developers

Entity–relationship modeling was developed by Peter Chen. An entity-relationship model (ERM) is an abstract conceptual data model (or semantic data model) used in software engineering to represent structured data

The object-relationship pair can be represented graphically using an ER diagram. An entity represents an object.

Normalization is a technique for producing a set of suitable relations that support the data requirements of an enterprise. Normalization is the process of organizing the fields and tables of a relational database to minimize redundancy. Normalization usually involves dividing large tables into smaller (and less redundant) tables and defining relationships between them. The objective is to isolate data so that additions, deletions, and modifications of a field can be made in just one table and then propagated through the rest of the database using the defined relationships.

3.14 Glossary

Database: an organized collection of data. Outside the world of professional information technology, the term *database* is often used to refer to any collection of related data (such as a spreadsheet or a card index).

Database management system (DBMS): group of programs to manage database,

- Manipulates database
- Provides an interface between database and the user of the database and other application programs

Database management systems (**DBMS**s) are computer software applications that interact with the user, other applications, and the database itself to capture and analyze data. A general-purpose DBMS is designed to allow the definition, creation, querying, update, and administration of databases. Well-known DBMSs include MySQL, PostgreSQL, Microsoft SQL Server, Oracle, SAP and IBM DB2

Entity: An entity may be defined as a thing capable of an independent existence that can be uniquely identified. An entity is an abstraction from the complexities of a domain.

3.15 References

O'Brien, J (1999). *Management Information Systems – Managing Information Technology in the Internetworked Enterprise*. Boston: Irwin McGraw-Hill. ISBN 0-07-112373-3.

Laudon, Kenneth C.; Laudon, Jane P. (2009). *Management Information Systems: Managing the Digital Firm* (11 ed.). Prentice Hall/CourseSmart. p. 164.

Transaction processing systems (TPS) collect and record the routine transactions of an organization. Examples of such systems are sales order entry, hotel reservations, payroll, employee record keeping, and shipping.

Bidgoli, Hossein, (2004). The Internet Encyclopedia, Volume 1, John Wiley & Sons, Inc. p. 707.

Pant, S., Hsu, C., (1995), Strategic Information Systems Planning: A Review, Information Resources Management Association International Conference, May 21–24, Atlanta.

3.16 Further Readings

Taylor, Victoria. "Supply Chain Management: The Next Big Thing?". *Sept. 12, 2011*. Business Week. Retrieved 5 March 2014.

Lynn, Samara. "What is CRM?". PC Mag. Retrieved 5 March 2014.

Joshi, Girdhar (2013). *Management Information Systems*. New Delhi: Oxford University Press. p. 328. ISBN 9780198080992.

Laudon, K.,&Laudon, J. (2010). Management information systems: Managing the digital firm. (11th ed.). Upper Saddle River, NJ: Pearson Prentice Hall

3.17 Model Questions

Q:1 Discuss about the role of Information Technology for IS

Q:2 Explain in detail database management

Q:3 Explain the advantages of Database Approach.

Q:4 discuss in detail the E-R modeling.

Q:5 What do you understand by normalization. Explain its various forms.

ANSWERS TO SELF ASSESSMENT QUESTIONS

A. 2
B. 3
C. 4
D. 1

CHAPTER 4

Developing Access Database

4.0 Objectives

After studying this chapter you will be able to:

1. **Develop access database**
2. **Define fields and its properties**
3. **Define primary key**
4. entering and viewing data from tables
5. Create, run and build queries

4.1 Developing Access Database

In Microsoft Access a database consists of one single file. The file contains all the tables of the database, the relationships (the crow's feet), *queries* (computed tables), *forms* (user windows), and many other things.

Steps to Create a blank database

1. On the File tab, click New, and then click Blank Database.
2. Type a file name in the File Name box. ...
3. Click Create. ...
4. Begin typing to add data, or you can paste data from another source, as described in the section Copy data from another source into an Access table.

As a systems developer *design* tables and user windows. As a user enter data into the tables (usually through user windows) and get data out of the tables, for instance through the same windows or through printed reports.

In Access it is very easy to switch between the developer role and the user role. As a developer you will typically design some tables, then switch to the user role to enter data into them, then switch back to the developer role to change the design, design more tables, etc. Access can to a large extent restructure the data that already is in the database so that it matches the new table design.

1. Locate the Access program. Depending on the way the system is set up, you may find it under Pro-grams -> Microsoft Access or Programs -> Micro-soft Office -> Microsoft Access.
2. In Access **97** and **2000**: Open Access and ask for a "blank" database. In Access **2003**: Open Access and click the *New* icon (under the *File* menu). Then click *Blank database* in the help area to the far right.
3. Access now asks where to store the new database. Select the folder you want and give the database the name **hotel** (or **hotel.mdb**).

The screen now shows the *database window*. It should look like Figure 4.1A. (In Access 97 it looks slightly different). We have selected the *Tables* tab, but there are no tables or other things in the database as yet. However, you see three icons that can create tables for you. When you have created a table, it will appear in the table window and you can then *Open* it and enter data into it, or you can *Design* it, i.e. change the definition of it. (In Access 97 the database window looks like a traditional tab form. There are no create-icons, but function buttons for the same purpose.)

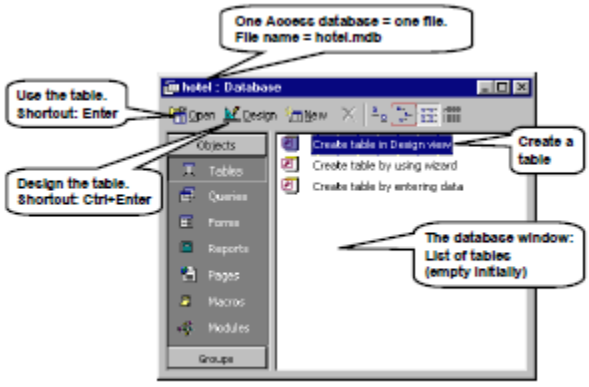


Figure: 4.1 A The Access database window

Double click on *Create table in Design view*.

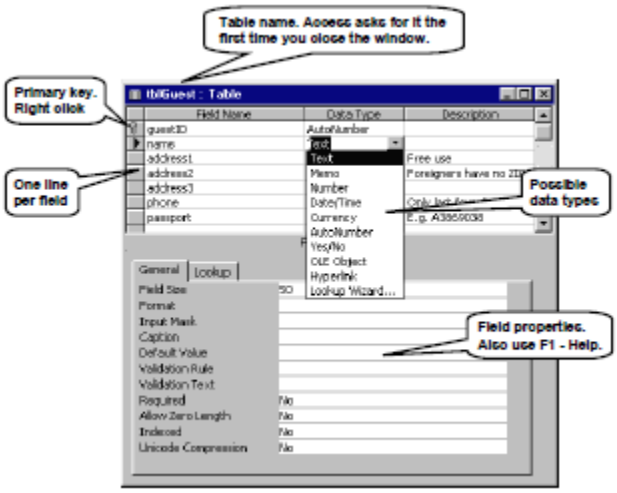


Figure: 4.1 B Define a table (design view)

Now you get a window as shown on Figure 4.1B. Here you define the fields (attributes) of the table. The list of fields runs downwards with one line per field. Initially there are only empty lines. The table hasn't got a name yet. Access asks for the name when you close the window.

The figure shows the finished guest table. You see the field names to the left. In the middle column is the type of the field - Data Type. The figure shows all the possible types as a combo box. The most important data types are Text, Number, Date/Time, and AutoNumber. An **AutoNumber** is a counter that Access increases for each new record, so that it serves as a unique key. The value is a *Long Integer* (32-bit integer). We explain more about data types in the next section.

1. Fill in all the field lines according to the attributes in the guest table (see the figure). All the fields are of type Text, except the guest ID which is of type AutoNumber.

Note that although we say phone *number* and passport *number*, these fields are texts because the "numbers" contain parentheses, dashes and maybe letters.

When you have chosen a data type, you can choose a number of other field properties. They are in the lower part of the window. On the figure you can see that the *name* field is a text field with space for 50 characters. You can also see that the user doesn't have to enter anything in the name field (*Required*=No). You should change this to *Yes* since it doesn't make sense to have a guest without a name.

Try to use Access's help to find more information about the data types and their properties. For instance, put the cursor in the Data Type of a field and click F1. Or point at one of the properties and click F1.

Lookup Wizard is not a field type. If you select Lookup Wizard, it makes the field into a combo box where the user can *select* a value instead of typing it into the field.

Key fields

Often you have to define a key field so that other tables can refer to this one. In our case, guestID must be the key field: Right-click somewhere in the guestID line. Then select *Primary Key*. Access now shows that the field is the key you can remove the key property again by once more selecting *Primary Key*. If the key consists of more than one field, you first select all the fields by clicking on their left-hand marker with Ctrl down. Then select Primary Key by right-clicking *inside* one of the field lines.

Close the window. Access asks you for the name of the table. Call it **tblGuest**. (The prefix *tbl* will help you remember that it is a table. As the system grows, there will be guest windows, guest buttons and many other things. Without discipline on your part, it becomes a mess.)

If you have not defined a primary key, Access will warn you and suggest that it makes one for you. Don't let it - do it yourself. Or at least check what Access makes in its excessive helpfulness.

Enter data

After these efforts, it is time to record some guests. Fortunately it is easy:

Select the guest table in the database window. Click *Open* or just use *Enter*.

Now the system shows the table in **user mode** (*Da-tasheet view*) so that you can enter guest data.

Enter the guests shown on Figure 4.1C. You add a new guest in the empty line of the table - the one marked with a star. Notice that as soon you start entering something, the record indicator changes to a pencil and a new star line appears. The pencil shows that you are editing the record, and the record you see is not yet in the database.

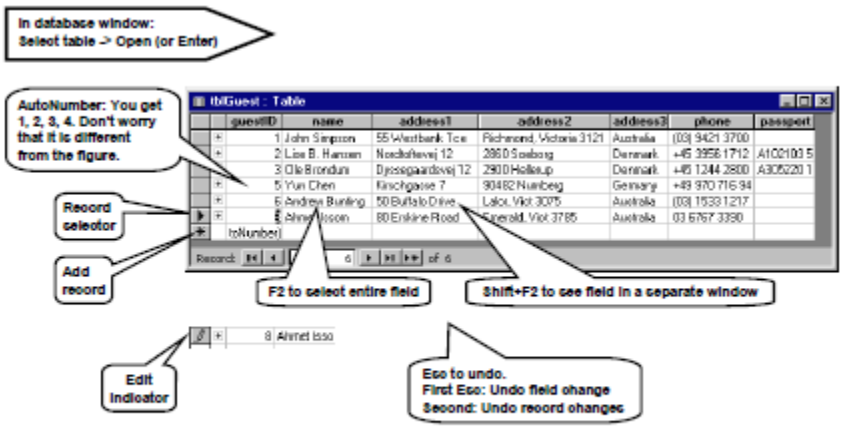


Figure: 4.1 C Enter data in user mode (datasheet view)

On Figure 4.1C we originally entered a guest that got guestID 4, later deleted this guest. Access will never reuse number 4 for a guest.

Close and reopen the database

To feel confident with Access, it is a good idea to close and open the database now.

1. Close the large Access window. (Not the small database window inside the Access window.)

Notice that Access doesn't ask whether you want to save changes. Access saves them all along, for instance when you define a table or when you enter a record in the table.

Find your database file (*hotel.mdb*) in the file folders. Use Enter or double click to open it.

Access **2003** is very security concerned and asks you several questions when you open the file. The dialog may vary from one installation to another, but is something like this:

1. *The file may not be safe. Do you want to open it*? Your database is safe, so answer *Open*.
2. *Unsafe expressions are not blocked. Do you want to block them?* You want full freedom, so answer *No.*
3. Access warns you one more time whether you want to open. Say *Open* or *Yes*. (In some versions the question is a very long text box, and you cannot understand it. Say yes anyway.)

As an alternative, you may say yes to blocking the unsafe expressions. This will save you some questions when you open the file in the future. However, some installations don't allow you to block expressions.

Note that Access 2003 shows that your database is in Access 2000 format. This is all right. It allows you to use it also from Access 2000. You can convert it to other formats with Tools -> Database Utilities -> Convert Database.

4.2 Defining fields for a table

The data model on Figure 4.2 shows the tables we will use. To simplify your job, we have shown all the keys, including the foreign keys and the artificial keys.

1. Close the guest table.
2. Create all the remaining tables in the same way as you created the guest table (from the *Tables* tab use *Create table in Design view* - or click *New*).

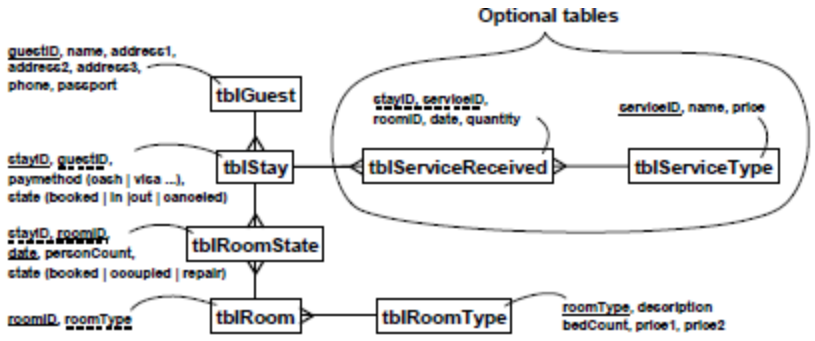


Figure: 4.2 Create Table

Make sure you define all the fields. Otherwise you will get stuck when later constructing the user interface. Here are a few notes about the various tables:

tblStay:

stayID is the primary key of tblStay. Make it an Auto-Number.

guestID is a foreign key that refers to the AutoNumber in tblGuest. The foreign key must have a matching data type - a long integer. Choose Data Type = *Number* and Field Size = *Long Integer*. **Warning:** Don't make the foreign key an AutoNumber. This would cause Access to fill in the foreign key fields automatically, and you cannot change the numbers so that they point to the primary keys in the guest table.

paymethod is an enumeration type. Make it an integer (a 16-bit integer, *not* a long integer). Choose Data Type = *Number* and Field Size= *Integer*. We will use the value 1 to denote Cash, the value 2 to de-note Visa, etc. We will look closer at this in section 2.4.

state must also be an enumeration type. Make it an in-teger. Here the value 1 will denote *booked*, 2 *in*, etc.

tblRoomType:

Contains one record for each type of room, for instance one for double rooms, one for single rooms, etc. (In the book *User Interface Design*, we added this table late in the design process to illustrate the normalization concept.)

roomType is an artificial key. An AutoNumber is okay.

description is a short text, for instance "double room, bath".

bedCount is the number of beds in the room, including temporary beds.

price1 and price2 are the standard price and a possible discount price. The price should be a decimal num-ber. Choose Data Type=*Number*, Field Size= *Sin-gle*, Decimal Places =2.

tblRoom:

roomID is a natural key - the number on the door. So don't use an AutoNumber. Use an integer.

roomType is a foreign key that refers to tblRoomType. (You should by know how to deal with it.)

tblRoomState:

stayID and roomID are foreign keys. Ensure their types match what they refer to. Notice that roomID refers to a natural key, not to an AutoNumber.

date should be a Date/Time field with Format = *Short Date*.

personCount is the number of persons staying in the room. An integer should suffice.

state is similar to state for tblStay, although the values are slightly different.

The key consists of two fields: roomID and date. It is a bit tricky to specify this: select both fields by clicking on the left-hand marker (hold down Ctrl while selecting the second field). Then right-click somewhere on the text *inside* the line.

Optional tables

The following two tables are needed for the full system.

tblServiceType:

serviceID is an artificial key. Should be an Auto-Number.

name and price should be obvious. The price should be a decimal number. Choose Data Type=*Number,* Field Size= *Single*, Decimal Places =2.

tblServiceReceived:

stayID and serviceID are foreign keys that refer to AutoNumbers. The foreign keys must thus be long integers.

roomID is an optional reference to a room. An integer should suffice. (This reference is needed when a waiter records a service for a specific room and the guest has more than one room.)

date should be a Date/Time field. Choose Format = *Short Date*.

quantity is the number of items the guest has got - an integer should suffice.

ACTIVITIES

1. Reflect back to an office that you are familiarized with, if the office was using office automation systems, do you think that the use of OAS increased or decreased printed papers? Why?
2. What are the primary factors that facilitate or inhibit the trend towards a paperless office? What were the problems faced by management? What management, organization and technology factors were responsible for those problems?

4.3 Choosing required data types

Data types

Data is stored in the computer according to its type. Here is a description of the most important types in the data base. Visual Basic deals with almost the same types **Text**. The field can contain any characters. The Field Size property defines the maximum number of characters. The maximum cannot be above 255 characters.

Memo: Like a text field, but the maximum number of characters is 65,535. Access takes more time to process a memo field, so use text fields if adequate.

Number: The field can contain a number. The Field Size property defines what kind of number:

. **Integer**. A small integer. It must be in the range -32,768 to +32,767 (a 16-bit integer).

. **Long Integer**. It must be in the range from around -2,140 million to +2,140 million (a 32-bit integer).

. **Single**. A decimal number in the range from -3.4*1038 to +3.4*1038 with an accuracy of 6 or 7 significant digits (a 32-bit floating point number).

. **Double**. A decimal number in the range from -1.8*10308 to +1.8*10308 with 14 significant digits (a 64-bit floating point number).

. **Decimal**. A very long integer with a decimal point placed somewhere. Intended for monetary calculations where rounding must be strictly controlled. In the book we use Single or Double instead.

Numbers can be shown in many ways depending on the format property of the field. You may for instance show them with a fixed number of decimals, with a currency symbol, etc.

Some formats show data in a way that depends on the regional settings of the computer. If you for instance specify the format of a number as *Currency*, the number will show with a $ on a US computer and with a £ on a British computer.

Date/Time: The field gives a point in time. In the computer it is stored as the number of days since 30/12-1899 at 0:00. It is really a Double number, so the number of days may include a fraction of a day. In this way the field specifies the date as well as the time with high precision. As an example, the number 1 corresponds to 31/12-1899 at 0:00, the number 1.75 to 31/12-1899 at 18:00 (6 PM).

Usually we don't show a date field as a number, but as a date and/or a time. The format property specifies this. Also here you can choose a format that adapts to the regional setting.

Yes/No: The field contains a Boolean value shown either as Yes/No, True/False, or On/Off. The format property specifies this.

AutoNumber: The field is a long integer (32 bits) that Access generates itself as a unique number in the table. Access numbers the records 1, 2, as you enter the records. However, you cannot trust that the sequence is unbroken. For instance when you add a record and undo the addition before having completed it, Access uses the next number in the sequence anyway.

The following table shows limitations on Microsoft Access data types.

Data type	**Description**
BINARY, VARBINARY, and VARCHAR	Creating a BINARY, VARBINARY, or VARCHAR column of zero or unspecified length actually returns a 510-byte column.
BYTE	Even though a Microsoft Access NUMBER field with a FieldSize equal to BYTE is unsigned, a negative number can be inserted into the field when using the Microsoft Access driver.
CHAR, LONGVARCHAR, and VARCHAR	A character string literal can contain any ANSI character (1-255 decimal). Use two consecutive single quotation marks (") to represent one single quotation mark ('). Procedures should be used to pass character data when using any special character in a character data type column.
DATE	Date values must be either delimited according to the ODBC canonical date format or delimited by the date time delimiter ("#"). Otherwise, Microsoft Access will treat the value as an arithmetic expression and will not raise a warning or error. For example, the date "March 5, 1996" must be represented as {d '1996-03-05'} or #03/05/1996#; otherwise, if only 03/05/1993 is submitted, Microsoft Access will evaluate this as 3 divided by 5 divided by 1996. This value rounds up to the integer 0, and since the zero day maps to 1899-12-31, this is the date used. A pipe character (|) cannot be used in a date value, even if enclosed in back quotes.

GUID	Data type limited to Microsoft Access 4.0
NUMERIC	Data type limited to Microsoft Access 4.0

A **foreign key** is a field (or several fields) that refer to something unique in another table - usually the primary key. Be careful here. The foreign key and the primary key must have the same type. However, when the primary key is an AutoNumber, the foreign key must be a long integer.

Changing a data type: Access is quite liberal with changing a data type to something else - even if there are data in the records. It can also change an Auto-Number field to a number field, but not the other way around. If you need to change field B to an Auto-Number, create a new field C and make it an Auto-Number. Then delete field B and rename field C to B.
If you for some reason want to store a record with an AutoNumber of your own choice (for instance create a stay with stayID=728), you need to append the record with an INSERT query (see section 7.1). You cannot just type in the stayID.

The majority of publically available Internet information sources are: 1. created in XML. 2. structured information. 3. normal information. 4. unstructured information. 5. none of the above.

SELF ASSESSMENT QUESTIONS

Please select the correct option:

A. Records management:
   1. is a discipline limited to digitised paper documents.
   2. is a discipline limited to library books.
   3. is a discipline limited to paper documents.
   4. is a discipline limited to information contained in databases.
   5. none of the above.

B. Which of the following should be represented on an information flow diagram?
   1. Entity
   2. Source
   3. Process
   4. Attribute
   5. Database

C. A web blueprint depicts:

1. the layout of an individual web page
2. the layout of the home page
3. the layout of an index page
4. the layout of a website
5. the layout of a sitemap

D. UML depicts information systems as a collection of:

3. Entities 2. Processes 3. Data 4. Information 5. Objects

4.4 Defining fields properties

Every field in a table has properties. These properties define the field's characteristics and behavior. The most important property for a field is its data type. A field's data type determines what kind of data it can store. For example, a field whose data type is Text can store data that consists of either text or numerical characters, but a field whose data type is Number can store only numerical data.

A field's data type determines many other important field qualities, such as:

- How you can use the field in expressions.
- The maximum size of a field value.
- Whether the field can be indexed.
- Which formats can be used with the field.

When you create a new field in Design view, you specify the field's data type and, optionally, its other properties.

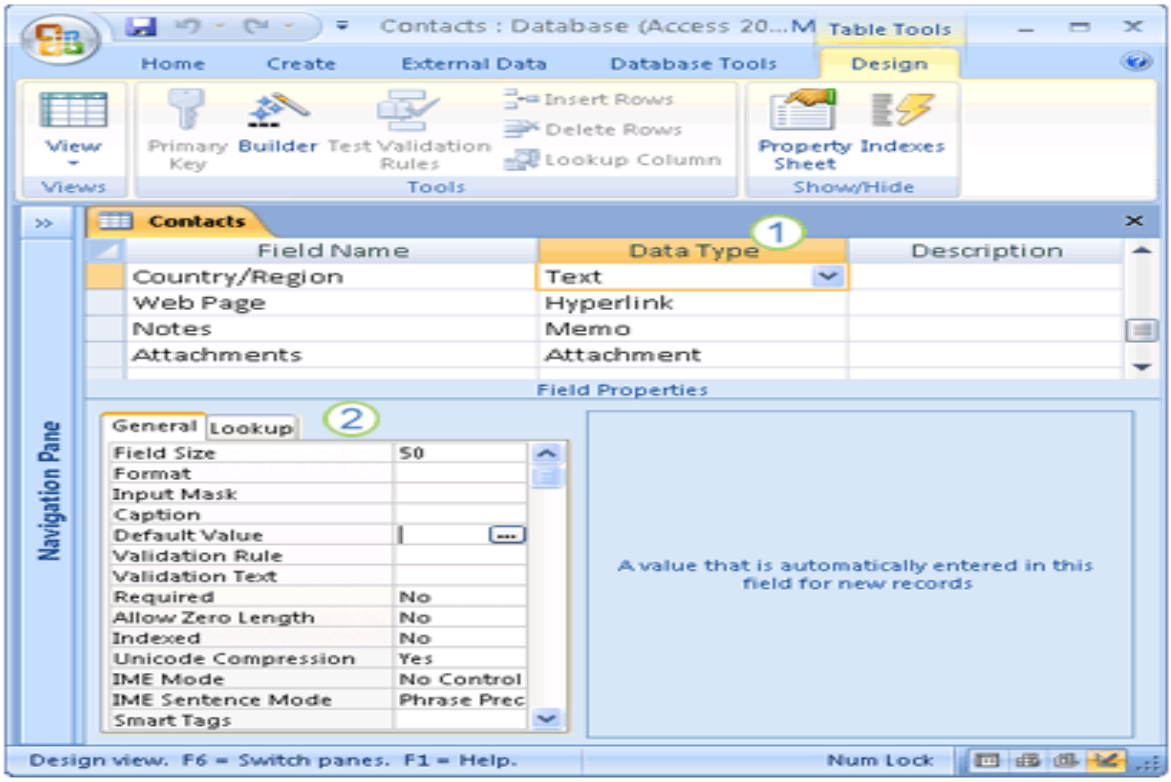


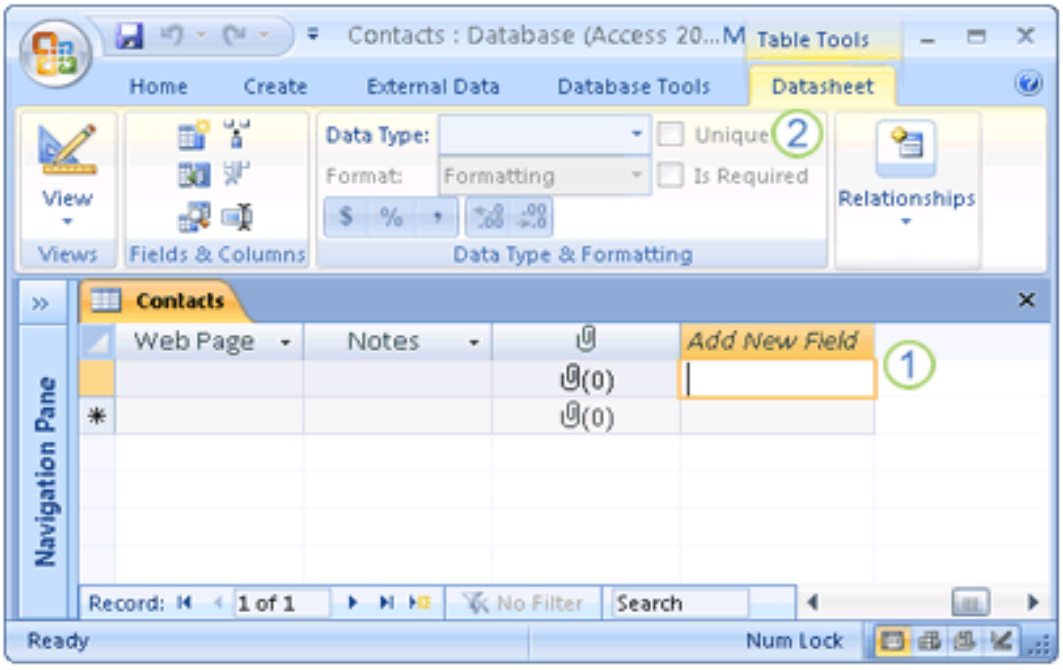
Figure: Contacts table open in Design view

1. Data type

2. Field properties

When you create a field in Datasheet view, the field's data type is defined for you. If you create a field in Datasheet view by using a field template or by using an existing field from another table, the data type is already defined in the template or in the other table. If you create a field by entering data in Datasheet view, Microsoft Office Access assigns a data type to the field based on the values that you enter. If you enter values that are of different data types in the field, Access may prompt you to make a decision about the data type.

You can change the field's data type and its **Format**, **Indexed**, and **Required** field properties in Datasheet view.



Contacts table open in Datasheet view

1. Create a field by entering data in an empty column.

2. Adjust the field's data type and other properties by using the **Datasheet** tab on the Ribbon.

4.5 Setting primary key

A **primary key**, also called a **primary** keyword, is a **key** in a relational database that is unique for each record. It is a unique identifier, such as a driver license number, telephone number (including area code), or vehicle identification number (VIN). A relational database must always have one and only one **primary key**.

The term Primary Key (PK) in a Relational Database System refers to a field (or combination of fields) whose values uniquely identify each record in that table. In Access, it is common to define a Primary Key field in each table, usually as a field that is of AutoNumber data type, to serve as the PK. This AutoNumber field will automatically create a unique number for each record as it is created, so that those records can be identified from the other records in the table.

A primary key's main features are:

- It must contain a unique value for each row of data.
- It cannot contain null values.

A primary key is either an existing table column or a column that is specifically generated by the database according to a defined sequence.

The primary key concept is critical to an efficient relational database. Without primary key and closely related foreign key concepts, relational databases would not work.

Almost all individuals deal with primary keys frequently but unknowingly in everyday life. For example, students are routinely assigned unique identification (ID) numbers, and all U.S. citizens have government-assigned and uniquely identifiable Social Security numbers.

For example, a database must hold all of the data stored by a commercial bank. Two of the database tables include the CUSTOMER_MASTER, which stores basic and static customer data (name, date of birth, address, Social Security number, etc.) and the ACCOUNTS_MASTER, which stores various bank account data (account creation date, account type, withdrawal limits or corresponding account information, etc.).

To uniquely identify customers, a column or combination of columns is selected to guarantee that two customers never have the same unique value. Thus, certain columns are immediately eliminated, e.g., surname and date of birth. A good primary key candidate is the column that is designated to hold Social Security numbers. However, some account holders may not have Social Security numbers, so this column's candidacy is eliminated. The next logical option is to use a combination of columns such as the surname to the date of birth to the email address, resulting in a long and cumbersome primary key.

The best option is to create a separate primary key in a new column named CUSTOMER_ID. Then, the database automatically generates a unique number each time a customer is added, guaranteeing unique identification. As this key is created, the column is designated as the primary key within the SQL script that creates the table, and all null values are automatically rejected.

The account number associated with each CUSTOMER_ID allows for the secure handling of customer queries and also demonstrates why primary keys offer the fastest method of data searching within tables. For example, a customer may be asked to provide his surname when conducting a bank query. A common surname (such as Smith) query is likely to return multiple results. When querying data, utilizing the primary key uniqueness feature guarantees one result.

Foreign Key

The term Foreign Key (FK) in a Relational Database System refers to a field that in a table that stores values from a Primary Key field in another table, to relate the two records to each other. In the context of relational databases, a **foreign key** is a field (or collection of fields) in one table that uniquely identifies a row of another table or the same table. In simpler words, the **foreign key** is defined in a second table, but it refers to the **primary key** in the first table. The table containing the foreign key is called the **child table**, and the table containing the candidate key is called the **referenced** or **parent table**. In database relational modeling and implementation, a unique key is a set of zero or more attributes, the value(s) of which are guaranteed to be unique for each tuple (row) in a relation. The value or combination of values of unique key attributes for any tuple cannot be duplicated for any other tuple in that relation.

When more than one column is combined to form a unique key, their combined value is used to access each row and maintain uniqueness. Values are not combined, they are compared using their data types.

Since the purpose of the foreign key is to identify a particular row of the referenced table, it is generally required that the foreign key is equal to the candidate key in some row of the primary table, or else have no value (the NULL value). This rule is called a **referential integrity constraint** between the two

tables.[5] Because violations of these constraints can be the source of many database problems, most database management systems provide mechanisms to ensure that every non-null foreign key corresponds to a row of the referenced table.

For example, consider a database with two tables: a CUSTOMER table that includes all customer data and an ORDER table that includes all customer orders. Suppose the business requires that each order must refer to a single customer. To reflect this in the database, a foreign key column is added to the ORDER table (e.g., CUSTOMERID), which references the primary key of CUSTOMER (e.g. ID). Because the primary key of a table must be unique, and because CUSTOMERID only contains values from that primary key field, we may assume that, when it has a value, CUSTOMERID will identify the particular customer which placed the order. However, this can no longer be assumed if the ORDER table is not kept up to date when rows of the CUSTOMER table are deleted or the ID column altered, and working with these tables may become more difficult. Many real world databases work around this problem by 'inactivating' rather than physically deleting master table foreign keys, or by complex update programs that modify all references to a foreign key when a change is needed.

Foreign keys play an essential role in database design. One important part of database design is making sure that relationships between real-world entities are reflected in the database by references, using foreign keys to refer from one table to another.[9] Another important part of database design is database normalization, in which tables are broken apart and foreign keys make it possible for them to be reconstructed.

Multiple rows in the referencing (or child) table may refer to the same row in the referenced (or parent) table. In this case, the relationship between the two tables is called a one to many relationship between the referenced table and the referencing table.

In addition, the child and parent table may, in fact, be the same table, i.e. the foreign key refers back to the same table. Such a foreign key is known in SQL:2003 as a **self-referencing** or **recursive** foreign key. In database management systems, this is often accomplished by linking a first and second reference to the same table.

A table may have multiple foreign keys, and each foreign key can have a different parent table. Each foreign key is enforced independently by the database system. Therefore, cascading relationships between tables can be established using foreign keys.

Another example, assume that you have some contacts for which you want to store some information, and that information happens to include the phone numbers for each contact. You realize that some contacts have more phone numbers than others and you want to store the data in your database application as efficiently as possible. So, you decide to create an Access database application that has a Contacts table and a PhoneNumbers table to store this data. In your database application, you need a way to identify any number

of phone number records that relate to each contact record. If the PK in the Contacts table is an AutoNumber type field, then the FK in the PhoneNumbers table should store the ContactID (AutoNumber) from the Contacts table and will need to be the type of Long Integer (Number). This means that a record in the PhoneNumbers table has the potential to contain the ContactID value of the contact record in the related Contacts table.

1 Open a table in Design view.

2 Click the Row Selector for each field you want to set for the primary key. The Row Selector is the small box that allows you to select an entire row
3 Click on the Primary Key icon in the toolbar, or select Primary Key from the Edit menu.
4 If you've selected a field for a Primary Key that already contains data or returns a null value, MS Access will not allow you to create a primary key.

4.6 Defining indexes

An index is a list of data, such as group of files or database entries. It is typically saved in a plain text format that can be quickly scanned by a search algorithm. This significantly speeds up searching and sorting operations on data referenced by the index. Indexes often include information about each item in the list, such as metadata or Glossary that allows the data to be searched via the index instead of reading through each file individually.

For example, a database program such as Microsoft Access may generate an index of entries in a table. When an SQL query is run on the database, the program can quickly scan the index file to see what entries match the search string. Search engines also use indexes to store a large list of Web pages. These indexes, such as those created by Google and Yahoo!, are necessary for quickly generating search results. If search engines had to scan through millions of pages each time a user submitted a search, it would take roughly forever. Fortunately, by using search indexes, Web searches can be performed in less than a second instead of several hours.

The term "index" can also be used as a verb, which not surprisingly means to create an index. It may also refer to adding a new item to an existing index. For example, Mac OS X 10.4 and later indexes the hard disk to create a searchable index for Apple's Spotlight search utility. Google's "Googlebot" crawls the Web on a regular basis, adding new Web pages to the Google index. While most database and hard disk indexes are updated on-the-fly, search engine indexes are only updated every few hours, days, or even weeks. This is

why newly published Web pages may not show up in search engine results. While it may be a frustration for Web developers, it is a small price to pay for the convenience of super-fast Web searches.

Indexing a table is simply a way of organizing the data in the table to allow Access to complete query searches and sorting more rapidly. Indexing can help speed up the time that it takes to complete queries in Access, given a few criteria are met first. First of all, index tables that have a variety of different data types within their fields. Second, indexing is more efficient if the data in the indexed fields give each record a more unique identification (like a primary key field). Indexing is not usually necessary on fields that have multiple repeating values. Third, it is required to index fields which are used for criteria in queries. For example, if you are creating many queries that find records based on phone numbers, you may want to create an index on the field which contains the phone numbers. Assuming your table fields have met these criteria, it can be useful to apply an index to the desired fields to increase the sorting and processing capabilities of data in queries.

Unfortunately, if users apply an index to a table that contains multiple similar data types, or has multiple duplicate information in each indexed field, you may actually slow down the speed at which Access updates table information. In these cases, it is actually preferable to not index these fields, as indexing these types of fields will rob you of more time than the time saved by using the indexed fields in the query.

Also, Access decides when and if it will choose to use the indexes that you create. All that you can do as the user is simply create the indexes for the table fields. Access will decide when and if to use the indexes when performing a query. Many times, when a query is performed on a table which has very little information, Access will perform a “table scan” on the table records, looking at all of the data in all of the fields and then extracting the requested results. However, as you add more and more data to the tables, Access may find that it is easier to index (sort) the table by one of the available indexes first, and then extract the requested records. However, it cannot do this if there are no indexes for it to take advantage of.

To view the indexes that has been created for a table, simply open up the table in design view. Next, click the “Indexes” button in the “Show/Hide” group on the “Design” tab of the “Table Tools” contextual tab to view the indexes for your selected table in a separate “Indexes” dialog box. Almost every table will have at least one index: the primary key. The primary key is a type of index. Notice that when you open a table in datasheet view, it sorts the view by the values in the primary key column, by default.

Creating Indexes

When you are creating indexes, you want to try and use field values that will identify each record in your table as uniquely as possible. If you are a good database designer, there will already be a single field in your table that already does this: your primary key field. However, you can create additional indexes on other fields to use in queries for faster query processing.

To create an index, open up the table which you would like to index in table design view. Then click the "Indexes" button in the "Show/Hide" group on the "Design" tab of the "Table Tools" contextual tab to show the "Indexes" dialog box. Click into the next available row under the "Index Name" column, and type a name for your new index. To the right, click into the "Field Name" column, and select the name of the field within the table which you wish to index. To the right of that, select whether that field should be sorted in "Ascending" (A-Z, 1-9) or "Descending" (Z-A, 9-1) order.

In the "Index Properties" section at the bottom of the dialog box, you have three drop-down text boxes into which you can set the properties of the index. The first property is "Primary," and can accept either a "Yes" or "No" value. Whatever index is the primary key of the table will show "Yes" in this property, and all others will show "No." There can only be one primary key field in a table.

The next property, "Unique," asks if the values within the field will always be unique (like the values in a primary key field are). Once again, you can select either "Yes" or "No," as appropriate.

You can then also set the "Ignore Nulls" property to "Yes" or "No" to either include or exclude "Null" (empty) values from the sorting. Nulls occur when there has been no data entry in the field for a record. For example, if you skipped entering an address into a customer record, the address field would contain a "null" value. It is not equivalent to zero, as zero is still a value. "Null" is simply "unknown."

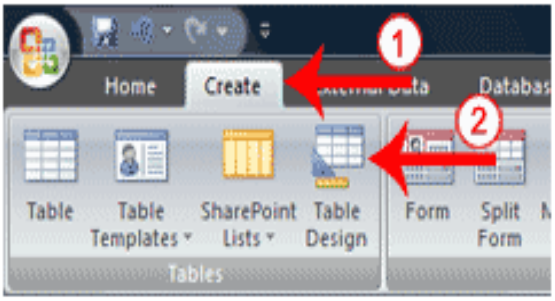
When you have finished creating the index, you can close the "Indexes" dialog box. Then click the "Save" button in the Quick Access toolbar to save your structural modifications to the table.

4.7 Saving the table

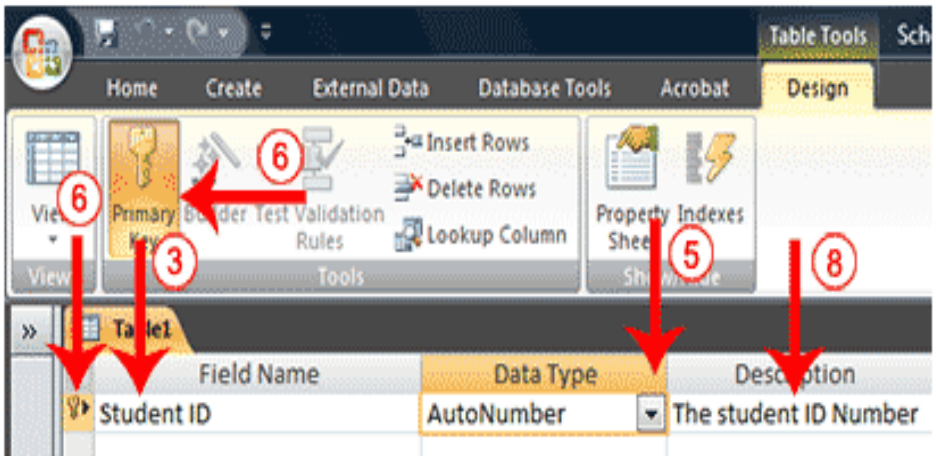
To save the table follow the steps:

1. Click the Save button on the Quick Access toolbar. The Save As dialog box appears.
2. Type the name you want to give your table.
3. Click OK. Access names your table.

To use Design view to create a new table:

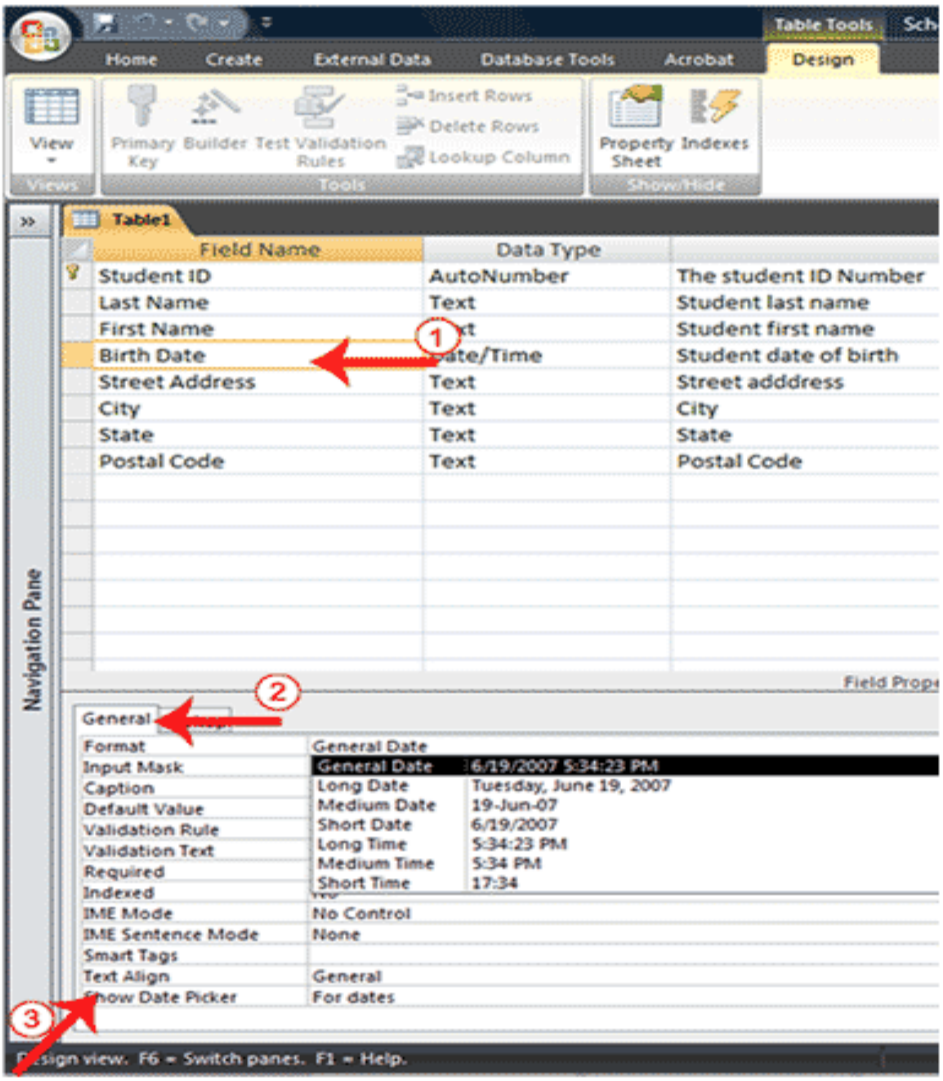


1. Activate the Create tab.
2. Click Table Design in the Tables group. Access changes to Design view and the Table Tools become available



3. Type the first field name in the Field Name field.
4. Press the Tab key.
5. Click the down-arrow that appears when you click in the Data Type field and then select a data type.
6. Click Primary Key if the column you created is a primary key. A small key appears next to the field name.
7. Press the Tab key.
8. Type a description. The description is optional.
9. Press the Tab key. Access moves to the Field Name field.
10. Repeat steps 3 through 10 until you have created all of your fields

4.7 Modifying table structure



1. Click the field for which you want to set the field properties.
2. Activate the General tab in the Field Properties area.
3. Set the properties you want to set.
4. Repeat steps 1 through 3 until you have set all the properties for all fields.

You can use Design view to create or modify a table. After you finish the task, you must save the table by clicking the Save button on the Quick Access toolbar.

1. Click the Save button on the Quick Access toolbar. Access saves the table unless you are saving for the first time. If you are saving for the first time, the Save As dialog box appears.
2. Type the name you want to give your table.
3. Click OK. Access saves the table. You can now access the table by using the Navigation pane.

4.9 Establishing relationship among tables

When we have several tables, we can make relation-ships (crow's feet). Then we get an E/R model instead of a simple collection of tables. The relationships allow Access to help us retrieve data across tables, check referential integrity, etc.

Figure below shows the hotel relationships in Access. It resembles the crow's feet model quite well. You define the relationships in this way:

1. Start in the database window and right-click somewhere.
2. Choose *Relationships*.

Now you see an empty *Relationship Window*. You have to tell Access which tables to show here. Some-times a *Show Table* window pops up by itself. Other-wise you have to invoke it with a right-click in the relationship window.

3. In the *Show Table* window, select the tables you want to include. In the hotel system it is all the tables.
4. Click *Add* and close the window. Now the tables should be in the relationship window.
5. Create the relationship between tblGuest and tblStay by dragging guestID from one table to guestID in the other.
6. An edit-relationship window pops up. If not, right-click on the relationship connector and choose the edit window.

In the edit-relationship window, you can specify foreign keys that consist of several fields. You can also specify that the relationship has referential integrity, so that all records on the m-side point to a record on the 1-side.

7. In our case, all stays must point to a guest, so mark the connector *enforce referential integrity*. (If Access refuses this, it is most likely because you have not defined the foreign key as a long integer.)
8. Close the relationship window. The relationship connector now appears in the window between the foreign key and its target.

The referential integrity makes Access show the con-nector as 1-∞ (1:m). Based on referential integrity and whether the connected fields are primary keys, Access may also decide that it is a 1:1 relationship. It is not important what Access decides in these matters. You can later tell it otherwise when you want to use the connector.

9. Create the remaining relationships too. Note that there is no referential integrity between tblStay and tblRoomState. It is on purpose - if the room is in repair state there is no connected stay.

Partial integrity. Access provides a more relaxed version of referential integrity. It allows the foreign key to be either empty (Null) or point to a record on the 1-side. This is the case for the relationship between tblStay and tblRoomState. Give it partial integrity in this way:

10. Open tblRoomState in design view. For stayID (the foreign key) set the Default Value to empty (delete all characters in the field). Also set *Required* to *No*.

11. In the relationship window, right-click on the connector and choose the edit window. Select *enforce referential integrity*.

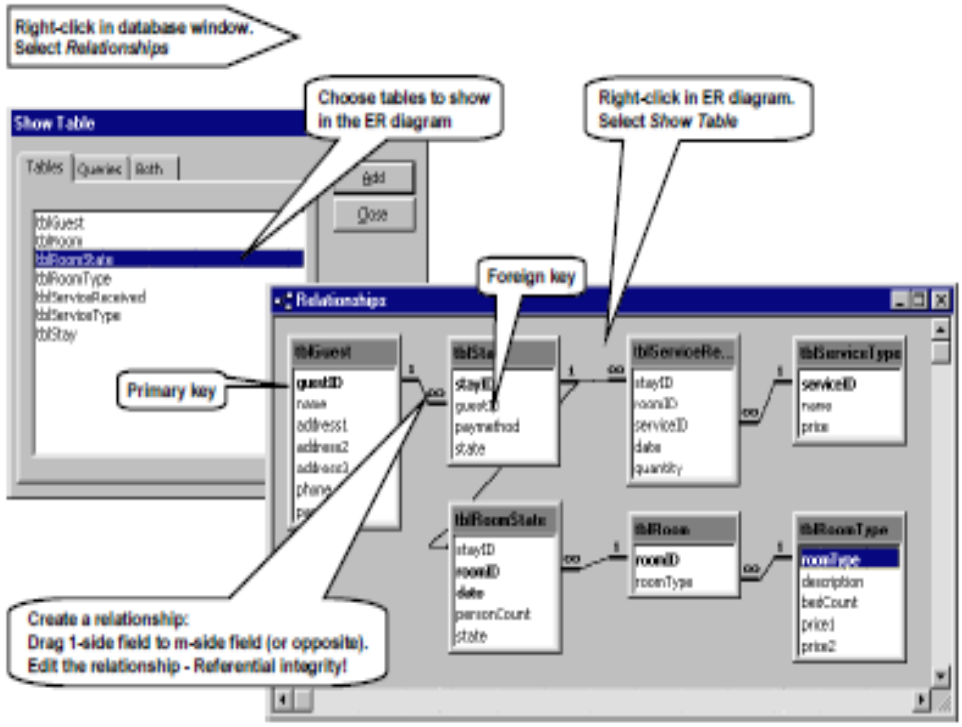


Figure: Create Relationships

Note that you cannot see in the relationship window whether the relationship has full or partial referential integrity.

Deleting a relationship. If you need to delete a relationship, click it and press Del

4.10 Entering and viewing data from tables

Views are different ways of looking at the same object. Tables have four views: Datasheet view, Pivot Table view, Pivot Chart view, and Design view. You use Datasheet view to create a table, edit data, or view data; Pivot Table view to create a pivot table; Pivot Chart view to create a pivot chart; and Design view to create a table or modify an existing table.

To change the view:

1. Activate the Home tab.
2. Click the down-arrow under the View button. A menu appears.
3. Click the view you want. Access changes to the view you chose.

Tip: You can also use a template to create a table. Access has several templates from which you can choose. When using a template, you create the table and then modify it to suit your needs

1. Activate the Create tab.
2. Click the Table Templates button in the Tables group. A menu appears.
3. Click the template you want to use. Access creates a table based on the template

4.11Creating queries by example

1. Start in the database window. Select *Queries* and *Create query in Design view*. (In Access 97 select *New* and then *Design view*.)
2. Access asks you to select the tables you want to combine. Select tblGuest and tblStay (Figure 4.11). Click *Add* and then *Close*.

User now sees the query design window (middle of Figure 4.11). The top part of the window is an E/R-model, in our case consisting of tblGuest and tblStay. Access has included the relationship from the full E/R-model. It shows that the tables will be combined according to guestID. This is just what we want in our case, but in other cases you have to remove the relationships you don't need and add new ones that you need for the query. These changed relationships are only used in the query; they don't influence the full E/R-model.
User may delete tables from the query window or add further tables by right-clicking in the E/R-model.
In the lower part of the window, you see the **query grid** where we will make a column for each field in the computed table.

3. Drag *stayID* from tblStay to the grid. Then drag *name*, *address1* and *phone* from tblGuest. Finally, drag *state* from tblStay. (You may also double-click the fields.)

User may rearrange the columns by selecting a column and dragging it to another place.

4. Switch to datasheet view. The query table should look like the bottom of the figure. It contains all stays recorded in the database with guest information attached. In the example, John Simpson has three stays and Yun Chen two stays.
5. Save the query and give it the name **qryStayList**. (The standard prefix for queries is *qry*.)

This looks almost too easy. What happens really? Access has made a so-called **join** of tblGuest and tblStay. According to the E/R-model, each record in tblStay has a connecting string to a record in tblGuest. In the query table, there will be one record for each of these strings. If one of the stays didn't have a string to a guest, this stay would not occur in the result.

Since each query record corresponds to a string between two source records, we can include arbitrary fields from both source tables. This is what we have done.

Star = all fields? Note that the data model at the top of the query window has a star in each box. It means "all fields". You may drag it to the grid and all source fields will be included in the query table. This may be con-venient, but don't do it yet. Why? Assume that you drag the star from both tables. Then you will have two guestID fields in the result. You then have to refer to them as tblStay.guestID and tblGuest.guestID. This leads to endless confusion later, particularly because some of Access's built-in Wizards cannot figure out about these names and screw things up.

Dynaset and data entry to query table

The query not only shows data, it can also be used for data entry. Try these experiments:

6. Open qryStayList and tblGuest at the same time.
7. In the *query* table, change the name of one of the guests. As soon as you move the cursor to the next line, you will see the change in the query table as well as the guest table.
8. In tblGuest, try to change the guest name again. The query table will be updated immediately.

The query table we look at is a *dynaset* because it is updated automatically. We can enter data into it be-cause it is a simple query where Access can find out how to store the data into the source tables. For more complex queries, this is not possible.

The query has a property called *Recordset Type*. It is *Dynaset* as a standard, but you can change it to *Snap-shot*. Then Access computes the record list when you open the query, and doesn't update it dynamically. In this case you cannot enter data through the query. (How to find the Recordset Property? From the query design window, use *View -> Properties*. But don't change anything right now.)

Adding/deleting records in a dynaset

1. In the query table, enter a guest name in the last line (with star-indication). When you move the cursor up, you have created a new guest record (but not a new stay record). You cannot see it in tblGuest, but if you close and open tblGuest, you will see it. (Using the sort button A/Z on the tool-bar will also show the new guest.)
2. In the query table, enter a state in the last line and move the cursor up. Access refuses to do it. It tries to create a stay record, but lacks the foreign key to the guest and cannot preserve the referential integrity. If

you had included the foreign key in the query, you could set it now and succeed. Never mind. (Remember to use Esc to get out of the in-consistent data update.)

3. In the query table, try to delete a line. What you delete is the stay record, not the guest record.

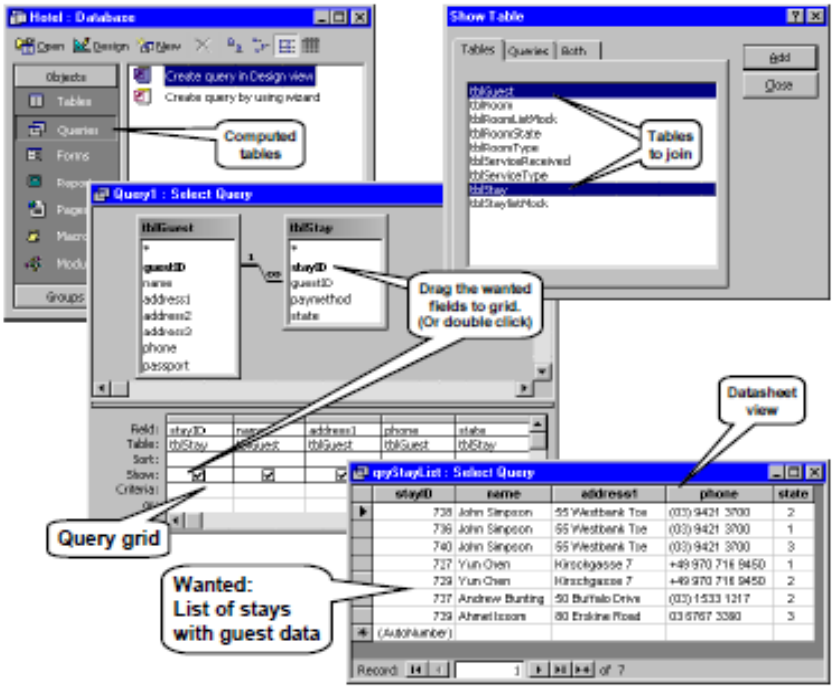
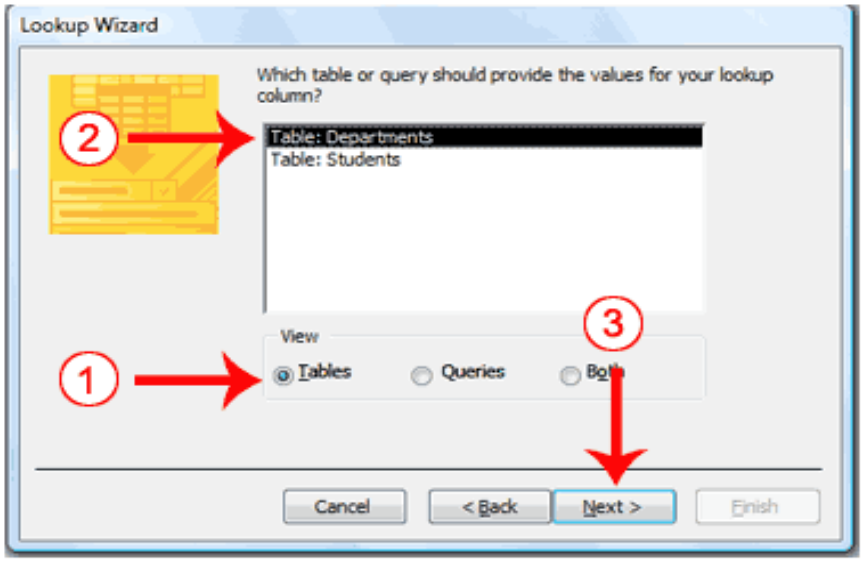


Figure: 4.11 **Query - join two tables**

Summary. The dynaset is suitable for editing data produced by a simple query. As soon as you fill in a field in the new record line (star-marked), Access will try to create a new record. If you fill in a field from the guest table, Access will make a guest record. If you fill in a field from the stay table, Access will (also) make a stay record. When you move the cursor to another record, Access will check that referential integrity and other rules are met.

4.12 Running a query

A lookup column can be based on a table, a query, or a list of values you type. If you base your lookup column on a table or query, you must create the table or query before creating the lookup column. A query is a list of rows and columns based on one or more tables. A query only displays the rows and columns you specify.



1. Click a radio button to select what you want to base your lookup column on. Choose from Tables, Queries, or Both.
2. Click to select the table or query you want.
3. Click Next. The Lookup Wizard moves to the next page

4.13 Summary

In Microsoft Access a database consists of one single file. The file contains all the tables of the database, the relationships (the crow's feet), *queries* (computed tables), *forms* (user windows), and many other things.

As a systems developer you will *design* tables and user windows. As a user you will enter data into the tables (usually through user windows) and get data out of the tables, for instance through the same windows or through printed reports.

In Access it is very easy to switch between the developer role and the user role. As a developer you will typically design some tables, then switch to the user role to enter data into them, then switch back to the developer role to change the design, design more tables, etc. Access can to a large extent restructure the data that already is in the database so that it matches the new table design.

The following two tables are needed for the full Optional table system.

tblServiceType:

serviceID is an artificial key. Should be an Auto-Number.

name and price should be obvious. The price should be a decimal number. Choose Data Type=*Number*, Field Size= *Single*, Decimal Places =2.

tblServiceReceived:

stayID and serviceID are foreign keys that refer to AutoNumbers. The foreign keys must thus be long integers.

roomID is an optional reference to a room. An integer should suffice. (This reference is needed when a waiter records a service for a specific room and the guest has more than one room.)

date should be a Date/Time field. Choose Format = *Short Date*.

quantity is the number of items the guest has got - an integer should suffice.

4.15 Glossary

Memo: Like a text field, but the maximum number of characters is 65,535. Access takes more time to process a memo field, so use text fields if adequate.

Number: The field can contain a number. The Field Size property defines what kind of number:

Integer. A small integer. It must be in the range -32,768 to +32,767 (a 16-bit integer).

Long Integer. It must be in the range from around -2,140 million to +2,140 million (a 32-bit integer).

. **Single**. A decimal number in the range from -3.4*1038 to +3.4*1038 with an accuracy of 6 or 7 significant digits (a 32-bit floating point number).

Double. A decimal number in the range from -1.8*10308 to +1.8*10308 with 14 significant digits (a 64-bit floating point number).

Decimal. A very long integer with a decimal point placed somewhere. Intended for monetary calculations where rounding must be strictly controlled. In the book we use Single or Double instead.

Primary Key (PK): it in a Relational Database System refers to a field (or combination of fields) whose values uniquely identify each record in that table. In Access, it is common to define a Primary Key field in each table, usually as a field that is of AutoNumber data type, to serve as the PK. This AutoNumber field will automatically create a unique number for each record as it is created, so that those records can be identified from the other records in the table.

Foreign Key (FK): The term Foreign Key (FK) in a Relational Database System refers to a field that in a table that stores values from a Primary Key field in another table, to relate the two records to each other.

Index: An index is a list of data, such as group of files or database entries. It is typically saved in a plain text format that can be quickly scanned by a search algorithm. This significantly speeds up searching and sorting operations on data referenced by the index. Indexes often include information about each item in the list, such as metadata or Glossary that allows the data to be searched via the index instead of reading through each file individually.

4.16 References

O'Brien, J (1999). *Management Information Systems – Managing Information Technology in the Internetworked Enterprise*. Boston: Irwin McGraw-Hill. ISBN 0-07-112373-3.

Laudon, Kenneth C.; Laudon, Jane P. (2009). *Management Information Systems: Managing the Digital Firm* (11 ed.). Prentice Hall/CourseSmart. p. 164.

Transaction processing systems (TPS) collect and record the routine transactions of an organization. Examples of such systems are sales order entry, hotel reservations, payroll, employee record keeping, and shipping.

Bidgoli, Hossein, (2004). The Internet Encyclopedia, Volume 1, John Wiley & Sons, Inc. p. 707.

Pant, S., Hsu, C., (1995), Strategic Information Systems Planning: A Review, Information Resources Management Association International Conference, May 21–24, Atlanta.

4.17 Further Readings

1. Taylor, Victoria. "Supply Chain Management: The Next Big Thing?". *Sept. 12, 2011*. Business Week. Retrieved 5 March 2014.
2. Lynn, Samara. "What is CRM?". PC Mag. Retrieved 5 March 2014.
3. Joshi, Girdhar (2013). *Management Information Systems*. New Delhi: Oxford University Press. p. 328. ISBN 9780198080992.
4. Laudon, K.,&Laudon, J. (2010). Management information systems: Managing the digital firm. (11th ed.). Upper Saddle River, NJ: Pearson Prentice Hall

4.18 Model Questions

Q:1 Explain how to create a database in MS Access.

Q:2 Explain the various data types in MS Access

Q:3 Define the fields properties.

Q:4 Explain the steps to create and save table in MS Access

Q:5 Explain the concept of primary and secondary key.

ANSWERS TO SELF ASSESSMENT QUESTIONS

A. 4, B- 5, C- 2, D- 4, E-5

CHAPTER 5

Building IS

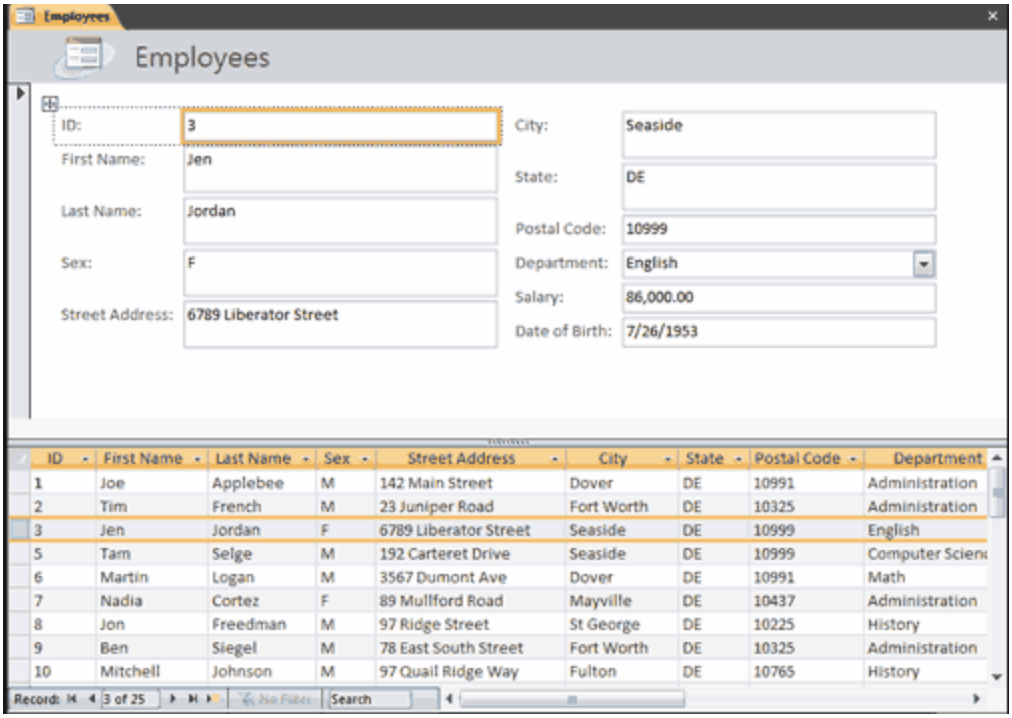
STRUCTURE

5.0Objectives

5.1 Basics of MS-Access Forms and Reports

Forms

Definition: A form in Access is a database object that you can use to create a user interface for a database application. A "bound" form is one that is directly connected to a data source such as a table or query, and can be used to enter, edit, or display data from that data source.



Although information in a database can be entered and edited directly in a table, most people find it simpler to use a form. We use forms all the time in everyday life as a way of recording information so forms are familiar to us. The Form design tools in Access are very flexible and allow you to customize a form with many features to make it easy to use. Forms can be created in a few different ways.

- Use an Auto form to create a form based on a standard layout.
- Use the Form Wizard.
- Use one of the above methods and then modify the form in Design View.
- Create a form completely from scratch using the Design View tools.

As you become more and more familiar with working with Forms in Access you may find yourself using the latter methods more and more. To begin with though, it is best to use the Auto forms and Form Wizards until you are more comfortable with designing forms.

Using Auto forms

Auto forms allow you to create a form quickly based on a standard layout. In the first exercises, you actually used an auto form when you used a form for data entry. There are several auto form layouts to choose from and we will try out three of them in the following exercises.

Use the Form tool to create a new form

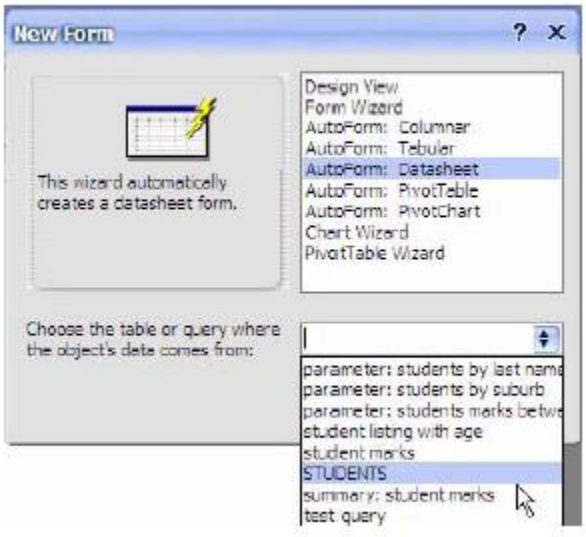
1. In the Navigation Pane, click the table or query that contains the data you want to see on your form.
2. On the **Create** tab, in the **Forms** group, click **Form**.

Access creates the form and displays it in Layout view. In Layout view, you can make design changes to the form while it is displaying data. For example, you can adjust the size of the text boxes to fit the data, if necessary.

If Access finds a single table that has a one-to-many relationship with the table or query that you used to create the form, Access adds a datasheet to the form that is based on the related table or query. For example, if you create a simple form that is based on the Employees table, and there is a one-to-many relationship that is defined between the Employees table and Orders table, the datasheet displays all the records in the Orders table that relate to the current Employee record. You can delete the datasheet from the form if you decide you do not need it. If there is more than one table with a one-to-many relationship to the table that you used to create the form, Access does not add any datasheets to the form.

Creating a Datasheet Auto Form

1) Make sure your Student List database is open.
2) Select the Forms section from the Database Window.
3) Click the New button at the top of the Database Window.
4) When the New Form dialog appears, click on the list at the bottom as shown below. A list of all your tables and queries appears since forms can be based on either.
5) Select your *STUDENTS* table from the list (it's easy to tell which one's a table because we named tables in uppercase and queries in lowercase).



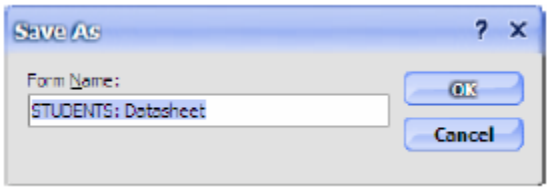
6) From the list of options at the top of the dialog, select AutoForm: Datasheet.

7) Click OK to create the form.

The end result is a form that looks and acts the same as a table. It may not seem very useful to have a form that's the same as a table but it can be very useful for *subforms* as you will see later on. Sometimes it is handy to have a list inside a form. Since a table can't be placed inside a form, another form that looks like a table can be used instead.

8) Close the form.

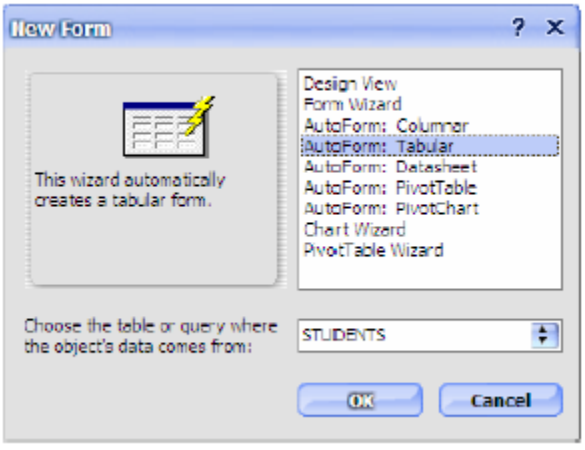
9) When you are prompted to save the form, click Yes.



10) Enter *STUDENTS: Datasheet* as the form name and click OK.

Creating a Tabular Auto form

1) Click the New button at the top of the Database Window.



2) When the New Form dialog appears, select the options shown above (*STUDENTS* as the table and *AutoForm: Tabular* as the type of form).

3) Click OK to create the form.

4) Close the form.

5) When you are prompted to save the form, click Yes.

6) Enter *STUDENTS: Tabular* as the form name and click OK.

A tabular form can be used in the same was as a table, but it can be formatted and customized a lot more than a table. This form could be neatened up in Design View.

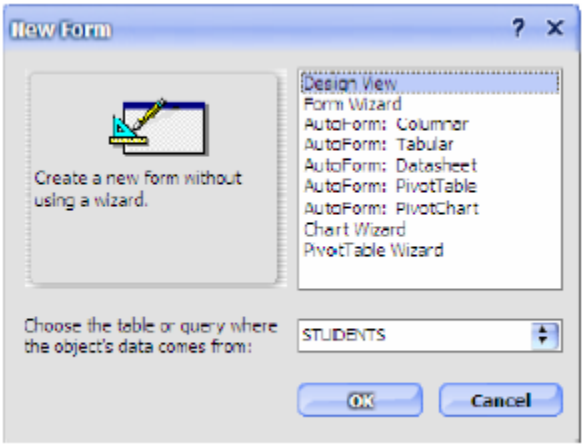
Custom Forms

Using the Form Wizard and form Design View allow you to have a lot more control over how your form looks and functions. These methods allow you to select fields from more than one table/query as well as giving you a lot of choice about how the form will appear.

Using a Form Wizard

1) Click the New button at the top of the Database Window.

2) Choose Form Wizard from the list of options. Choose *STUDENTS* for the table.



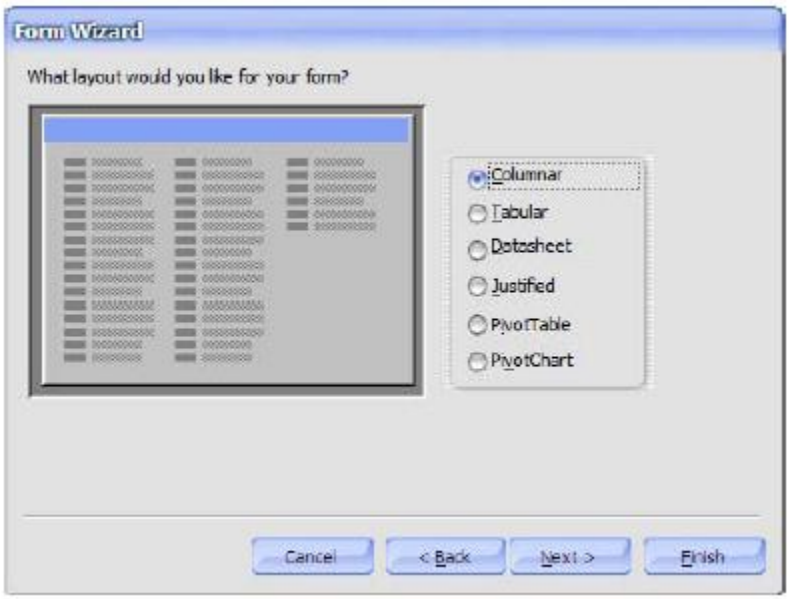
3) Click OK to begin the wizard.

In the first step of the wizard, you are asked to specify which fields will be used in the form.

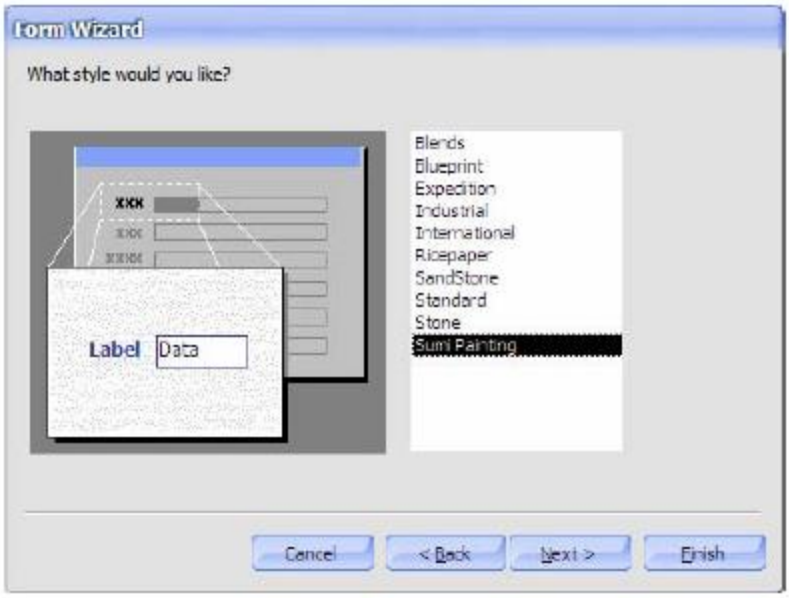
4) Click the  button to select all of the fields for use in the form. All of the fields will now be listed on the right side.

5) Double-click on *Student Number* to move it back over to the left, since we won't need to have that displayed in our form.

6) Click Next to move to the next step of the wizard.

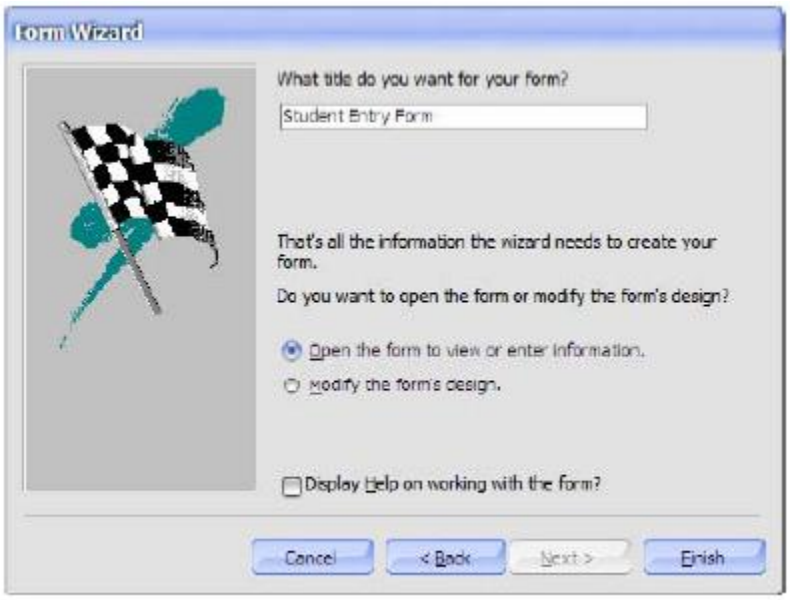


7) Leave Columnar selected for the Form Layout and click Next.



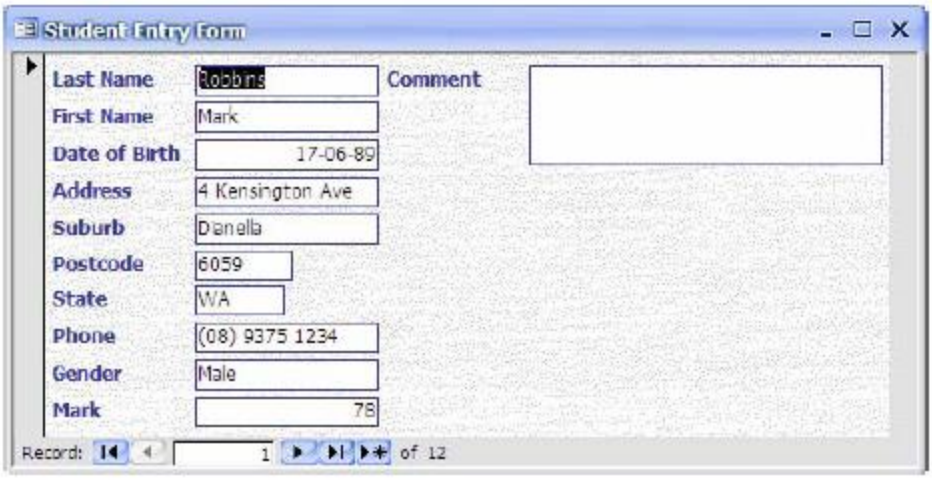
8) Click on each of the different form styles to see the preview for each one.

9) Select a style you like and click Next.



The last step in the wizard asks you to specify a name for the form. When you click Finish, the form is automatically saved with that name.

10) Type *Student Entry Form* for the form title and click Finish.



11) Close the Form. There is no need to save it since it was saved at the end of the wizard.

Reports

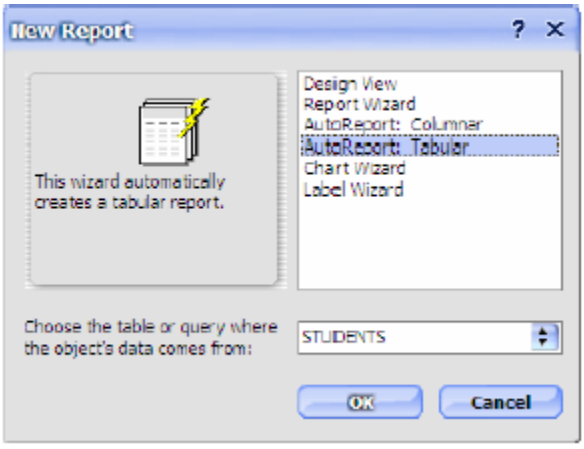
Reports are used in a database to present information in a neat and organised format that is ready for printing. When a report is opened in Access, it is opened in Print preview for this reason. Creating a report is very similar to creating a form and like a form, can be done using any of the following methods:

- Use an Auto report to create a form based on a standard layout.
- Use the Report Wizard.
- Use one of the above methods and then modify the report in Design View.
- Create a report completely from scratch using the Design View tools.

Creating a Tabular Auto Report

1) Make sure you are in the forms section of the Database Window.

2) Click the New button at the top of the Database Window.



3) Select AutoReport: Tabular and make sure the *STUDENTS* table is selected as the source.

4) Click OK to create the report. The report will appear in Print Preview ready for printing.

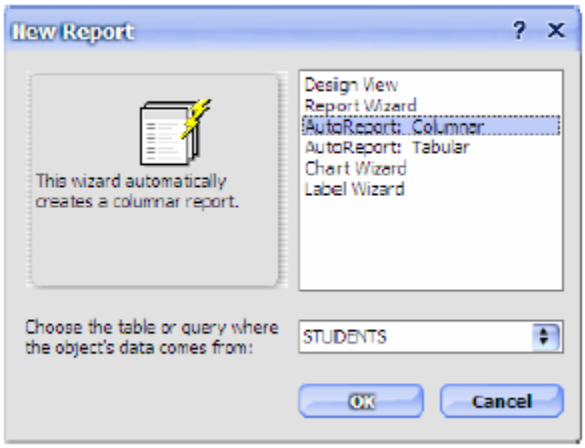
STUDENTS

t Number	Last Name	First Name	e of Birth	Address	Suburb	Postc	State	Phone	Gender
1	Robbins	Mark	17-06-89	4 Kensington A	Dianella	6059	WA	(08) 9375 1234	Male
2	Stevens	Sarah	10-04-89	24 Browne Ave	Yokine	6060	WA	(08) 9249 6127	Female
3	Andrews	Claire	01-11-89	322 Walter Rd	Morley	6059	WA	(08) 9275 1937	Female
4	McKay	Tim	02-08-89	54 Coode St	Dianella	6059	WA	(08) 9375 5610	Male

5) Close the report. When prompted, save the report as *Student Report: Tabular*.

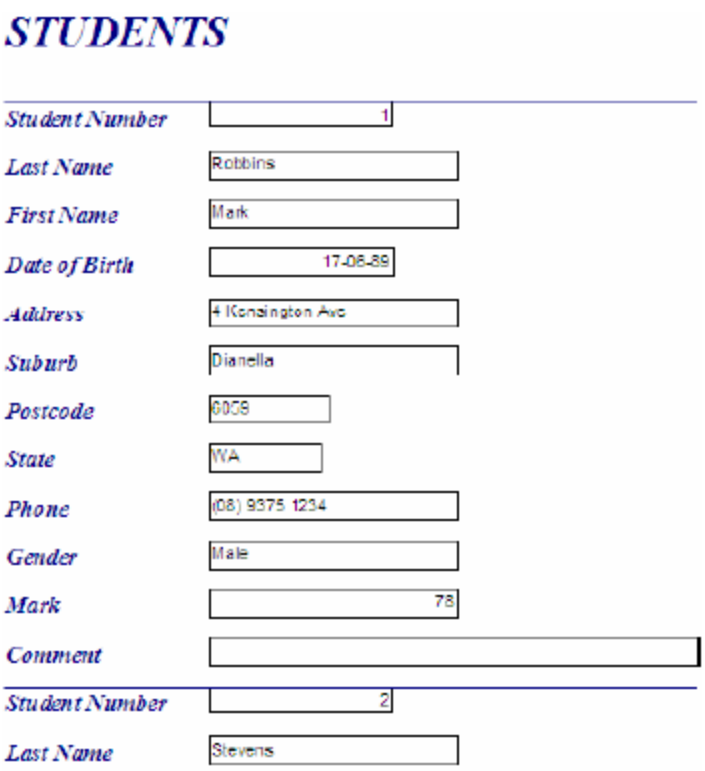
Creating a Columnar Auto Report

1) Click the New button at the top of the Database Window.



2) Select AutoReport: Tabular and make sure the *STUDENTS* table is selected as the source.

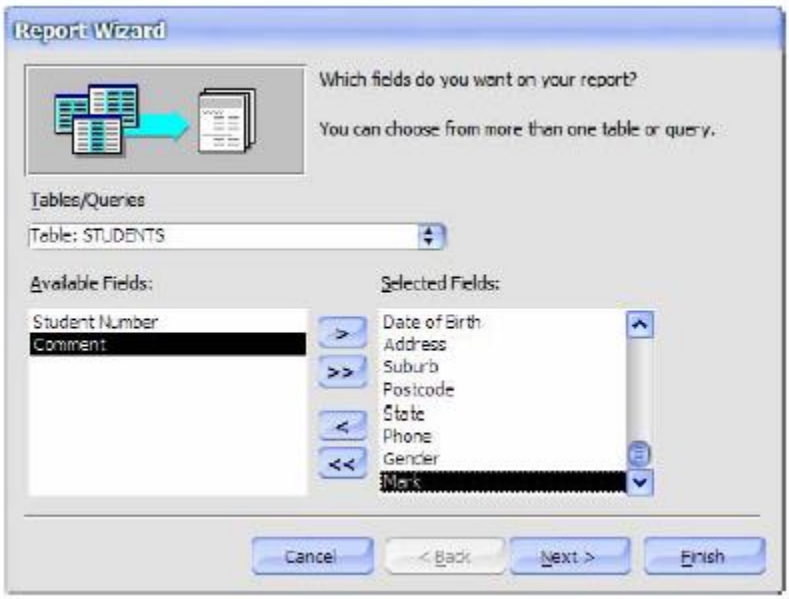
3) Click OK to create the report. The report will appear in Print Preview ready for printing.



4) Close the report. When prompted, save the report as *Student Report: Columnar*.

Using a Report Wizard

1) From the Database Window, click the option that says Create report by using wizard

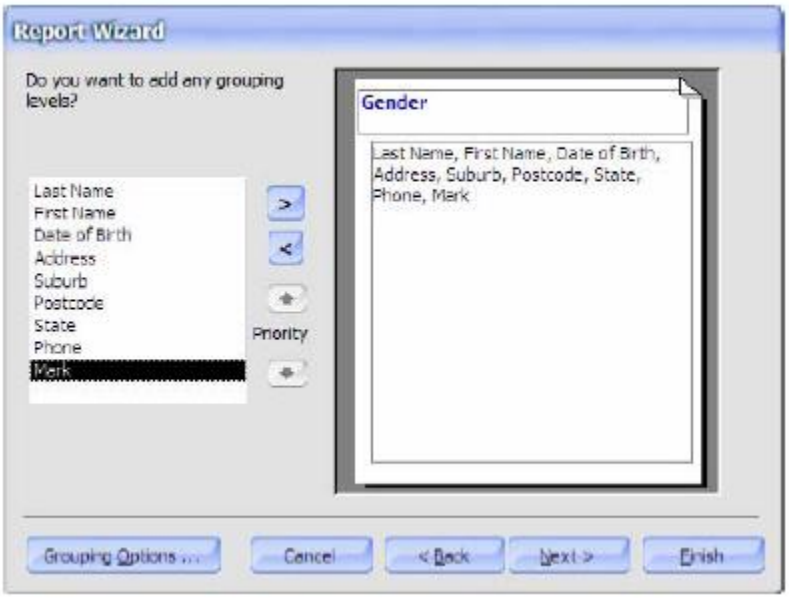


2) In the Tables/Queries list, make sure that *table: STUDENTS* is selected.

3) Click the >> button to select all of the fields for use in the form. All of the fields will now be listed on the right side.

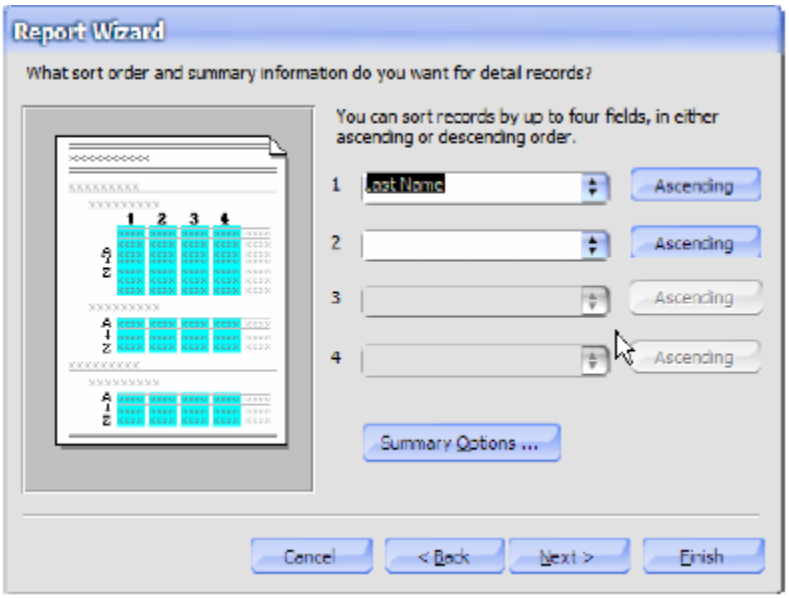
4) Double-click on *Student Number* and *Comment* to move them back over to the left as shown above.

5) Click Next when ready.



6) The next step allows you to choose grouping levels for your report. Double click on Gender to select that as the grouping field. This means that all of the female students will be grouped together in the report and all the male students will be grouped together.

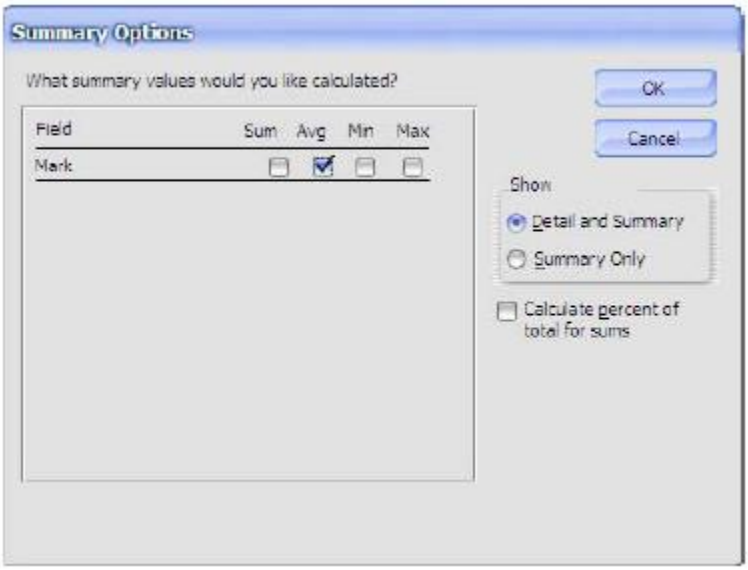
7) Click Next to continue.



This step in the wizard allows you to choose how the records in the report will be sorted. There are also options for adding totals and subtotals to your report.

8) In the first sort box, select Last Name as shown above. You can also select additional fields for sorting in case there are any records with the same last name.

9) Click the Summary Options button.

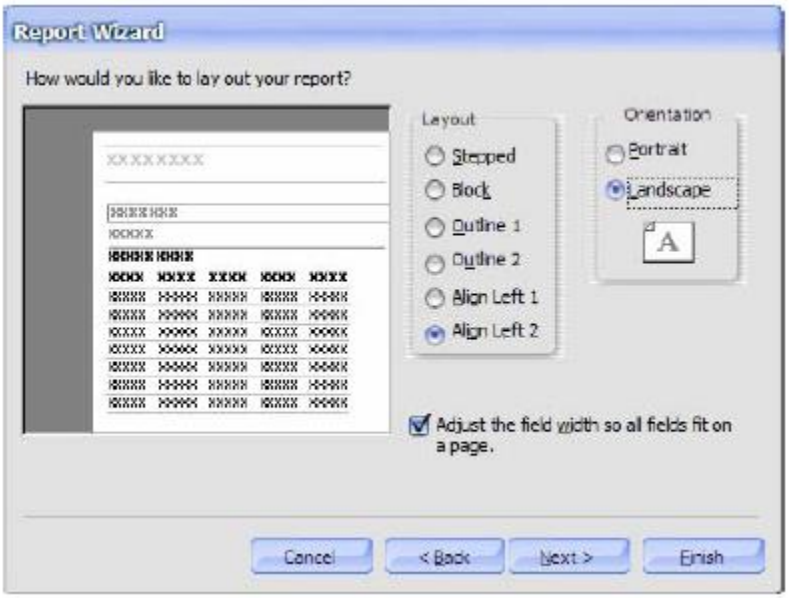


You can use these options to add totals for any number fields. The only field available here is *Mark*.

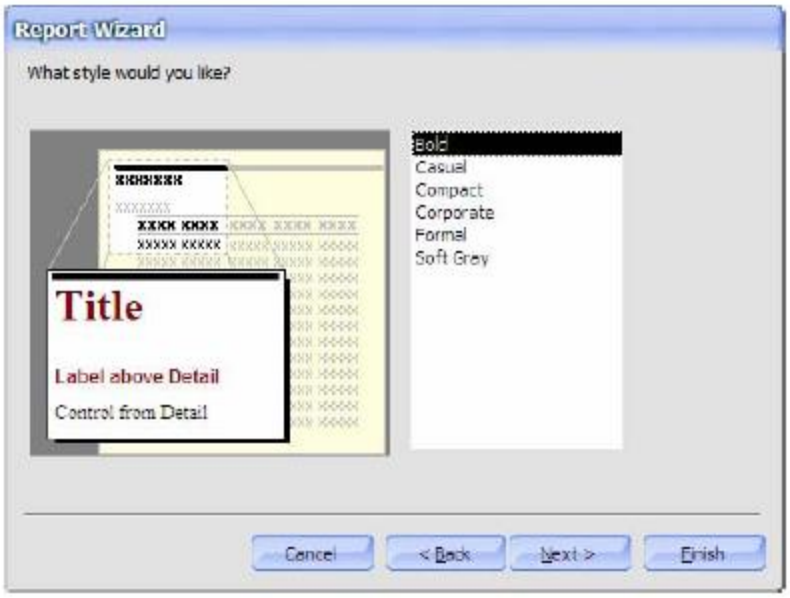
10) Click the box to put a tick under the Avg option as shown above. This will add an average mark figure to the report.

11) Click OK to return to the wizard.

12) Click Next to move to the next step.

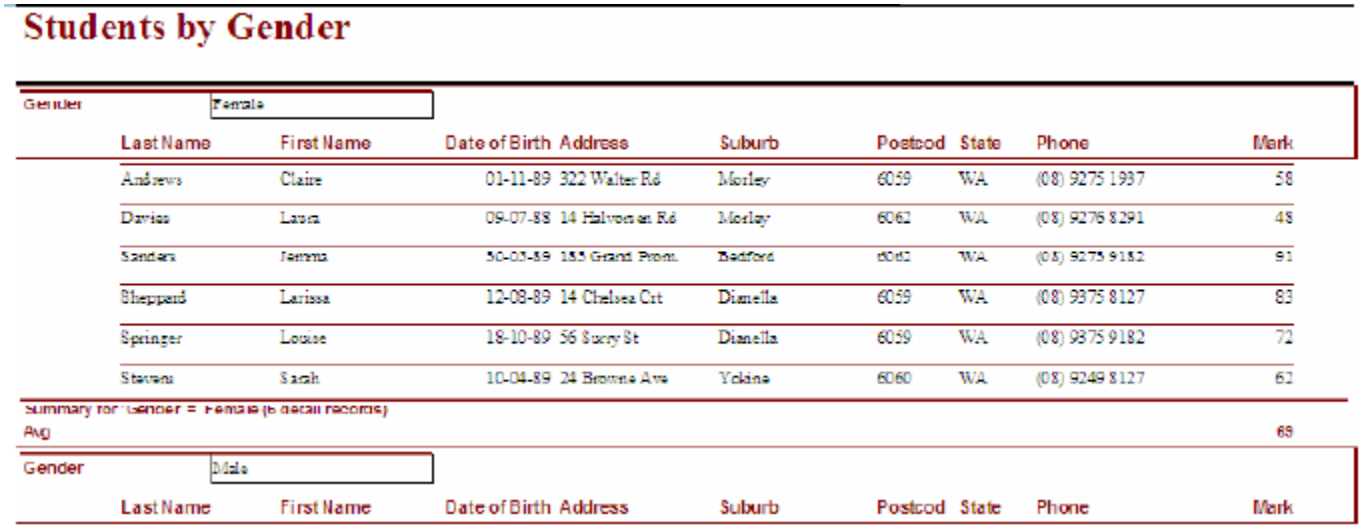


13) Click Align Left 2 for the report layout with Landscape selected as the orientation and click Next.



14) Select a report style and click Next.

15) Enter *Students by Gender* for the report name and click Finish



Students by Gender

Gender: Female

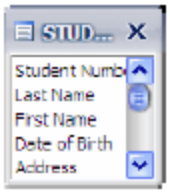
Last Name	First Name	Date of Birth	Address	Suburb	Postcod	State	Phone	Mark
Andrews	Claire	01-11-89	322 Walter Rd	Morley	6059	WA	(08) 9275 1937	58
Davies	Laura	09-07-88	14 Halvorsen Rd	Morley	6062	WA	(08) 9276 8291	48
Sanders	Serena	30-03-89	155 Grand Prom.	Bedford	6062	WA	(08) 9275 9182	91
Sheppard	Larissa	12-08-89	14 Chelsea Crt	Dianella	6059	WA	(08) 9375 8127	83
Springer	Louise	18-10-89	56 Surry St	Dianella	6059	WA	(08) 9375 9182	72
Stevens	Sarah	10-04-89	24 Bromus Ave	Yokine	6060	WA	(08) 9249 8127	62
Summary for 'Gender = Female (6 detail records)								
Avg								69

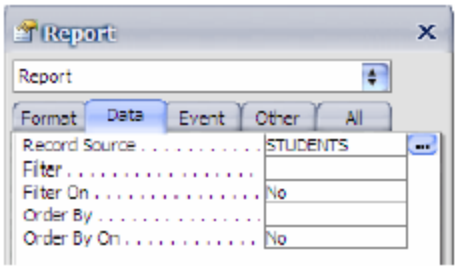
Gender: Male

Last Name	First Name	Date of Birth	Address	Suburb	Postcod	State	Phone	Mark

Report Design

Several things may appear in the Report Design window. Some may appear to begin with and others may need to be turned on before they will appear.





Field List 

Turn on or off with the icon or with the Field List option in the View menu

Shows the fields in the table/query the report is linked to. This can be used to add more fields to the report if needed.

Toolbox 

Turn on or off with the icon or with the Toolbox option in the View menu

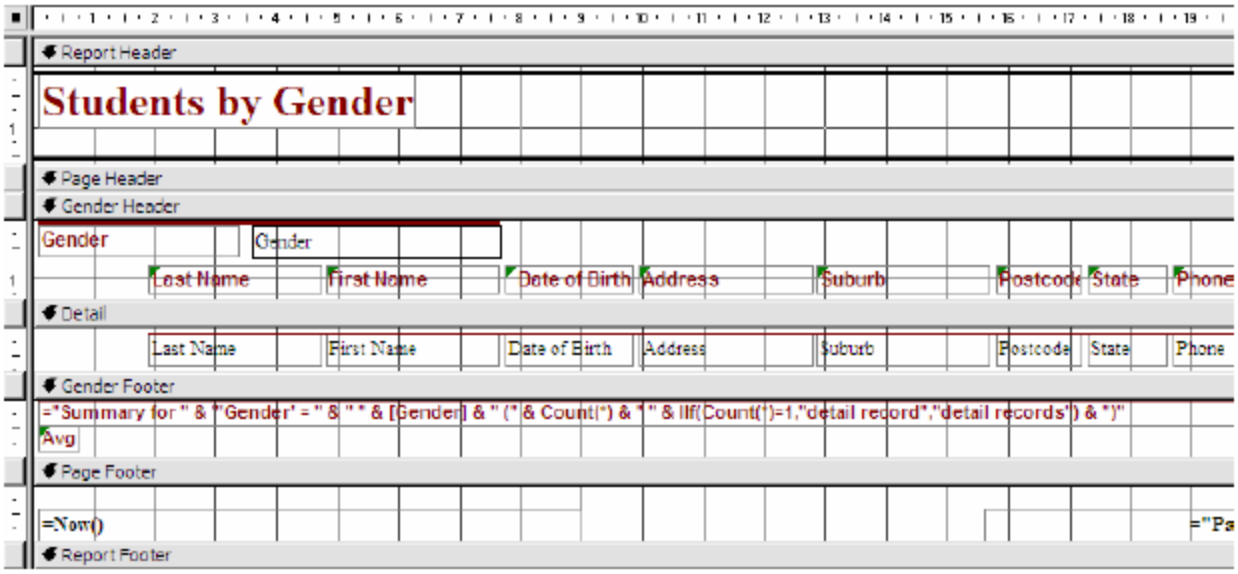
This toolbar has icons for the common Report Design tools. It usually sits towards the side of the screen and can be attached to the side of the screen so it's not in the way.

Properties Window 

Turn on or off with the icon or with the Properties option in the View menu

This shows properties that can be modified for the report or for whichever part of the form is currently selected.

The design window itself has several sections as shown below.



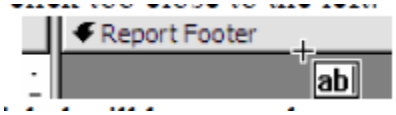
Rulers	Placed along the top and side of the form to help you position and select objects on the report.
Report Header	The contents of this section will be displayed at the beginning of the report.
Page Header	This section is displayed at the top of each page in the report.
Gender Header	When you have used a field to group the results, this area includes everything that will appear before each group.
Detail	This is the section of the report that is repeated for each record.
Gender Footer	Summary information for each grouped field. In this case, the average for each gender will be displayed here.
Page Footer	Repeated at the bottom of each page. In this case, the current date and page number are displayed.
Report Footer	Displayed at the end of the report and usually contains totals for the report.

Customising a Report in Report Design

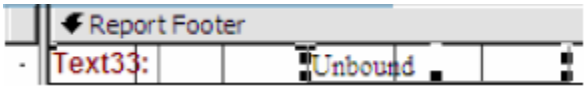
1) With the *Students by Gender* report still open, click the View icon to go to Design View.

2) Click on the Text Box icon in the toolbox to the left of the screen.

3) Click in the Report Footer area to place the textbox as shown. A textbox will usually have a label next to it so make sure you don't click too close to the left.



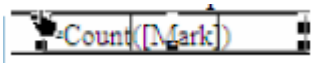
4) When you click, a textbox and a label will be created.



5) Click in the label (the one that says *Text33:*). Edit the label so that it says *Total Students*.

6) Edit the textbox (the one that says *unbound*) and change it so it says *=count([Mark])*.

This is a function that will count all of the records using the *Mark* field. The square brackets are used because they are necessary for any formula that refers to a field name.

7) If you are still editing the textbox, press [Enter] to exit edit mode. The whole box should be selected. Move your mouse over the top-left corner of the box. Your mouse pointer will change to a hand shape as shown. 

8) Use this corner to drag the textbox towards the right edge of the report.

Normally dragging a textbox will move it's related label as well. Using this corner will move only the textbox.

Note When moving and resizing in a report, be careful not to go too close to the edge of the report. The wizard will make the report just the right width to fit on a page. If you move things beyond the edge, even by a small amount, your report will become bigger and won't fit on a page. The result is that when the report is printed is that you may get what appears like blank pages between every page in the report as the small overflow gets printed.

9) Click the view icon to preview the changes to the report and see your new total.

10) Close the report and save any changes.

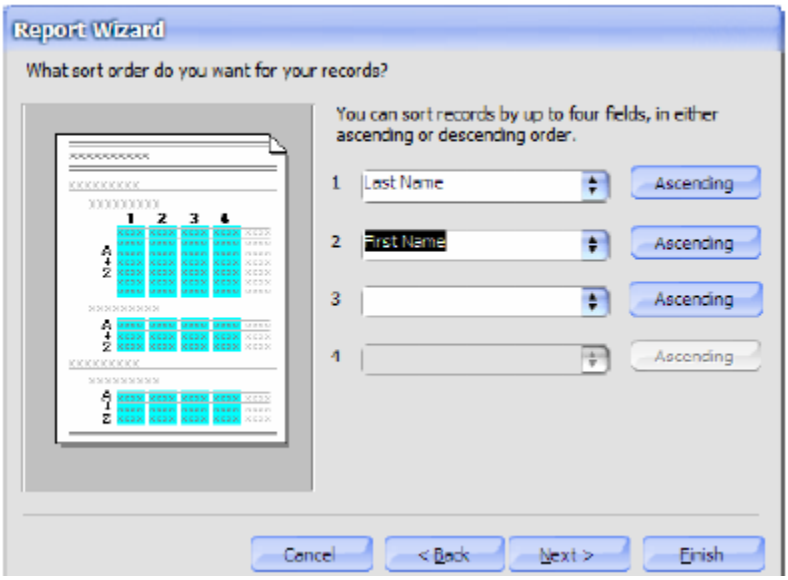
Creating a Report from a Query

1) From the Database Window, click the option that says Create report by using wizard.

2) From the Tables/Queries list, select *Query: parameter: students by suburb* as shown.

3) Click the icon to select all of the fields in the query and then click Next.

4) Click Next again to skip the grouping options.



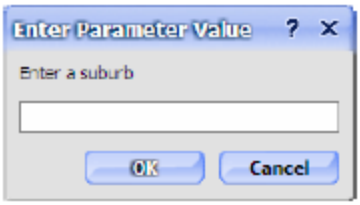
5) Choose to sort by *Last name* and *First name* as shown and click Next.

6) Leave the report layout as Tabular and click Next.

7) Select a report style and click Next.

8) For the name of the report, enter *Students by Suburb* and click Finish.

Because the report is based on a parameter query, the parameter criteria prompt will appear.



9) Enter Morley as the suburb and click OK to see the report.

10) Click the View icon to enter Design view.

The Report Header section contains a label with the text *Students by Suburb*. We will change this in to a textbox that shows the name of the suburb being displayed in the report results.

11) Click the *Students by Suburb* label to select it.

12) From the Format menu choose Change To and then Text Box as shown to the right. When the label changes to a textbox, its contents will change to *Unbound*, which just means that we haven't specified what will appear in the textbox yet. In the textbo

="Student Listing for " & [Suburb]T

13) x enter the following his is a formula that will take the text inside the quotation ze the textbox to make it wider. 5) Click the view icon

marks (including the space after the word *for*) and then the *&* character will join it to the contents of the Suburb Field (since we are using the query to only show the results for one suburb.

14) Re-size the textbox to make it wider.

15) Click the view icon to preview the report.

16) Enter Morley for the suburb again. The modified report header should now include the suburb as shown below

Student Listing for Morley

Last Name	First Name	Address	Suburb	Postcode	State
Andrews	Claire	322 Walter Rd	Morley	6059	WA
Chang	Paul	89 Wellington Rd	Morley	6062	WA
Davies	Laura	14 Halvorsen Rd	Morley	6062	WA

17) Close the report and save the changes.

18) Double-click the report to test it again with a different suburb (such as Yokine or Bedford).

19) Close the report again when you are done.

5.2 Developing Database Applications and DSS

A **database** is a collection of objects that allow user to store data, organize it and retrieve it in any way user wants. What this means is that, with MS Access user create structures called **tables** that allow user to organize the data so that it's easy to find later, user create **forms** that let user input the data into the tables and then you create **reports** that print selected information from the tables. For example, if user run a store, user would create a Customers table, a Products table and an Invoices table. Then, when user open an account for a new customer users would have a Customer form to input a customer's data into the Customers table and an Order form to input the purchase information. Later, user can print any number of Sales reports, grouping and arranging the information from the Invoices, Customers and Products tables to analyze daily or weekly or monthly sales in all kinds of combinations.

5.3Excel Data Analysis Tools

The various MS-Excel data analysis tools are:

1 Sort: You can sort your Excel data on one column or multiple columns. You can sort in ascending or descending order.

2 Filter: Filter your Excel data if you only want to display records that meet certain criteria.

3 Conditional Formatting: Conditional formatting in Excel enables you to highlight cells with a certain color, depending on the cell's value.

4 Charts: A simple Excel chart can say more than a sheet full of numbers. As you'll see, creating charts is very easy.

5 Pivot Tables: Pivot tables are one of Excel's most powerful features. A pivot table allows you to extract the significance from a large, detailed data set.

6 Tables: Tables allow you to analyze your data in Excel quickly and easily.

7 What-If Analysis: What-If Analysis in Excel allows you to try out different values (scenarios) for formulas.

8 Solver: Excel includes a tool called solver that uses techniques from the operations research to find optimal solutions for all kind of decision problems.

9 Analysis ToolPak: The Analysis ToolPak is an Excel add-in program that provides data analysis tools for financial, statistical and engineering data analysis.

5.4 Goal seeking

In computing, **goal seeking** is the ability to calculate backward to obtain an input that would result in a given output. This can also be called **what-if analysis** or **back-solving**. It can either be attempted through trial and improvement or more logical means. Basic goal seeking functionality is built into most modern spreadsheet packages such as Microsoft Excel.

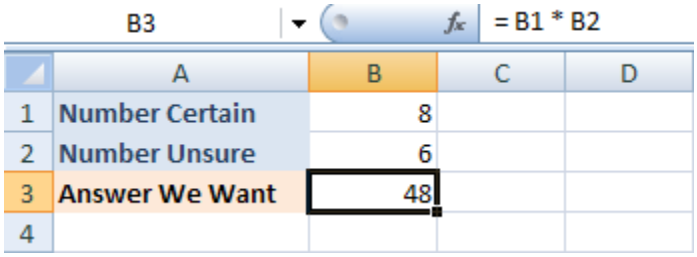
According to O'Brien and Marakas, **optimization analysis** is a more complex extension of goal-seeking analysis. Instead of setting a specific target value for a variable, the goal is to find the optimum value for one or more target variables, given certain constraints. Then one or more other variables are changed repeatedly, subject to the specified constraints, until you discover the best values for the target variables

Examples:

Suppose a family wanted to take out the biggest loan that they could afford to pay for. If they set aside $500 a month, the goal-seeking program would try to work out how big a loan the family could afford to take out. Even using simple trial and improvement, a computer could quickly determine that they could not afford $50,000 loan, but could afford a $48,000 loan. It would then repeat the process until it had reached a figure such as $48,476.34, which would give them a monthly repayment as close to, but not exceeding, $500, as possible.

A more efficient method, especially on more complicated calculations, would be for the program to logically work through the argument. By drawing up a simple equation, the program could come to the Summary that the output equalled one ninety-sixth of the input, and could then multiply the output (or goal) by ninety-six to find the necessary input.

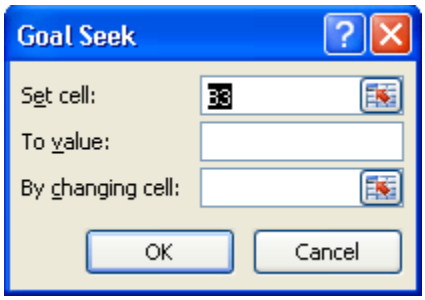
How to Create Excel spreadsheet and Goal Seeking



In the spreadsheet above, we know that we want to multiply the number in B1 by the number in B2. The number in cell B2 is the one we're not too sure of. The answer is going in cell B3. Our answer is wrong at the moment, because we has a Goal of 56. To use Goal Seek to get the answer, try the following:

- From the Excel menu bar, click on **Data**

- Locate the **Data Tools** panel and the **What if Analysis** item. From the What if Analysis menu, select **Goal Seek**
- The following dialogue box appears:



The first thing Excel is looking for is "Set cell". This is not very well named. It means "Which cell contains the Formula that you want Excel to use". For us, this is cell B3. We have the following formula in B3:

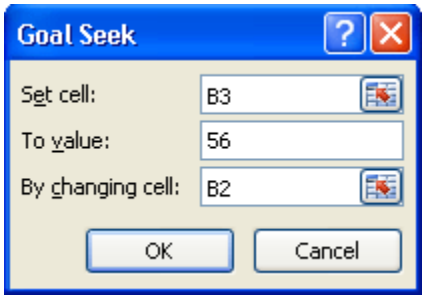
= B1 * B2

So enter B3 into the "Set cell" box, if it's not already in there.

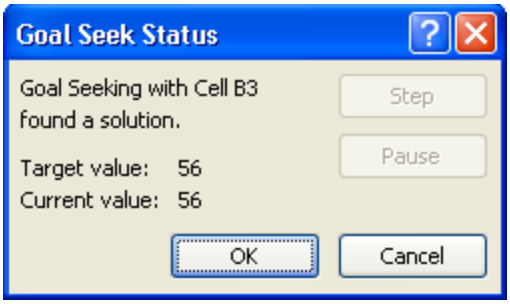
The "To value" box means "What answer are you looking for"? For us, this is 56. So just type 56 into the "To value" box

The "By Changing Cell" is the part you're not sure of. Excel will be changing this part. For us, it was cell B2. We're weren't sure which number, when multiplied by 8, gave the answer 56. So type B2 into the box.

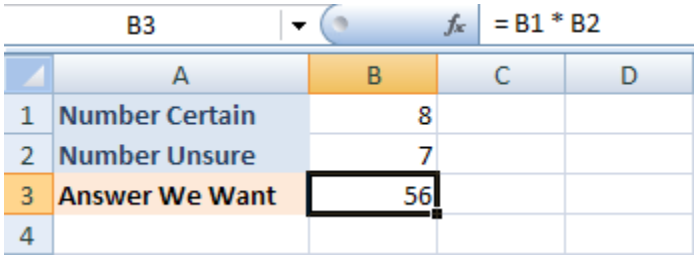
You Goal Seek dialogue box should look like ours below:



Click OK and Excel will tell you if it has found a solution:



Click OK again, because Excel has found the answer. Your new spreadsheet will look like this one:



As you can see, Excel has changed cell B2 and replace the 6 with a 7 - the correct answer.

5.5 Sensitivity Analysis

Sensitivity analysis is the study of how the uncertainty in the output of a mathematical model or system (numerical or otherwise) can be apportioned to different sources of uncertainty in its inputs. A related practice is uncertainty analysis, which has a greater focus on uncertainty quantification and propagation of uncertainty. Ideally, uncertainty and sensitivity analysis should be run in tandem.

Sensitivity analysis can be useful for a range of purposes including

- Testing the robustness of the results of a model or system in the presence of uncertainty.
- Increased understanding of the relationships between input and output variables in a system or model.
- Uncertainty reduction: identifying model inputs that cause significant uncertainty in the output and should therefore be the focus of attention if the robustness is to be increased (perhaps by further research).
- Searching for errors in the model (by encountering unexpected relationships between inputs and outputs).
- Model simplification – fixing model inputs that have no effect on the output, or identifying and removing redundant parts of the model structure.
- Enhancing communication from modelers to decision makers (e.g. by making recommendations more credible, understandable, compelling or persuasive).

- Finding regions in the space of input factors for which the model output is either maximum or minimum or meets some optimum criterion

The choice of method of sensitivity analysis is typically dictated by a number of problem constraints or settings. Some of the most common are

- **Computational expense:** Sensitivity analysis is almost always performed by running the model a (possibly large) number of times, i.e. a sampling-based approach.[4] This can be a significant problem when,
    - A single run of the model takes a significant amount of time (minutes, hours or longer). This is not unusual with very complex models.
    - The model has a large number of uncertain inputs. Sensitivity analysis is essentially the exploration of the multidimensional input space, which grows exponentially in size with the number of inputs. See the curse of dimensionality.

    Computational expense is a problem in many practical sensitivity analyses. Some methods of reducing computational expense include the use of emulators (for large models), and screening methods (for reducing the dimensionality of the problem). Another method is to use an event-based sensitivity analysis method for variable selection for time-constrained applications. This is an input variable selection method that assembles together information about the trace of the changes in system inputs and outputs using sensitivity analysis to produce an input/output trigger/event matrix that is designed to map the relationships between input data as causes that trigger events and the output data that describes the actual events. The cause-effect relationship between the causes of state change i.e. input variables and the effect system output parameters determines which set of inputs have a genuine impact on a given output. The method has a clear advantage over analytical and computational IVS method since it tries to understand and interpret system state change in the shortest possible time with minimum computational overhead.

- **Correlated inputs:** Most common sensitivity analysis methods assume independence between model inputs, but sometimes inputs can be strongly correlated. This is still an immature field of research and definitive methods have yet to be established.
- **Nonlinearity:** Some sensitivity analysis approaches, such as those based on linear regression, can inaccurately measure sensitivity when the model response is nonlinear with respect to its inputs. In such cases, variance-based measures are more appropriate.
- **Model interactions:** Interactions occur when the perturbation of two or more inputs *simultaneously* causes variation in the output greater than that of varying each of the inputs alone. Such interactions

are present in any model that is non-additive, but will be neglected by methods such as scatterplots and one-at-a-time perturbations.[7] The effect of interactions can be measured by the total-order sensitivity index.

- **Multiple outputs:** Virtually all sensitivity analysis methods consider a single univariate model output, yet many models output a large number of possibly spatially or time-dependent data. Note that this does not preclude the possibility of performing different sensitivity analyses for each output of interest. However, for models in which the outputs are correlated, the sensitivity measures can be hard to interpret.
- **Given data:** While in many cases the practitioner has access to the model, in some instances a sensitivity analysis must be performed with "given data", i.e. where the sample points (the values of the model inputs for each run) cannot be chosen by the analyst. This may occur when a sensitivity analysis has to be performed retrospectively, perhaps using data from an optimisation or uncertainty analysis, or when data comes from a discrete source.

SELF ASSESSMENT QUESTIONS

Please select the correct option:

A. Which of the following statements describes a taxonomy?
   1. A browsable hierarchy
   2. A list of equivalent terms
   3. A complex controlled vocabulary showing relationships

B. Which of the following statements about XML schemas is incorrect?
   1. Schemas can specify integer values
   2. Schemas are defined by XSD tag
   3. All XML documents must have a schema
   4. Schemas provide data oriented data types
   5. They offer more flexibility than DTDs

C. Which of the following relates to enterprise interoperability?
   1. DFD
   2. Information flow diagram
   3. XML
   4. Entity relationship diagram

D. Which of the following is not a type of navigation system for a web site?

1. National
2. Global
3. Contextual
4. Local
5. Regional

E. What should not be part of an acceptable use policy?

1. Allowable use of systems
2. Legal obligations
3. Encryption policies
4. User responsibilities
5. Account and password responsibilities

F. Use-case analysis focuses upon: 1. data 2. Objects 3. Entities 4. Actors

5.6 Filtering

A **filter** is a computer program or subroutine to process a stream, producing another stream. While a single filter can be used individually, they are frequently strung together to form a pipeline. Some operating systems such as Unix are rich with filter programs. Windows 7 and later are also rich with filters, as they include Windows PowerShell. In comparison, however, few filters are built into cmd.exe (the original command-line interface of Windows), most of which have significant enhancements relative to the similar filter commands that were available in MS-DOS. OS X includes filters from its underlying Unix base but also has Automator, which allows filters (known as "Actions") to be strung together to form a pipeline.

Apply a **filter** to view select records in an **Access** database. **Filtering** is a useful way to see only the data that you want displayed. F**ilters** can be used to display specific records in a form, report, query, or datasheet, or to print only certain records from a report, table, or query.

To create a filter from a selection:

1. Select the cell or data you want to create a filter with. ...
2. Select the Home tab on the Ribbon, locate the Sort & Filter group, and click the Selection drop-down arrow. ...
3. Select the type of filter you want to apply: ...
4. The filter will be applied.

Steps to filter in access

This option tells **Access** to show records except the one that was highlighted. User can see that the **toggle filter** button is now highlighted. User can use the button to **toggle** the last **filter** on and off. Go back to the city field again, click on the arrow and select the "clear **filter** from city" option.

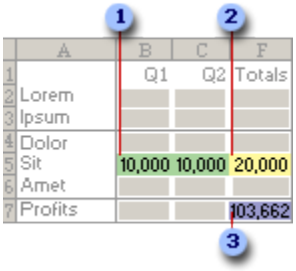
5.7 Solver

A **solver** is a generic term indicating a piece of mathematical software, possibly in the form of a stand-alone computer program or as a software library, that 'solves' a mathematical problem. A solver takes problem descriptions in some sort of generic form and calculate their solution. In a solver, the emphasis is on creating a program or library that can easily be applied to other problems of similar type.

Solver is part of a suite of commands sometimes called what-if analysis tools. With Solver, user can find an optimal (maximum or minimum) value for a formula in one cell — called the objective cell — subject to constraints, or limits, on the values of other formula cells on a worksheet. Solver works with a group of cells, called decision variables or simply variable cells, that participate in computing the formulas in the objective and constraint cells. Solver adjusts the values in the decision variable cells to satisfy the limits on constraint cells and produce the result you want for the objective cell.

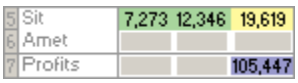
Example of a Solver evaluation

In the following example, the level of advertising in each quarter affects the number of units sold, indirectly determining the amount of sales revenue, the associated expenses, and the profit. Solver can change the quarterly budgets for advertising (decision variable cells B5:C5), up to a total budget constraint of $20,000 (cell F5), until the total profit (objective cell F7) reaches the maximum possible amount. The values in the variable cells are used to calculate the profit for each quarter, so they are related to the formula objective cell F7, =SUM(Q1 Profit:Q2 Profit).



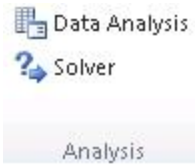
1. Variable cells
2. Constrained cell
3. Objective cell

After Solver runs, the new values are as follows.



Define and solve a problem

1. On the **Data** tab, in the **Analysis** group, click **Solver**.



If the **Solver** command or the **Analysis** group is not available, you need to load the Solver Add-in program.

How to load the Solver Add-in program

a. Click the **File** tab, click **Options**, and then click the **Add-Ins** category.

b. In the **Manage** box, click **Excel Add-ins**, and then click **Go**.

c. In the **Add-ins available** box, select the **Solver Add-in** check box, and then click **OK**.

2. In the **Set Objective** box, enter a cell reference or name for the objective cell. The objective cell must contain a formula.

3. Do one of the following:

a. If you want the value of the objective cell to be as large as possible, click **Max**.

b. If you want the value of the objective cell to be as small as possible, click **Min**.

c. If you want the objective cell to be a certain value, click **Value of**, and then type the value in the box.

4. In the **By Changing Variable Cells** box, enter a name or reference for each decision variable cell range. Separate the nonadjacent references with commas. The variable cells must be related directly or indirectly to the objective cell. You can specify up to 200 variable cells.

5. In the **Subject to the Constraints** box, enter any constraints that you want to apply by doing the following:

a. In the **Solver Parameters** dialog box, click **Add**.

b. In the **Cell Reference** box, enter the cell reference or name of the cell range for which you want to constrain the value.

c. Click the relationship (<=, =, >=, **int**, **bin**, or **dif**) that you want between the referenced cell and the constraint.

d. If you click **int**, **integer** appears in the **Constraint** box. If you click **bin**, **binary** appears in the **Constraint** box. If you click **dif**, **alldifferent** appears in the **Constraint** box.

e. If you choose <=, =, or >= for the relationship in the **Constraint** box, type a number, a cell reference or name, or a formula.

f. Do one of the following:

- To accept the constraint and add another, click **Add**.
- To accept the constraint and return to the **Solver Parameters** dialog box, click **OK**.

Note You can apply the **int**, **bin**, and **dif** relationships only in constraints on decision variable cells.

You can change or delete an existing constraint by doing the following:

g. In the **Solver Parameters** dialog box, click the constraint that you want to change or delete.

h. Click **Change** and then make your changes, or click **Delete**.

6. Click **Solve** and do one of the following:

a. To keep the solution values on the worksheet, in the **Solver Results** dialog box, click **Keep Solver Solution**.

b. To restore the original values before you clicked **Solve**, click **Restore Original Values**.

Notes

c. You can interrupt the solution process by pressing ESC. Microsoft Excel recalculates the worksheet with the last values that are found for the decision variable cells.

d. To create a report that is based on your solution after Solver finds a solution, you can click a report type in the **Reports** box and then click **OK**. The report is created on a new worksheet in your workbook. If Solver doesn't find a solution, only certain reports or no reports are available.

e. To save your decision variable cell values as a scenario that you can display later, click **Save Scenario**in the **Solver Results** dialog box, and then type a name for the scenario in the **Scenario Name** box.

Step through Solver trial solutions

1. After you define a problem, click **Options** in the **Solver Parameters** dialog box.
2. In the **Options** dialog box, select the **Show Iteration Results** check box to see the values of each trial solution, and then click **OK**.
3. In the **Solver Parameters** dialog box, click **Solve**.
4. In the **Show Trial Solution** dialog box, do one of the following:

- To stop the solution process and display the **Solver Results** dialog box, click **Stop**.
- To continue the solution process and display the next trial solution, click **Continue**.

Types of problems with existing dedicated solvers include:

- Linear and non-linear equations. In the case of a single equation, the "solver" is more appropriately called a root-finding algorithm.
- Systems of linear equations.
- Nonlinear systems.
- Systems of polynomial equations, which are a special case of non linear systems, better solved by specific solvers.

- Linear and non-linear optimization problems
- Systems of ordinary differential equations
- Systems of differential algebraic equations
- Logic/satisfiability problems
- Constraint satisfaction problems
- Shortest path problems
- Minimum spanning tree problems
- Search algorithms

5.8Summary

Creating reports in MS Access

Once the user have opened the database, the follow the steps:

1. **Choose the Reports menu.** Once you've opened Northwind, choose the Create tab on the Microsoft Office ribbon. In the "Reports" selection, you'll see a number of methods that Access supports for creating a report. If you'd like, feel free to click on a few of these and get a feel for what reports look like and the various types of information that they contain.
2. **Create a new report.** After you've satisfied your curiosity, go ahead and click on "Report Wizard" and we'll begin the process of creating a report. The wizard will walk us through the creation process step-by-step. After you've mastered the wizard, you might want to return to this step and explore the flexibility provided by the other creation methods.
3. **Choose a table or query.** The first screen of the Report Wizard asks us to choose the source of data for our report. If you want to retrieve information from a single table, you can select it from the drop-down box below. Alternatively, for more complex reports, we can choose to base our report on the output of a query that we previously designed. For our example, all of the data we need is contained within the Employees table, so choose "Table:Employees" from the drop-down menu.
4. **Select the fields to include.** Notice that after you select the table from the drop-down menu, the bottom section of the screen changes to show the fields available in that table. Use the '>' button to move the fields you would like to include in your report to the "Selected Fields" section. Note that the order you place the fields in the right column determines the default order they will appear in your report. Remember that we're creating an employee telephone directory for our senior management. Let's keep the information contained in it simple -- the first and last name of each employee, their title and their home telephone number. Go ahead and select these fields. When you are satisfied, click the Next Button

5. **Select the grouping levels.** At this stage, you can select one or more grouping levels to refine the order in which our report data is presented. For example, we may wish to break down our telephone directory by department so that all of the members of each department are listed separately. However, due to the small number of employees in our database, this is not necessary for our report. Go ahead and simply click on the Next button to bypass this step. You may wish to return here later and experiment with grouping levels.
6. **Choose your sorting options.** In order to make reports useful, we often want to sort our results by one or more attributes. In the case of our telephone directory, the logical choice is to sort by the last name of each employee in ascending (A-Z) order. Select this attribute from the first drop-down box and then click the Next button to continue.
7. **Choose the formatting options.** In the next screen, we're presented with some formatting options. We'll accept the default tabular layout but let's change the page orientation to landscape to ensure the data fits properly on the page. Once you've completed this, click the Next button to continue.
8. **Add the title.** Finally, we need to give the report a title. Access will automatically provide a nicely formatted title at the top of the screen, with the appearance shown in the report style you selected during the previous step. Let's call our report "Employee Home Phone List." Make sure that the "Preview the report" option is selected and click on Finish to see our report!

5.9Glossary

Goal seeking: **Goal seeking** is the ability to calculate backward to obtain an input that would result in a given output. This can also be called **what-if analysis** or **back-solving**. It can either be attempted through trial and improvement or more logical means. Basic goal seeking functionality is built into most modern spreadsheet packages such as Microsoft Excel.

Filtering : A **filter** is a computer program or subroutine to process a stream, producing another stream. While a single filter can be used individually, they are frequently strung together to form a pipeline. Some operating systems such as Unix are rich with filter programs.

Solver: A **solver** is a generic term indicating a piece of mathematical software, possibly in the form of a stand-alone computer program or as a software library, that 'solves' a mathematical problem. A solver takes problem descriptions in some sort of generic form and calculate their solution. In a solver, the emphasis is on creating a program or library that can easily be applied to other problems of similar type.

Access: A form in Access is a database object that you can use to create a user interface for a database application. A "bound" form is one that is directly connected to a data source such as a table or query, and can be used to enter, edit, or display data from that data source.

5.9 References

1. O'Brien, J (1999). *Management Information Systems – Managing Information Technology in the Internetworked Enterprise*. Boston: Irwin McGraw-Hill. ISBN 0-07-112373-3.
2. Laudon, Kenneth C.; Laudon, Jane P. (2009). *Management Information Systems: Managing the Digital Firm* (11 ed.). Prentice Hall/CourseSmart. p. 164.
3. *Transaction processing systems (TPS)* collect and record the routine transactions of an organization. Examples of such systems are sales order entry, hotel reservations, payroll, employee record keeping, and shipping.
4. Bidgoli, Hossein, (2004). The Internet Encyclopedia, Volume 1, John Wiley & Sons, Inc. p. 707.
5. Pant, S., Hsu, C., (1995), Strategic Information Systems Planning: A Review, Information Resources Management Association International Conference, May 21–24, Atlanta.

5.10 Further Readings

6. Taylor, Victoria. "Supply Chain Management: The Next Big Thing?". *Sept. 12, 2011*. Business Week. Retrieved 5 March 2014.
7. Lynn, Samara. "What is CRM?". PC Mag. Retrieved 5 March 2014.
8. Joshi, Girdhar (2013). *Management Information Systems*. New Delhi: Oxford University Press. p. 328. ISBN 9780198080992.
9. Laudon, K.,&Laudon, J. (2010). Management information systems: Managing the digital firm. (11th ed.). Upper Saddle River, NJ: Pearson Prentice Hall

5.11 **Model Questions**

Q:1 Discuss how will you make forms and reports in MS Access.

Q:2 How will you develop Database Applications.

Q:3 Discuss the various tools of Excel used for data analysis.

Q:5 Discuss in detail Goal Seeking and sensitivity analysis

ANSWERS TO SELF ASSESSMENT QUESTIONS

A. 1

B. 3

C. 3

D. 5

E. 3

F. 4

UNIT –III

CHAPTER 6

Information Systems in Business

Structure

6.0 Objectives

After studying this chapter you will be able :

1. **Discuss about the Decision making and its levels.**
2. **Discuss in detail about TPS, MIS, EIS, DSS, ES and OAS**
3. **Understand about the Collaboration Technologies.**
4. **Understand about various functional areas of Information Systems**

Organizational Systems:

6.1Decision making

Decision-making can be regarded as the cognitive process resulting in the selection of a belief or a course of action among several alternative possibilities. Every decision-making process produces a final choice that may or may not prompt action. Decision-making is the study of identifying and choosing alternatives based on the values and preferences of the decision maker. Decision-making is one of the central activities of management and is a huge part of any process of implementation.

Following elements can be derived from the above mentioned definitions:

1. Decision–making is a selection process and is concerned with selecting the best type of alternative.
2. The decision taken is aimed at achieving the organizational goals.
3. It is concerned with the detailed study of the available alternatives for finding the best possible alternative.
4. Decision making is a mental process. It is the outline of constant thoughtful consideration.
5. It leads to commitment. The commitment depends upon the nature of the decision whether short term or long term.

Features or Characteristics of Decision-Making:

From definitions and elements we can draw the following important features of managerial decisions:

1. Rational Thinking:

It is invariably based on rational thinking. Since the human brain with its ability to learn, remember and relate many complex factors, makes the rationality possible.

2. Process:

It is the process followed by deliberations and reasoning.

3. Selective:

It is selective, i.e. it is the choice of the best course among alternatives. In other words, decision involves selection of the best course from among the available alternative courses that are identified by the decision-maker.

4. Purposive:

It is usually purposive i.e. it relates to the end. The solution to a problem provides an effective means to the desired goal or end.

5. Positive:

Although every decision is usually positive sometimes certain decisions may be negative and may just be a decision not to decide. For instance, the manufacturers of VOX Wagan car once decided not to change the model (body style) and size of the car although the other rival enterprise (i.e. the Ford Corporation) was planning to introduce a new model every year, in the USA.

That a negative decision and is equally important was stressed by Chester I. Bernard-one of the pioneers in Management Thought-who observed, "The fine art of executive decision consists in not deciding questions that are not now pertinent, in not deciding prematurely, in not making decisions that cannot be made effective, and in not making decisions that other should make. "

6. Commitment:

Every decision is based on the concept of commitment. In other words, the Management is committed to every decision it takes for two reasons- viz., (/) it promotes the stability of the concern and (ii) every decision taken becomes a part of the expectations of the people involved in the organisation.

Decisions are usually so much inter-related to the organisational life of an enterprise that any change in one area of activity may change the other areas too. As such, the Manager is committed to decisions not only from the time that they are taken but upto their successfully implementation.

7. Evaluation:

Decision-making involves evaluation in two ways, viz., (i) the executive must evaluate the alternatives, and (ii) he should evaluate the results of the decisions taken by him.

Human performance with regard to decisions has been the subject of active research from several perspectives:

Psychological: examining individual decisions in the context of a set of needs, preferences and values the individual has or seeks.
Cognitive: the decision-making process regarded as a continuous process integrated in the interaction with the environment.
Normative: the analysis of individual decisions concerned with the logic of decision-making and rationality and the invariant choice it leads to.
Decision-making can also be regarded as a problem-solving activity terminated by a solution deemed to be satisfactory. It is, therefore, a reasoning or emotional process which can be rational or irrational and can be based on explicit assumptions or tacit assumptions. Rational choice theory encompasses the notion that people try to maximize benefits while minimizing costs.

Some have argued that most decisions are made unconsciously. Jim Nightingale states that "we simply decide without thinking much about the decision process." In a controlled environment, such as a classroom, instructors might try to encourage students to weigh pros and cons before making a decision. This strategy is known as Franklin's rule. However, because such a rule requires time, cognitive resources and full access to relevant information about the decision, this rule may not best describe how people make decisions.

Logical decision-making is an important part of all science-based professions, where specialists apply their knowledge in a given area to make informed decisions. For example, medical decision-making often involves a diagnosis and the selection of appropriate treatment. Some[which?] research using naturalistic methods shows, however, that in situations with higher time pressure, higher stakes, or increased ambiguities, experts use intuitive decision-making rather than structured approaches – following a recognition primed decision that fits their experience – and arrive at a course of action without weighing alternatives. Recent robust decision research has formally integrated uncertainty into its decision-making model. Decision analysis recognized and included uncertainties in its theorizing since its conception in 1964.

A major part of decision-making involves the analysis of a finite set of alternatives described in terms of evaluative criteria. Information overload occurs when there is a substantial gap between the capacity of information and the ways in which people may or can adapt. The overload of information can be related to problem≠ processing and tasking, which effects decision-making. These criteria may be benefit or cost in nature. Then the problem might be to rank these alternatives in terms of how attractive they are to the decision-maker(s) when all the criteria are considered simultaneously. Another goal might be to just find the best alternative or to determine the relative total priority of each alternative (for instance, if alternatives represent projects competing for funds) when all the criteria are considered simultaneously. Solving such problems is the focus of multi-criteria decision analysis (MCDA), also known as multi-criteria decision-making (MCDM). This area of decision-making, although very old, has attracted the interest of many researchers and practitioners and is still highly debated as there are many MCDA/MCDM methods which may yield very different results when they are applied on exactly the same data. This leads to the formulation of a decision-making paradox.

In regards to management and decision-making, each level of management is responsible for different things. Top level managers look at and create strategic plans where the organization's vision, goals, and values are taken into account to create a plan that is cohesive with the mission statement. For mid-level managers, tactical plans are created with specific steps with actions that need to be executed to meet the strategic objective. Finally, the front-line managers are responsible for creating and executing operational plans. These

plans include the policies, processes, and procedures of the organization. Each must take into account the overall goals and processes of the organization.

The environment can also play a part in the decision making process. It is important to know that environmental complexity is a factor that influences cognitive function and well being. A complex environment is an environment with a large number of different possible states which come and go over time. It is in different states at different times and different in different places as opposed to the same all over. Peter Godfrey-Smith, professor at Stamford University, states "whether a particular type of complexity is relevant to an organism depends on what the organism is like- size, needs, habits and physiology."

Decision making in Management:

Decision-making is an integral part of modern management. Essentially, Rational or sound decision making is taken as primary function of management. Every manager takes hundreds and hundreds of decisions subconsciously or consciously making it as the key component in the role of a manager. Decisions play important roles as they determine both organizational and managerial activities. A decision can be defined as a course of action purposely chosen from a set of alternatives to achieve organizational or managerial objectives or goals. Decision making process is continuous and indispensable component of managing any organization or business activities. Decisions are made to sustain the activities of all business activities and organizational functioning.

Figure: Decision making at Organizations

Decisions are made at every level of management to ensure organizational or business goals are achieved. Further, the decisions make up one of core functional values that every organization adopts and implements to ensure optimum growth and drivability in terms of services and or products offered.

Trewatha & Newport defines decision making process as follows:, "**Decision-making involves the selection of a course of action from among two or more possible alternatives in order to arrive at a solution for a given problem"**.

As evidenced by the foregone definitions, decision making process is a consultative affair done by a comity of professionals to drive better functioning of any organization. Thereby, it is a continuous and dynamic activity that pervades all other activities pertaining to the organization. Since it is an ongoing activity, decision making process plays vital importance in the functioning of an organization. Since intellectual minds are involved in the process of decision making, it requires solid scientific knowledge coupled with skills and experience in addition to mental maturity.

Further, decision making process can be regarded as check and balance system that keeps the organisation growing both in vertical and linear directions. It means that decision making process seeks a goal. The goals are pre-set business objectives, company missions and its vision. To achieve these goals, company may face lot of obstacles in administrative, operational, marketing wings and operational domains. Such problems are sorted out through comprehensive decision making process. No decision comes as end in itself, since in may evolve new problems to solve. When one problem is solved another arises and so on, such that decision making process, as said earlier, is a continuous and dynamic.

A lot of time is consumed while decisions are taken. In a management setting, decision cannot be taken abruptly. It should follow the steps such as

1. Defining the problem
2. Gathering information and collecting data
3. Developing and weighing the options
4. Choosing best possible option
5. Plan and execute
6. Take follow up action

Since decision making process follows the above sequential steps, a lot of time is spent in this process. This is the case with every decision taken to solve management and administrative problems in a business setting. Though the whole process is time consuming, the result of such process in a professional organization is magnanimous.

Decision Making in an Organizational Context

Decision making is part of everyone's life and all of us have to make decisions every moment. **In an organizational context, it is worthwhile to note that decision making needs the right kind of**

information, the complete information and the ability to synthesize and make sense of the information. While the first two attributes depend on external sources, the ability to make informed decisions is a personality trait. Hence, successful CEO's are those who can take into account the different viewpoints and divergent perspectives and arrive at the right decision.

The business landscape of the current times is littered with examples of companies that have made strategic errors and these are mostly to do with lack of proper decisions taken by the CEO's and managers in these firms. For instance, the failure of Chrysler and Ford (the automobile majors in the United States) to meet the challenge of competition from Japanese auto majors like Toyota was mostly due to the lack of imaginative decisions that would have responded to the threat in a coherent manner. Of course, it is another matter that these companies (Chrysler in particular) under the stewardship of Lee Iacocca were able to successfully meet the competition by the Japanese because of firm decisions taken by him.

The other aspect that relates to decision making in an organizational context is that there must be complete and accurate information made available to the decision maker. In Economics, there is a term called "asymmetries of information" that indicates how incomplete and insufficient information leads to poor decisions and wrong choices. What this concept means is that having partial information or faulty information often leads to "analysis paralysis" which is another term for poor decision making abilities. Finally, even with reliable and accurate information, the decision maker ought to have good problem solving skills and astute decision making abilities to arrive at sound judgments regarding the everyday problems and issues.

The overriding rule in decision making is that the decision maker ought to have legitimacy and authority over the people who he or she is deciding upon. In other words, decision makers succeed only when their decisions are honored and followed by the people or groups that the decision impacts. The reason for mentioning this towards the end is that in many cases, the fragmented nature of the organizations with different interests represented by factions often undermines the decision making capabilities of the decision maker. Hence, it is worth mentioning that such authority must be vested with the decision maker.

Rational and irrational decision-making:

In economics, it is thought that if humans are rational and free to make their own decisions, then they would behave according to rational choice theory. This theory states that people make decisions by determining the likelihood of a potential outcome, the value of the outcome, multiplying the two, and then choosing the more positive of the two outcomes.

In reality, however, there are some factors that affect decision-making abilities and cause people to make irrational decisions, one of them being availability bias. Availability bias is the tendency for some items that are more readily available in memory to be judged as more frequently occurring. For example, someone who

watches a lot of movies about terrorist attacks may think the frequency of terrorism to be higher than it actually is.

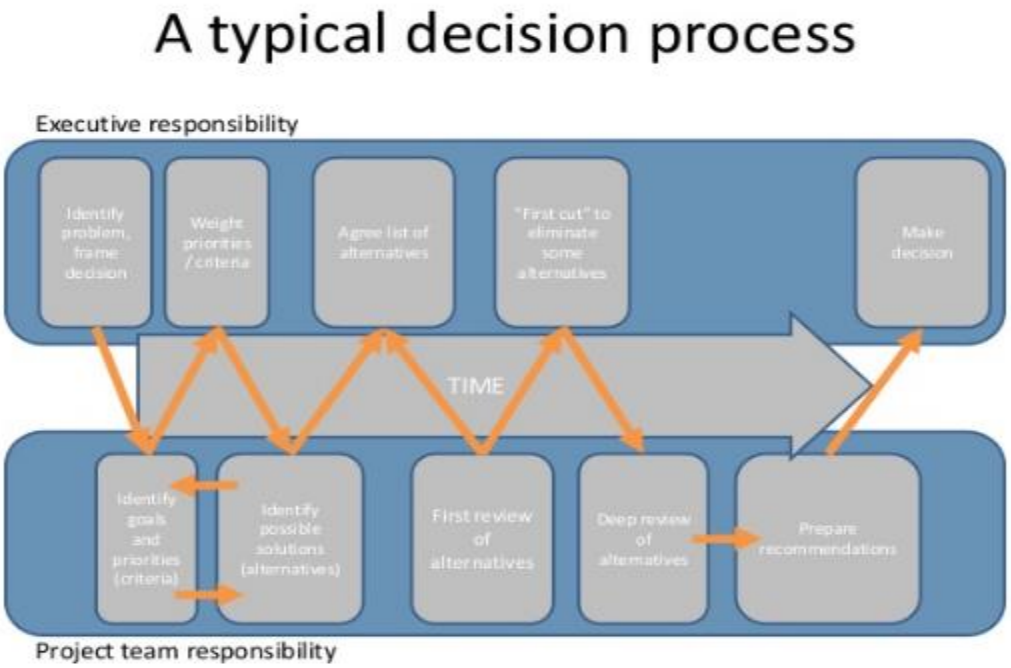


Figure: Decision making process

6.2Types of Decision

The word "**decision** "is derived from the Latin word "decido" which means "A decision, therefore is

- · A Settlement
- · A fixed intuition to bringing to a conclusive result
- · A judgment
- · A resolution

A decision is the choice out of several options made by the decision maker to achieve some objective in a given situation.

Business Decision: Business decisions are those which are made in the process of conducting business to achieve its objective in a given situation.

Characteristic of Business Decision Making:

a) Sequential in nature.

b) Exceedingly complex due to risk and trade off.

c) Influenced by personal values.

d) Made in institutional setting and business environment.

Rational Decision Making: A rational decision is the one which, effectively and efficiently, ensure the achievement of the goal for which the decision is made .In reality there is no right or wrong decision but a rational decision or irrational decision which depends on situation.

Type of Rationality

Objectively: Maximum the value of the objectives.

Subjective: If it is minimize the attainment of value in relation to the knowledge and awareness of subject.

Consciously: Extent the process of the decision making is a conscious one

Organizationally: degree of the orientation towards the organization.

Personal: Rational to the extent is achiever's an individual's personal reason (goals).

Type of Decision Making System

There are two types of decision making system on the basis of knowledge about the environment.

(i) **Closed:** If the manager operates in a known environment then it is called closed decision making system.

Conditions:

a) Manager knows the set of decision alternative and know their outcome in term of values.

b) Manager has a model, by which decision alternatives can be generated, tested and ranked.

c) The manager can choose one of them, based on some goal or objective.

(ii) **Open:** If the manager operates in unknown environment then it is called open decision making.

Conditions:

a) Manager does not know all alternatives.

b) Outcome is not known.

c) No methods or models are used.

d) Decide objective or goal; select one where his aspirates or desire are met best.

Types of Decision: Types of decision are based on the degree of knowledge about the outcome of the events which are yet to take place.

Certainty: If the manager has full knowledge of event or outcome then it is a situation of certainty.

Risk: If the manager has partial knowledge or probabilistic knowledge then it is decision under risk.

Uncertainty: If the manager does not have any knowledge, it is decision making under uncertainty MIS converts the uncertainty to risk and risk to certainty. The decision at the low level management is certain, at middle level of the management the decision is under risk and at the top level management the decision is in under uncertain.

Nature of decision: Decision making is a complex task. To resolve the complexity the nature of decision are of two types:

Programmed and Non-Programmed Decision :

a) If a decision can be based on a rule, methods or even guidelines, it is called the programmed decision.

b) A decision which cannot be made by using a rule or model is the non programmed decision.

Decision-making techniques can be separated into two broad categories:

- Group decision-making and individual decision-making techniques.

Group decision-making techniques:

- Consensus decision-making tries to avoid "winners" and "losers". Consensus requires that a majority approve a given course of action, but that the minority agree to go along with the course of action. In other words, if the minority opposes the course of action, consensus requires that the course of action be modified to remove objectionable features.
- Voting-based methods.

  Range voting lets each member score one or more of the available options. The option with the highest average is chosen. This method has experimentally been shown to produce the lowest Bayesian regret among common voting methods, even when voters are strategic.

  Majority requires support from more than 50% of the members of the group. Thus, the bar for action is lower than with unanimity and a group of "losers" is implicit to this rule.
- Plurality, where the largest block in a group decides, even if it falls short of a majority.
- Delphi method is structured communication technique for groups, originally developed for collaborative forecasting but has also been used for policy making.
- Dotmocracy is a facilitation method that relies on the use of special forms called Dotmocracy Sheets to allow large groups to collectively brainstorm and recognize agreement on an unlimited number of ideas they have authored.

Individual decision-making techniques

- Pros and cons: listing the advantages and disadvantages of each option, popularized by Plato and Benjamin Franklin. Contrast the costs and benefits of all alternatives. Also called "rational decision-making".
- Simple prioritization: choosing the alternative with the highest probability-weighted utility for each alternative (see Decision analysis).

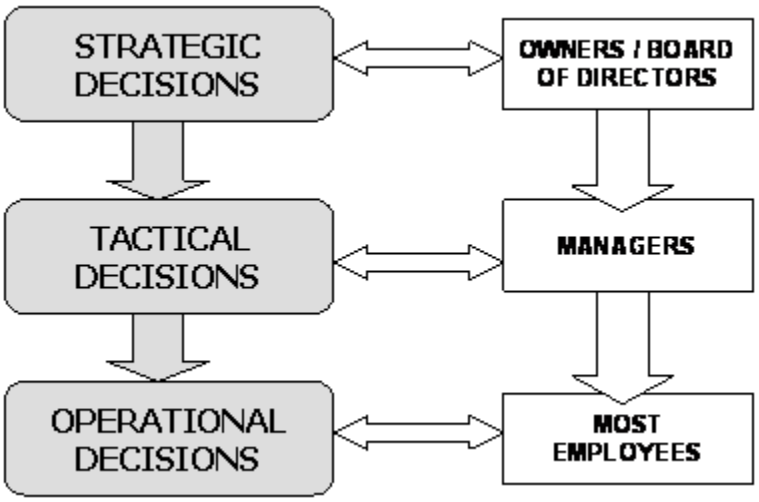
- Satisfying: examining alternatives only until an acceptable one is found. Contrasted with maximizing, in which many or all alternatives are examined in order to find the best option.
- Elimination by aspects: choosing between alternatives using Mathematical psychology. The technique was introduced by Amos Tversky in 1972. It is a covert elimination process that involves comparing all available alternatives by aspects. The decision-maker chooses an aspect; any alternatives without that aspect are then eliminated. The decision-maker repeats this process with as many aspects as needed until there remains only one alternative
- Preference trees: In 1979, Tversky and Shmuel Sattach updated the elimination by aspects technique by presenting a more ordered and structured way of comparing the available alternatives. This technique compared the alternatives by presenting the aspects in a decided and sequential order. It became a more hierarchical system in which the aspects are ordered from general to specific
- Acquiesce to a person in authority or an "expert"; "just following orders".
- Flipism: flipping a coin, cutting a deck of playing cards, and other random or coincidence methods
- Prayer, tarot cards, astrology, augurs, revelation, or other forms of divination.
- Taking the most opposite action compared to the advice of mistrusted authorities (parents, police officers, partners...)
- Opportunity cost: calculating the opportunity cost of each options and decide the decision.
- Bureaucratic: set up criteria for automated decisions.
- Political: negotiate choices among interest groups.
- Participative decision-making (PDM): a methodology in which a single decision-maker, in order to take advantage of additional input, opens up the decision-making process to a group for a collaborative effort.
- Use of a structured decision-making method.
- Individual decision-making techniques can often be applied by a group as part of a group decision-making technique.

A need to use software for a decision-making process is emerging for individuals and businesses. This is due to increasing decision complexity and an increase in the need to consider additional stakeholders, categories, elements or other factors that effect decisions.

6.3 Decision-Making Levels

The management process is executed through a variety of decisions taken at each step of planning organizing, staffing, directing, coordinating and control. As discussed earlier the MIS aids decision making. If the management is able to spell out the decisions required to taken in these steps are tabulated below:

Steps in management	Decision
Planning	A selection from various alternatives- strategies, resources, methods, etc.
Organization	A selection of a combination out of several combinations of the goals, people, resources, method, and authority.
Staffing	Providing a proper manpower complement.
Directing	Choosing a method from the various methods of directing the efforts in the organization.
Coordinating	Choice of the tools and the techniques for coordinating the efforts for optimum results.
Controlling	A selection of the exceptional conditions and the decision guidelines.



6.4Transaction Processing System (TPS)

The most fundamental computer based system in an organization pertains to the processing of business transactions. A transaction processing system can be defined as a system that captures, classifies, stores, maintains, updates and retrieves transaction data for record keeping and input to the other types of CBIS. Transaction Processing System is aimed at improving the routine business activities. A transaction is any event or activity that affects the whole organization. Placing order, billing customers, hiring of employees and depositing cheques are some of the common transactions. Types of transactions that occur vary from organization to organization but this is true that all organizations process transaction as a major part of their

daily business activities. Transaction Processing System provides speed and accuracy and can be programmed to follow routines without any variance.

It is a type of information system. TPSs collect, store, modify, and retrieve the transactions of an organization. A transaction is an event that generates or modifies data that is eventually stored in an information system. It is recommended that a transaction processing system should pass the ACID test. The essence of a transaction program is that it manages data that must be left in a consistent state, e.g. if an electronic payment is made, the amount must be both withdrawn from one account and added to the other; it cannot complete only one of those steps. Either both must occur, or neither. In case of a failure preventing transaction completion, the partially executed transaction must be 'rolled back' by the TPS. While this type of integrity must be provided also for batch transaction processing, it is particularly important for online processing: if e.g. an airline seat reservation system is accessed by multiple operators, after an empty seat inquiry, the seat reservation data must be locked until the reservation is made, otherwise another user may get the impression a seat is still free while it is actually being booked at the time. Without proper transaction monitoring, double bookings may occur. Other transaction monitor functions include deadlock detection and resolution (deadlocks may be inevitable in certain cases of cross-dependence on data), and transaction logging (in 'journals') for 'forward recovery' in case of massive failures.

Types

Contrasted with batch processing

Batch processing is a form of transaction processing. Batch processing involves processing several transactions at the same time, and the results of each transaction are not immediately available when the transaction is being entered, there is a time delay. Transactions are accumulated for a certain period (say for day) where updates are made especially after work. Online transaction processing is the form of transaction processing that processes data as it becomes available.

Real-time and batch processing

There are a number of differences between **real-time** and **batch processing**. These are outlined below:

Each transaction in real-time processing is unique. It is not part of a group of transactions, even though those transactions are processed in the same manner. Transactions in real-time processing are stand-alone both in the entry to the system and also in the handling of output.

Real-time processing requires the master file to be available more often for updating and reference than batch processing. The database is not accessible all of the time for batch processing.

Real-time processing has fewer errors than batch processing, as transaction data is validated and entered immediately. With batch processing, the data is organised and stored before the master file is updated. Errors can occur during these steps.

Infrequent errors may occur in real-time processing; however, they are often tolerated. It is not practical to shut down the system for infrequent errors.

More computer operators are required in real-time processing, as the operations are not centralised. It is more difficult to maintain a real-time processing system than a batch processing system.

Features of TPS

- **Rapid response**

Fast performance with a rapid response time is critical. Businesses cannot afford to have customers waiting for a TPS to respond, the turnaround time from the input of the transaction to the production for the output must be a few seconds or less.

- **Reliability**

Many organizations rely heavily on their TPS; a breakdown will disrupt operations or even stop the business. For a TPS to be effective its failure rate must be very low. If a TPS does fail, then quick and accurate recovery must be possible. This makes well–designed backup and recovery procedures essential.

- **Inflexibility**

A TPS wants every transaction to be processed in the same way regardless of the user, the customer or the time for day. If a TPS were flexible, there would be too many opportunities for non-standard operations, for example, a commercial airline needs to consistently accept airline reservations from a range of travel agents, accepting different transactions data from different travel agents would be a problem.

- **Controlled processing**

The processing in a TPS must support an organization's operations. For example if an organization allocates roles and responsibilities to particular employees, then the TPS should enforce and maintain this requirement. An example of this is an ATM transaction

6.5 Management Information System (MIS)

MIS stands for management information systems. Management information system is that system which allows the managers to make decisions for the successful operation of the businesses. MIS also refers to the organizations that develop and maintain most of the computer system in the enterprise helpful for the managers to make decisions.

MIS is a combination of 3 English Letters:

M which stands for Management

I which stands for Information

S which stands for System

With the help of these 3 letters we can make a number of combinations:

MI = Management Information means information regarding management.

MS= management System means the basic structure of the management like the hierarchical order of management.

IS= Information System which provides the information

SM= System Management means how to manage a system whether it is a business, organization etc

SI = System Information means the information regarding the system like what are the different parts of the system, how they relate to each other etc.
IM = Information Management means how to manage a particular information.
MIS = Management Information System.

The MIS can be defined as a system which helps to provides information support for decision making in the organization.

The MIS is defined as an integrated system of man and machine for providing the information to support the operations, the management and the decision making function in the organization.

The MIS can also be defined as a system based on the database of the organization evolved for the purpose of providing information to the people in the organization.

Another definition of MIS is that it is a Computer based Information System. It refers to a computer-based system that provides managers with the tools for organizing, evaluating and efficiently running the departments of the organization. it helps in providing past, present and prediction information. The use of software helps in fast decision making, data resources, decision support systems, people management, project management etc to help the departments run efficiently. In today's world MIS is a computerized business processing system which is helpful in generating information for the people in the organization to meet the information. It also helps in better decision making so as to achieve the corporate objective of the organization.

The main goal of MIS in an organization is to deliver information systems to the various levels of corporate managers. MIS professionals create and support the computer system throughout the company. MIS is a complex system comprising of different subsystems in which data is processed and the information generated as a result is used at different levels of management from top to bottom level. The management analyzes the environment and then sets goals and objectives to be accomplished. To perform the task information is required. To perform functions like planning, organizing, staffing, directing and controlling the information is provided or supplied by the MIS

ACTIVITY

1. Can you think of an MIS which could be user independent and business dependent?
2. Take an organization of your knowledge and give a conceptual view of MIS and physical view of MIS.

6.6 Executive Information System (EIS)

It is a reporting tool that provides quick access to summarized reports coming from all company levels and departments such as accounting, human resources and operations.

6.7 Decision Support System (DSS)

Decision support system (**DSS**) is a computer-based information system that supports business or organizational decision-making activities. DSSs serve the management, operations, and planning levels of an organization and help to make decisions, which may be rapidly changing and not easily specified in advance. Decision support systems can be either fully computerized, human or a combination of both.

DSSs include knowledge-based systems. A properly designed DSS is an interactive software-based system intended to help decision makers compile useful information from a combination of raw data, documents, and personal knowledge, or business models to identify and solve problems and make decisions.

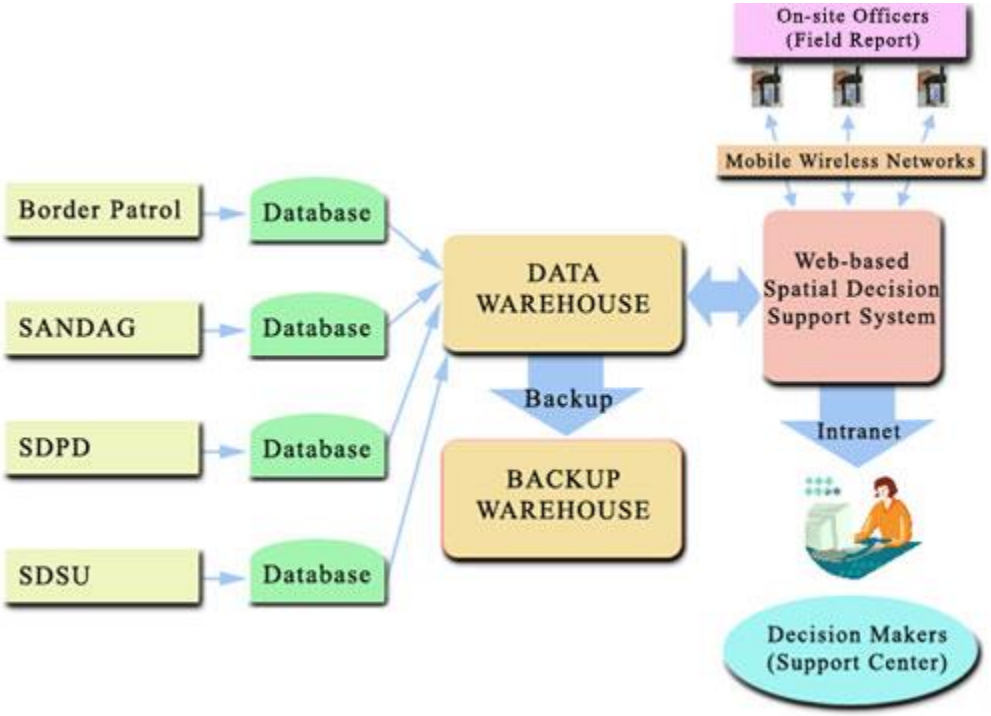


Fig: DSS

Typical information that a decision support application might gather and present includes:

- inventories of information assets (including legacy and relational data sources, cubes, data warehouses, and data marts),
- comparative sales figures between one period and the next,
- projected revenue figures based on product sales assumptions

Decision Support System refers to a class of systems which support in the process of decision making and does not always give a decision itself.

Decision Support Systems (DSS) are a specific class of computerized information system that supports business and organizational decision-making activities. A properly designed DSS is an interactive software-based system intended to help decision makers compile useful information from raw data, documents, personal knowledge, and/or business models to identify and solve problems and make decisions.

DSS is an application of Hebert Simon model, as discussed, the model has three Phases**:**
i) Intelligence
ii) Design
iii) Choice

The DSS basically helps in the information system in the intelligence phase where the objective is to identify the problem and then go to the design phase for solution. The choice of selection criteria varies from problem to problem.

It is therefore, required to go through these phases again and again till satisfactory solution is found. In the following three phase cycle, you may use inquiry, analysis, and models and accounting system to come to rational solution.

These systems are helpful where the decision maker calls for complex manipulation of data and use of several methods to reach an acceptable solution using different analysis approach. The decision support system helps in making a decision and also in performance analysis. DSS can be built around the rule in case of programmable decision situation. The rules are not fixed or predetermined and requires every time the user to go through the decision making cycle as indicated in Herbert Simon model.

Attribute:
i) DSS should be adaptable and flexible
ii) DSS should be interactive and provide ease of use
iii) Effectiveness balanced with efficiency (benefit must exceed cost).
iv) Complete control by decision-makers
v) Ease of development by (modification to suit needs and changing environment) end users
vi) Support modeling and analysis

vii) Data access

viii) Standalone, integration and Web-based

DSS Characteristics:

i) Support for decision makers in semi structured and unstructured problems.

ii) Support managers at all levels.

iii) Support individuals and groups.

iv) Support for interdependent or sequential decisions.

v) Support intelligence, design, choice, and implementation.

vi)Support variety of decision processes and styles

6.8 Expert System (ES)

Expert System is used to provide expert advice to the end user for decision making. It is further used for providing critical information for executives and managers

Example: credit application advisor

An expert system is a knowledge-based information system; that is, it uses its knowledge about a specific area to act as an expert consultant to users. The components of an expert system are a knowledge base and software modules that perform inferences on the knowledge and offer answers to a user's questions.

Expert systems provide answers to questions in a very specific problem area by making human like inferences about knowledge contained in a specialized knowledge base. Expert systems can provide decision support to end users in the form of advice from an expert consultant in a specific problem area.

Expert systems are being used in many different fields, including medicine, engineering, the physical sciences, and business. For example, expert systems now help diagnose illnesses, search for minerals, analyze compounds, recommend repairs, and do financial planning. Expert systems can support either operations or management activities.

Expert Systems Structure

The components of an expert system include a knowledge base and software modules that perform inferences on the knowledge in the knowledge base and communicate answers to a user's questions.

The knowledge base of an expert system contains Facts about a specific area, Heuristics (thumbs of rule) that express the reasoning procedures of an expert on the subject. There are many ways that knowledge is represented in expert systems:-

Case-based reasoning: Representing knowledge in an expert system's knowledge base in the form of cases.

- **Frame-based knowledge:** Knowledge represented in the form of a hierarchy or network of frames. A frame is a collection of knowledge about an entity consisting of a complex package of data values describing its attributes.
- **Object-based knowledge:** Knowledge represented as a network of objects. An object is a data element that includes both data and the methods or processes that act on those data.
- **Rule-based knowledge:** Knowledge represented in the form of rules and statements of fact. Rules are statements that typically take the form of a premise and a Summary such as: IF (condition), Then (Summary).
- **Software resources:** An expert system software package contains an inference engine and other programs for refining knowledge and communicating with users. The inference engine program processes the knowledge (such as rules and facts) related to a specific problem. It then makes associations and inferences resulting in recommended courses of action for a user. User interface programs for communicating with end-users are also needed, including an explanation program to explain the reasoning process to a user if requested.

Differences between DSS and ES

It is possible to integrate ES with DSS. There may be some components which may look similar in DSS and ES. But one should understand the differences between them. It then becomes clear as to how integration of ES with DSS can be realized.

- A DSS helps manager to take a decision whereas an ES acts as a decision maker or an advisor to the manager.
- A DSS is meant only for decision making whereas an ES provides expertise to the manager.
- The spectrum of complexity is high in DSS and low in ES since ES addresses issues related to specific areas only.
- DSS does not capability to reason whereas an ES has.
- A DSS cannot provide detailed explanation about the results whereas an ES can.

Hence by integrating the two it is possible the blend their advantages and derive the best out of the two.

Expert Systems Business Applications

Expert systems help diagnose illness, search minerals, analyze compounds, recommend repairs, and do financial planning. So from a strategic business point, expert systems can and are being used to improve every step of the product cycle of a business, from finding customers to shipping products to providing customer service. ES provides a cost reduced solution, consistent advice with low level of errors, solution to

handle equipments without the interference of human. It provides a high degree of reliability and faster response time. It helps to solve complex problem with in a small domain.

It is capable of analyzing the problem and can construct a business model appropriate to the characteristics of the application. Based on the model necessary objectives and constraints are identified. It identifies appropriate tools to solve the model. It uses the tools to solve the problem and also does the what –if analysis aimed at understanding the sensitivity of the model.

6.9 Office Automation System (OAS)

Office Automation Systems are among the newest and most rapidly expanding computer based information systems. They are being developed with the hope and expectation that they will increase the efficiency and productivity of office workers, typists, secretaries, administrative assistants, staff professionals, managers and others.

Office automation refers to the varied compute machinery and software used to digitally create, collect, store, manipulate, and relay office information needed for accomplishing basic tasks. Raw data storage, electronic transfer, and the management of electronic business information comprise the basic activities of an office automation system. Office automation helps in optimizing or automating existing office procedures.

The backbone of office automation is a LAN, which allows users to transmit data, mail and even voice across the network. All office functions, including dictation, typing, filing, copying, fax, Telex, microfilm and records management, telephone and telephone switchboard operations, fall into this category. Office automation was a popular term in the 1970s and 1980s as the desktop computer exploded onto the scene. ADVANTAGES are:- 1.office automation can get many task accomplished faster. 2.it eliminates the need for a large staff. 3.less storage required for data to store. 4.multiple people can updated data Simultaneously in the event of schedule change

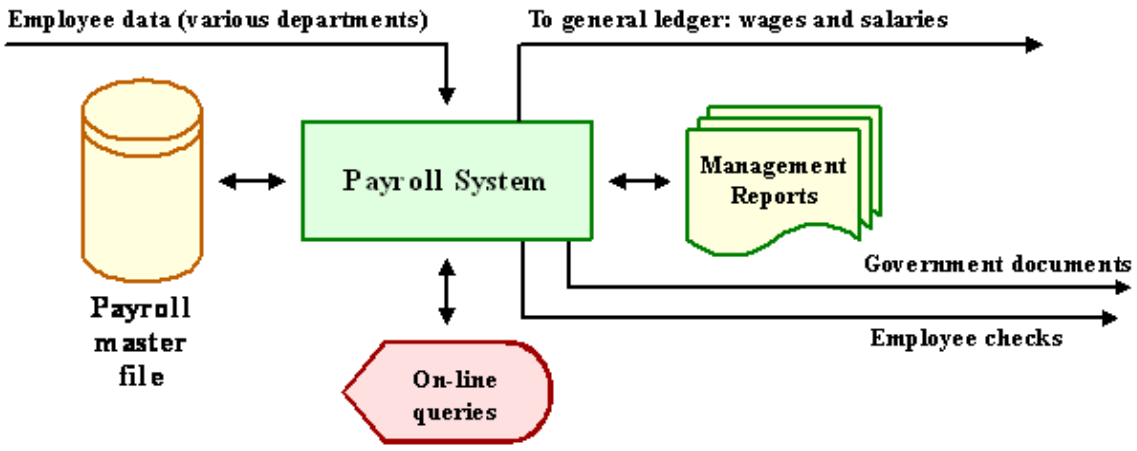


Fig: Office Automation System

Office Automation system collects, processes, stores and transmits data and information in the form of electronic office communication.

OAS could also be considered as computer-based information system that collects, process, store and transmit electronic message, document and other form of communication among individual, work group and organizations.

6.10 Collaboration Technologies

Collaboration Technologies is dedicated to deliver good quality services in Software development. Collaborative software or groupware is an application software designed to help people involved in a common task to achieve goals. One of the earliest definitions of collaborative software is 'intentional group processes plus software to support them.'

The design intent of collaborative software is to transform the way documents and rich media are shared to enable more effective team collaboration. Collaboration, in terms of information technology, seems to have several definitions. Understanding the differences in human interactions is necessary to ensure that appropriate technologies are employed to meet interaction needs.

Collaboration requires individuals working together in a coordinated fashion, towards a common goal. Accomplishing the goal is the primary purpose for bringing the team together. Collaborative software helps facilitate action-oriented teams working together over geographic distances by providing tools that aid communication, collaboration, and the process of problem solving. Additionally, collaborative software may support project management functions, such as task assignments, time-managing deadlines, and shared calendars. The artifacts, the tangible evidence of the problem solving process, and the final outcome of the collaborative effort, require documentation and may involve archiving project plans, deadlines, and deliverables.

Collaborative software is a broad concept that overlaps considerably with Computer-supported cooperative work (CSCW). Some authors argue they are equivalent. According to Carstensen and Schmidt (1999) groupware is part of CSCW. The authors claim that CSCW, and thereby groupware, addresses "how collaborative activities and their coordination can be supported by means of computer systems." Software products such as email, calendaring, text chat, wiki, and bookmarking belong to this category whenever used for group work. Whereas the more general term social software applies to systems used outside the workplace, for example, online dating services and social networking sites like Twitter and Facebook. It has

been suggested that Metcalfe's law — the more people who use something, the more valuable it becomes — applies to these types of software.

The use of **collaborative software** in the work space creates a collaborative working environment (CWE). A collaborative working environment supports people in both their individual and cooperative work thus evolving into a new class of professionals, e-professionals, who can work together irrespective of their geographical location.

Finally, collaborative software relates to the notion of collaborative work systems, which are conceived as any form of human organization that emerges any time that collaboration takes place, whether it is formal or informal, intentional or unintentional. Whereas the groupware or collaborative software pertains to the technological elements of computer-supported cooperative work, collaborative work systems become a useful analytical tool to understand the behavioral and organizational variables that are associated to the broader concept of CSCW.

6.11Information Systems in Functional Areas:

6.11.1 HR

A human resource information system (HRIS) is defined as a computer based application for assembling and processing data related to the **human resource management** (HRM) function. As in other types of information systems, an HRIS consists of a database, which contains one or more files in which the data relevant to the system are maintained, and a **database management system,** which provides the means by which users of the system access and utilize these data. The HRIS thus contains tools that allow users to input new data and edit existing data; in addition, such programs provide users with the opportunity to select from an array of predefined reports that may either be printed or displayed on a monitor. Reports may address any of a number of different HRM issues (e.g., succession planning, compensation planning, equal employment opportunity monitoring). HRISs also generally include tools by which users or system administrators may generate ad hoc reports and select specific cases or subsets of cases for display.

In order to understand the types of applications available to HRIS users, it is best to consider the evolving nature of human resource information systems applications. The HRM field lagged behind a number of other functional areas of **management** in the utilization of computer applications, but beginning in the late 1980s extensive use of sophisticated applications began to appear. Prior to that time, manual record systems often dominated in personnel or human resource departments. Computer applications used in the field were generally limited to basic record keeping and payroll management systems. Virtually all such systems were based on mainframe **computers** and required extensive support from information systems professionals.

Thus, human resource managers had little opportunity to design sophisticated reports and computer-based analytical tools to aid in managerial **decision making.** In general, uses of computers in HRM fell into the category of electronic data processing applications, which generally involve the automation of relatively routine tasks (e.g., calculating pay and printing checks).

A number of trends seem to have contributed to a growing reliance on computers as information-processing and decision-aiding tools in HRM. The emergence of the human resource management field (versus personnel administration) gave the human resource function greater credibility within the managerial hierarchy, necessitating more sophisticated use of information, especially as it related to the strategic management function. Firms have experienced increased competitive pressures that have translated into greater cost containment demands from upper management, leading to greater automation of the record-keeping function in the HRM field. The ready availability of microcomputers and relatively user friendly software means, that to an increasing extent, human resource managers are no longer dependent on information system professionals to develop and implement applications (which might be assigned a lower priority than other management functions). Many HRM departments in larger organizations have also developed internal information system capabilities, so that HRIS units have been established.

Another important factor has been the development of numerous HRIS products by external vendors. There are several full-featured, human resource-dedicated database management systems available, both for mainframe and micro platforms. Perhaps the best known of these is PeopleSoft, although numerous other such products exist. Many of these utilize client-server architectures, where databases reside on a central server and are accessed from individual workstations, connected to a network, via local client applications. There is also a trend toward enterprise-wide applications that integrate information system applications for various managerial functions (e.g., **marketing, finance,** human resources), which facilitates communication across functional areas, economizes on information system development at the enterprise level, and allows firms to collate information from multiple sources to facilitate strategic planning at the business unit and corporate levels. Examples of commonly used integrated systems that include HRIS modules are SAP and Oracle.

Specialized applications, intended to supplement HRISs, are also widely available. These include modules to aid in such areas as succession planning, benefits administration, applicant tracking, job evaluation, employee performance evaluation, grievance handling, and labor relations. These products, coupled with declining costs of computer systems (especially microcomputer systems) and the increasing user-friendliness of computer applications, have meant that the use of HRISs is increasingly attractive to practitioners.

As a result of the rapid change in computer and software technology, HRIS-related products are constantly upgrading and changing. Advanced Personnel Systems, a California-based HRIS consulting firm, markets a fairly up-to-date database listing a wide-range of HRISs and supplementary programs—along with descriptions of the capabilities of these systems—to guide managers in selecting appropriate products. There are also several conferences and shows held annually around the United States that are dedicated to advances in HRM-related **information technology** applications, including HRISs. HRIS vendors often demonstrate products at such shows.

Perhaps the most significant development in the HRIS area currently is the growing use of organizational intranets as a means of managing many aspects of a firm's HRIS. An intranet is an internal network that makes use of World Wide Web technology (browsers, servers, etc.) to gather and disseminate information within the firm. Intranets may be linked to the external Internet, but are usually secured in a variety of ways so that only authorized users can access the information on the internal components. While it is quite easy to generate static extracts of HRIS data tables, queries, forms, and reports for posting on an intranet, it is also quite feasible to establish live links between an intranet and a firm's HRIS. This allows real-time collection and display of information. Thus employees can complete forms online that enroll them in benefits programs, allow them to bid on job openings, let them submit suggestions, and facilitate filing of various claims. In addition, intranet displays can be tailored to the needs of specific users. The user may check on the current status of his or her fringe benefits, vacation time, training program enrollment, or **pension fund.** Intranets obviously require extensive security measures to prevent inappropriate changing or accessing of data. The issues seem to have been addressed, however, and HRIS products are increasingly emphasizing their functionality in intranet environments.

All indications are that HRISs will continue to play an increasingly important role in the HRM field. An important driving factor is that HRISs facilitate process reengineering in the HRM area, thus promoting greater efficiency. Functions that were carried out manually and in many steps can often be largely automated. And the business firm's continuing demand for information to facilitate planning and strategy formulation will necessitate further reliance on HRISs. Finally, the linking of HRISs to organizational intranets is apt to gain in popularity as a relatively inexpensive and appealing means of gathering and distributing human resource information.

6.11.2 Marketing

A **marketing IS** supports activities throughout the many activities of marketing departments. Some of the typical subsystems of a marketing MIS are marketing research, product development and delivery, promotion and advertising, product pricing and sales analysis.

One of the most common uses of a marketing IS is to produce sales reports. These are typically produced on a regular schedule, such as by week, month and quarter. Reports can be organized by sales representative, product, customer or geographic area. Such reports allow managers to see which aspects of sales are doing well and which ones need attention.

Perhaps one sales representative has suddenly experienced a drop in sales by losing one major customer and needs some support to develop some new leads. If there are only a handful of sales reps sharing one office, a manager might be able to pick up on this just by talking to everyone. However, what if a manager has to oversee more than 100 sales reps in 12 different offices around the nation? A specialized information system that provides regular updates in a meaningful format will make it a lot easier for the manager to make effective decisions

In order to pursue market opportunities as well as anticipate marketing problem, manager need to collect comprehensive and reliable information. Managers cannot carryout marketing analysis, planning, implementation and control without monitoring and researching customers, competitors, dealers and their sales and cost data. Every firm has many information flows of interest to marketing management. Many companies are studying their executive's information needs and design information system for marketing to meet these needs. Instead of plethora of unrelated data, an MIS combines various inputs and present integrated reports.

Definition : Marketing Information System is a continuing and interacting structure of people, equipments and procedures to gather, sort, analyze, evaluate, and distribute pertinent, timely and accurate information for use by marketing decision makers to improve their marketing planning, implementation and control activities.

Components of Marketing Information System : As shown in figure below, the box on the left shows components of the marketing environment that manager must monitor. Trends in the marketing environment are picked up and analyzed through four subsystems making up the marketing information system- Internal Accounting System, Marketing Intelligence System, Marketing Research System and Analytical Marketing System.

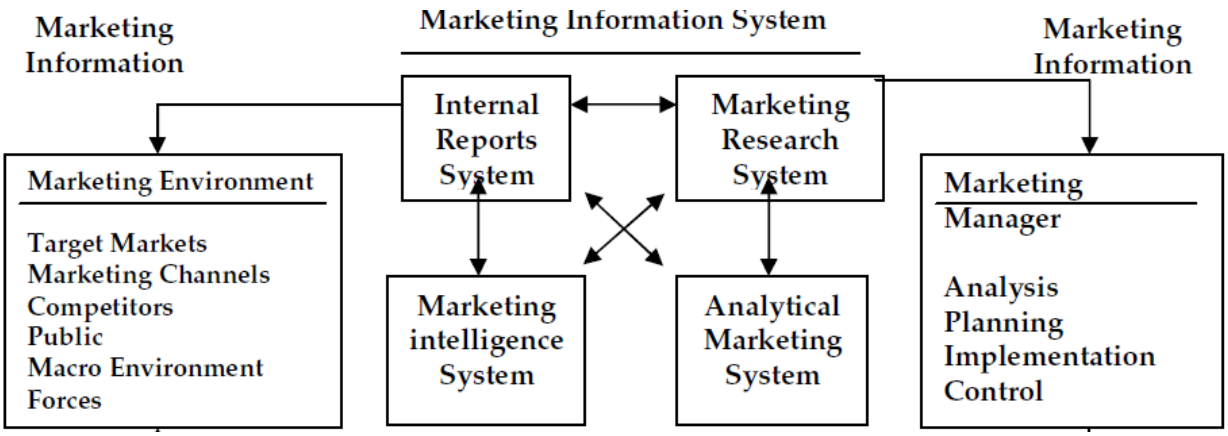


Fig: MIS and Marketing Information

Internal Accounting System is the most basic information system used by marketing executives. It is the system that reports orders, sales inventory levels, receivable, payable. By analyzing the information, marketing managers can spot important opportunities and problems.

_ **The Order Shipping Cycle:** Sales representatives, dealers and customers dispatch orders to the firm. The order department prepares multi-copy invoice and sends them to various departments. Out of stock items are back ordered. Shipped items are accompanied and sent to various departments. The company wants to carryout these steps quickly and accurately. The computer is harnessed to expedite the order shipping billing cycle.

_ **Improving the Timeliness or Sales Reports:** Marketing executives receive sales reports some times, after the sales have taken place. Many companies complain that sales are not reported fast enough in their company. Marketing information system can improve these things rapidly.

_ **Designing a User Oriented Report System:** In designing an advanced sales information system, the company should avoid certain pitfalls. The marketing information system should represent a cross between what managers think they need, what managers really need and what is economically feasible. Management information system should provide the reports for all marketing departments. Information system can delete the unwanted system from the survey and from other departments and prepare reports which are required by different persons of marketing department.

6.11.3 Finance

A financial IS provides financial information for managers to make daily decisions on operations within the organization. Most systems provide these functions:

- Integrate financial information from multiple sources
- Provide easy access to financial information in summarized form
- Enable financial analysis using easy-to-use tools
- Compare historic and current financial activity

A financial MIS often has a number of subsystems, depending on the type of organization. These include systems to analyze revenues, costs and profits, auditing systems for both internal and external purposes and systems to manage funds. A financial MIS can also be used to prepare reports for third parties, such as external auditors or shareholders.

Financial accounting systems and subsystems are just one type of institutional MIS. Financial accounting systems are an important functional element or part
of the total MIS structure. However, they are more narrowly focused on the internal balancing of an institution's books to the general ledger and other financial accounting subsystems. For example, accrual adjustments, reconciling and correcting entries used to reconcile the financial systems to the general ledger are not always immediately entered into other MIS systems.

Accordingly, although MIS and accounting reconcilement totals for related listings and activities should be similar, they may not necessarily balance.

Financial management function has a primary objective of meeting the financial needs of the business. The second objective of FM is to meet the statutory compliance by way of declaring the auditing financial result, submitting reports and returns to the govt. and Tax authorities and fulfill the obligations to the shareholders.

FM uses variety of tools and techniques like Break Even Analysis, ABC Analysis, Ratio Analysis, Management Accounting and Cost Analysis.

Input Documents :

o Receipts from customers, authorities, employees, share holders, financial institution and others.
o Payment to suppliers, authorities, share holders, financial institutions and others.
o Data from stock exchange on the shares prices consolidated financial results of the other companies etc.
Transactions are payments and receipts and they are documented through journal vouchers, bills, debit notes, credit notes, receipts and transfer documents.

Application of Financial Management Information System : The major application of financial management information system includes financial accounting system, which accounts for the financial

transactions of the company and produces financial results for the company. It produces balance sheet for the company where the performance of the company is published in standard format prescribed by the govt. The system is made so comprehensive that it not only collects financial data but also collects data on different matters such as job, department, and division and so on. It forms a basis for certain reports which are required by the top level management. The users of the financial data base are finance managers, cost controller, auditors, material managers, marketing managers, company secretaries and the top management.

6.11.4 Production

Manufacturing or production information system provides information on production /operation activities of an organization and thus facilitates the decision-making process of production managers of an organization. The main decisions to be taken in manufacturing system are: Product Design

SELF ASSESSMENT QUESTIONS

Please select the correct option:

A. Information systems that support the business functions that reach out to suppliers are known as:
   1. back office information systems
   2. decision support systems
   3. expert information systems
   4. front office information systems
   5. none of the above

B. Which of the following is not a class of information system applications?
   1. database management system
   2. decision support system
   3. expert system
   4. management information system
   5. office automation system

C. Who are the people that actually use the system to perform or support the work to be completed?
   1. system analysts
   2. system designers
   3. system owners
   4. system builders
   5. none of the above

D. Which is not a typical business function?
   1. Sales
   2. Service
   3. Manufacturing
   4. Accounting
   5. Benefits and Compensation

6.12 Sources of competitive advantages: IS for Competitive Advantages

Gaining competitive advantage is critical for organizations. Baltzan and Phillips define competitive advantage as 'a product or service that an organization's customers value more highly than similar offerings from its competitors' (in other words, you have something useful (i.e. products, services, capabilities) that your competitors do not have). Competitive advantages are typically temporary as competitors often seek ways to duplicate the competitive advantage. In order to stay a head of competition, organizations have to continually develop new competitive advantages.

This section discusses how an organization can analyze, identify, and develop competitive advantages using tools such as Porter's Five Forces, three generic strategies, and value chains. Michael Porter's Five Forces Model is a useful tool to assist in assessing the competition in an industry and determining the relative attractiveness of that industry. Porter states that in order to do an industry analysis a firm must analyse five competitive forces:

- Rivalry of competitors within its industry
- Threat of new entrants into an industry and its markets
- Threat posed by substitute products which might capture market share
- Bargaining power of customers
- Bargaining power of suppliers.

To survive and succeed, a business must develop and implement strategies to effectively counter the above five competitive forces. O'Brien and Marakas suggested that organizations can follow one of five basic competitive strategies, which are based on Porter's three generic strategies of broad cost leadership, broad differentiation, and focused strategy. The five competitive strategies are: cost leadership, differentiation, innovation, growth, and alliance. Meanwhile, information systems could be a critical enabler of these five competitive strategies (see Table).

Table: Competitive Strategies & Roles of Information Systems

Competitive Strategy	Roles of Information Systems
Cost Leadership	Organizations can use information systems to fundamentally shift the cost of doing business or reduce the costs of business processes or/and to lower the costs of customers or suppliers, i.e., using online business to consumer & business to business models, e-procurement systems to reduce operating costs

Differentiation	Organizations can use information systems to develop differentiated features or/and to reduce competitors' differentiation advantages, i.e., using online live chatting systems and social networks to better understand and serve customers; using technology to create inform diaries to offer value-added service and improve customers' stickiness to your web site/business; applying advanced and established measures for online operations to offline practices (i.e., more accurate and systematic ways of measuring efficiency and effectiveness of advertising
Innovation	Organizations can use information systems to identify and create (or assist in creating) new products and services or/and to develop new/niche markets or/and to radically change business processes via automation (i.e., using digital modeling and simulation of product design to reduce the time and cost to the market. They also can work on new initiatives of establishing pure online businesses/operations. At the same time, the Internet and telecommunications networks provide better capabilities and opportunities for innovation. "Combinational innovation" and Open innovation are two good examples. There are a large number of component parts on the networks that are very expensive or extremely different before the establishment of the networks, and organizations could combine or recombine components/parts on the networks to create new innovations. Meanwhile everyone is connected via personal computers, laptops and other mobile devices through cabled Internet or wireless networks or mobile networks, there are plenty of opportunities to co-create with customers, external partners and internal people.
Growth (including mergers and acquisitions)	Organizations can use information systems to expand domestic and international operations or/and to diversify and integrate into other products and services, i.e., establishing global intranet and global operation platform; establishing Omni-channel strategy to gain growth(Omni-channel strategy looks at leveraging advantages of both online (or digital) and offline (or non-digital) channels)
Strategic Alliance	Organizations can use information systems to create and enhance relations with partners via applications, such as developing virtual organizations and inter-organizational information systems.

On top of these five basic strategies, companies can also adopt other competitive strategies facilitated by information systems to shape their competitive advantage. Some examples include:

• Locking in customers or suppliers by enhancing relations and building valuable new relationships via customer/partner relationship management systems/ applications (i.e., providing a bank's customers with multiple touch points via telephones, Internet, fax machines, videos, mobile devices, ATMs, branches, the bank's agents).

• Building switching costs via extranets and proprietary software applications (i.e., Amazon's user-friendly and useful B2C website and Alibaba's B2B platform) so that a firm's customers or suppliers are reluctant to pay the costs in time, money, effort, and bear the inconvenience of switching to a company's competitors.

• Raising barriers to entry through improving operations or/and optimizing/flattening organizational structure by increasing the amount or the complexity of the technology required (i.e., Google's search engine and P & G's digitization strategy/efforts-P & G is working on digitizing almost every aspect of its operation to make it the world's most technologically enabled firm).

Benefits of Management Information System for Enterprise Management

A brief description of some reasons for deploying management information systems in companies and also the main benefits of MIS for management are as follows:

3. Motivation for MIS Deployment Nowadays, managing a company (unless talking about very small companies consisting of a couple of people) is almost unimaginable without a software support represented by information systems. The first enterprise information system that companies acquire is usually an ERP system, which can be in the form of accounting system, human resources management system or it can cover some other areas. Larger companies usually decide to buy some more complex ERP system solutions, capable of providing information support of everyday operations in all (or at least majority of) key areas. Typical representatives of such systems are SAP, Helios, MFG/PRO and others. These products, although often irreplaceable tools for everyday activities (including accounting, billing, stock details and others) aren't usually capable of providing sufficiently good support and data for a manager's work.

   Information stored in ERP systems are often well structured for everyday "operational" use, but unfortunately, often not very well structured for the needs of decision making support and other related tasks. Next step in the development of company's information systems is thus usually acquiring a management information system (MIS). It usually (at least partly) utilizes

the data of ERP system, but using the suitable representation, such data are given a new information value.

A high-quality MIS can provide the company some of the following benefits: • Comprehensible, quick and anytime-available (ad hoc) reporting • Reducing exhausting routine work and unblocking staff capacities for really creative activities • Properly structured information • Faster planning, capability of variant planning and modeling impacts of various market situations

4. Benefits for Reporting and Analyses Reporting can be conceived as a system of intercompany statements and reports, utilized not only for controlling and evaluating results achieved in past period. Outputs in the form of statements and reports should also be used for decisions about measures for improving company performance in the future. Unfortunately, even nowadays, many companies see the regular reporting as routine, exhausting and not very popular activity. We often see a "fight for data" every time after accounts final. These data are then filled into many more or less standardized tables.

   As the complexity and size of the company grows, the "bottom-up journey" of the data, ergo the process of getting the data from lower-level organization units to the top management, which should be the user of these reports, is more and more thorny and lengthy. As time goes on, most companies inevitably reach the point when reports created in the aforementioned way can no longer be used for efficient company management. The reason is simple – at the time the responsible managers get the reports, the information presented are usually obsolete or completely "dead" and unsatisfactory for the purposes of management. Especially nowadays, in the times of hard competition, hardly anyone can afford to manage a company using obsolete information.

   There are business branches where management based on "a couple of weeks old" information is almost the same as prophecy from crystal ball or so called "window method", which means a situation when a manager looks out of the window, thinks the things over a little bit and then makes a decision, without any supporting documents, based only on intuition.

   Properly designed MIS can be the right solution for the situation mentioned in the previous paragraph. The data can be available anytime, real-time, and, thanks to properly designed report structures and link to ERP system. MIS can provide necessary information virtually immediately, without any delays and without risk of mistakes, made by manual processing of huge amounts of data. The problems mentioned above also result in completely inadequate stress of controlling department staff in some companies. Although reporting is of course one

of controller's activities, it shouldn't be the main one, because contemporary controlling is far from being just reporting.

Unfortunately, we constantly face the fact that many companies see their controlling department as some kind of "reports factory" and the staff spend most of their time by absolutely routine activities, instead of doing analyses or providing really efficient management support. This means wasting of valuable potential, which could be used much better for further company development. Statistics show that "controllers" of some companies spend as much as 70 % of their time by processing data to requested form. In such situation, it's natural that controllers simply don't have enough time for creating qualified Summarys and making suggestions for the management, although it should be the most important part of their work. Therefore, valuable manpower and knowledge is suppressed by routine. To maximize the utility of available information, it's necessary to structure it properly. That's one of the reasons why the ERP systems aren't usually very suitable for higher level of management.

Management and decision-making requires, aside from precision and availability, also a proper structure of the information, enabling managers to perform required analyses and draw required Summarys from the information available. In this context, especially OLAP (On-Line Analytical Processing) technology based management information systems seem irreplaceable. Unlike classical "two dimensional" database tables, the data in OLAP database are represented in the form of so called "multidimensional data cubes", where every data can be described by a combination of elements of selected dimensions. As an example, we can assume this combination – "Customers total" (1st dimension – Customer), "Beer 10°" (2 nd dimension – Product), "Budget" (3 rd dimension – Plan version), "2010" (4 th dimension – Year), "January" (5 th dimension – Period) and "Volume sold (pieces)" (6 th dimension - Indicator). Proper combination of elements from various dimensions can then help to achieve creating various "views" of the same data, which is an invaluable feature for management support. Product manager can look at sales of different product lines for all customers, regional manager can look at sales in different regions for all products and so on. In doing so, creating a new "view" fully according to user's needs is a matter of a few moments, if the data structures are designed properly.

5. Benefits for Planning: Modern MIS can offer various tools for efficient planning support. Planning is a key activity in the majority of nowadays companies. A well known proverb says: "The key characteristic of winners is the ability to plan their success". Planning is a process of creating a plan, which should then represent an obligation for all the people in the

company, or a “standard” that everyone involved is trying to meet or even achieve results that are better than planned ones.

According to classical concept, plans can be classified by planning period length as long-term (strategic – usually a couple of years), mid-term (tactical – the most common example is annual plan) and short-term (operative – the planning period can be a couple of months in this case). The annual plan plays a key role in most companies, so we will focus on it further. Annual plan is a document that is created every year. It should cover all key areas of the enterprise (sales, production, investments, finance…). The importance of the annual plan is usually reflected in the laboriousness of its creation. In many companies, creating of annual plan is multi-level difficult process, which takes a couple of months to complete, requiring attention of company’s key employees. The work on creating the plan is often very exhausting, therefore it is quite common that after creation, the annual plan becomes a kind of “idol” for the next year and changes or corrections are made in very rare and exceptional cases.

Many companies then see their annual plan as a document which is created mainly because “it has to be made”, but its’ practical contribution is usually very little. Even a well-created annual plan may not be enough – companies in some branches simply can’t rely on the fact that plan based on the values from the end of a certain year will make any sense in the middle of the next year. It’s necessary to use other, more sophisticated approaches to planning, for example variant planning (creating multiple variants based on possible scenarios of future situation) or so called “sliding planning”, creating regular updated forecasts for the next period (i.e., 6 months) based on recent trends. Quality management information system can provide a substantial simplification of planning process, speed up the planning and provide tools for creating variants, updated forecasting or prognoses based on statistic methods. Thanks to proper tools and data structures, OLAP based MIS enables users to quickly create plans for the next period using for example extrapolation techniques. The data can then be further adjusted and “refined”, whereas OLAP data structure maintains data consistency in various views – if we change the total sum for products, the consistency is maintained and automatic calculation for “customers” view is done, to reflect product adjustment. Thanks to tools simplifying routine work and ensuring data consistency, it’s possible to perform easy recalculation according to updated future forecasts, create variant versions or use sliding planning techniques. Thanks to link between key areas (sales planning, cost and profit calculation, financial management…), the impact of changes made to sales plan immediately reflected in other areas.

6.13 Summary

Decision-making can be regarded as the cognitive process resulting in the selection of a belief or a course of action among several alternative possibilities. Every decision-making process produces a final choice that may or may not prompt action. Decision-making is the study of identifying and choosing alternatives based on the values and preferences of the decision maker.

The word "**decision** "is derived from the Latin word "decido" which means "A decision, therefore is

- A Settlement
- A fixed intuition to bringing to a conclusive result
- A judgment
- A resolution

Business Decision: Business decisions are those which are made in the process of conducting business to achieve its objective in a given situation

Types of Decision: Types of decision are based on the degree of knowledge about the outcome of the events which are yet to take place.

Certainty: If the manager has full knowledge of event or outcome then it is a situation of certainty.

Risk: If the manager has partial knowledge or probabilistic knowledge then it is decision under risk.

Uncertainty: If the manager does not have any knowledge, it is decision making under uncertainty MIS converts the uncertainty to risk and risk to certainty. The decision at the low level management is certain, at middle level of the management the decision is under risk and at the top level management the decision is in under uncertain.

Batch processing is a form of transaction processing. Batch processing involves processing several transactions at the same time, and the results of each transaction are not immediately available when the transaction is being entered, there is a time delay. Transactions are accumulated for a certain period (say for day) where updates are made especially after work. Online transaction processing is the form of transaction processing that processes data as it becomes available.

MIS is a combination of 3 English Letters:

M which stands for Management

I which stands for Information

S which stands for System

The MIS can also be defined as a system based on the database of the organization evolved for the purpose of providing information to the people in the organization.

Decision support system (**DSS**) is a computer-based information system that supports business or organizational decision-making activities. DSSs serve the management, operations, and planning levels of an organization and help to make decisions, which may be rapidly changing and not easily specified in advance. Decision support systems can be either fully computerized, human or a combination of both.

Expert System is used to provide expert advice to the end user for decision making. It is further used for providing critical information for executives and managers
Example: credit application advisor

Collaboration Technologies is dedicated to deliver good quality services in Software development. Collaborative software or groupware is an application software designed to help people involved in a common task to achieve goals. One of the earliest definitions of collaborative software is 'intentional group processes plus software to Support them.

6.14 Glossary

Computer Based Information System (CBIS) : This category of information system depends mainly on the computer for handling business applications. System analyst develops different types of information systems to meet variety of business needs. There is a class of system collectively known as computer based information system.
They can be classified as
· Transaction Processing System (TPS)
· Management Information System(MIS)
· Decision Support System (DSS)
· Office Automation System (OAS)

Transaction Processing System (TPS) : The most fundamental computer based
system in an organization pertains to the processing of business transactions. A transaction processing system can be defined as a system that captures, classifies, stores, maintains, updates and retrieves transaction data for record keeping and input to the other types of CBIS. Transaction Processing System is aimed at improving the routine business activities. A transaction is any event or activity that affects the whole organization. Placing order, billing customers, hiring of employees and depositing cheques are some of the common transactions. Types of transactions that occur vary from organization to organization but this is true

that all organizations process transaction as a major part of their daily business activities. Transaction Processing System provides speed and accuracy and can be programmed to follow routines without any variance.

Management Information System (MIS) : Data processing by computers has been extremely effective because of several reasons. The main reason is that huge amount of data relating to accounts and other transactions can be processed very quickly. MIS are more concerned with levels of management with information essential to the running of smooth business. This Information must be as relevant, timely, accurate, complete and concise as is economically feasible.

Decision Support System (DSS) : It is an information system that offers the kind of information that may not be predictable. Business professionals may need such information only once. These systems do not produce regularly scheduled management reports. Instead, they are designed to respond to wide range of requests. It is true that all the decisions in an organization are not of a recurring nature. Decision support systems assist managers, who make decisions that are not highly structured, often called unstructured or semi structured decision. The decision support systems support, but do not replace, judgments of managers.

Office Automation System (OAS) : Office Automation Systems are among the
newest and most rapidly expanding computer based information systems. They are being developed with the hope and expectation that they will increase the efficiency and productivity of office workers, typists, secretaries, administrative assistants, staff professionals, managers and others.

6.15 References

1. O'Brien, J (1999). *Management Information Systems – Managing Information Technology in the Internetworked Enterprise*. Boston: Irwin McGraw-Hill. ISBN 0-07-112373-3.
2. Laudon, Kenneth C.; Laudon, Jane P. (2009). *Management Information Systems: Managing the Digital Firm* (11 ed.). Prentice Hall/CourseSmart. p. 164.
3. *Transaction processing systems (TPS)* collect and record the routine transactions of an organization. Examples of such systems are sales order entry, hotel reservations, payroll, employee record keeping, and shipping.
4. Bidgoli, Hossein, (2004). The Internet Encyclopedia, Volume 1, John Wiley & Sons, Inc. p. 707.
5. Pant, S., Hsu, C., (1995), Strategic Information Systems Planning: A Review, Information Resources Management Association International Conference, May 21–24, Atlanta.

6.16 Further Readings

6. Taylor, Victoria. "Supply Chain Management: The Next Big Thing?". *Sept. 12, 2011*. Business Week. Retrieved 5 March 2014.
7. Lynn, Samara. "What is CRM?". PC Mag. Retrieved 5 March 2014.
8. Joshi, Girdhar (2013). *Management Information Systems*. New Delhi: Oxford University Press. p. 328. ISBN 9780198080992.
9. Laudon, K.,&Laudon, J. (2010). Management information systems: Managing the digital firm. (11th ed.). Upper Saddle River, NJ: Pearson Prentice Hall

6.17Model Questions

Q:1 Discuss about the Decision making and its levels.

Q:2Discuss in detail about TPS, MIS, EIS, DSS, ES and OAS

Q:3 Discuss in detail about the Collaboration Technologies.

Q:4 What are the various functional areas of Information Systems.

ANSWERS TO SELF ASSESSMENT QUESTIONS

A. 1
B. 1
C. 5
D. 5

CHAPTER 7

IS Security

Structure

7.0 Objectives

After studying this chapter you will be able to:

6. **Understand about IS security**
7. **Understand about IS vulnerability and computer crime**
8. **Understand how to protect information system**
9. **Understand about Disaster recovery plan**

7.1Information Systems security

Information security also referred as InfoSec, is defined as the practice of preventing unauthorized access, use, disclosure, disruption, modification, inspection, recording or destruction of **information**. It is a general term that can be used regardless of the form the data may take (e.g. electronic, physical). It is a set of strategies for managing the processes, tools and policies necessary to prevent, detect, document and counter threats to digital and non-digital information. Infosec responsibilities include establishing a set of business processes that will protect information assets regardless of how the information is formatted or whether it is in transit, is being processed or is at rest in storage. Infosec programs are built around the core objectives of the CIA triad: maintaining the confidentiality, integrity and availability of IT systems and business data. These objectives ensure that sensitive information is only disclosed to authorized parties (confidentiality), prevent unauthorized modification of data(integrity) and guarantee the data can be accessed by authorized parties when requested (availability).

Many large enterprises employ a dedicated security group to implement and maintain the organization's infosec program. Typically, this group is led by a chief information security officer. The security group is generally responsible for conducting risk management, a process through which vulnerabilities and threats to information assets are continuously assessed, and the appropriate protective controls are decided on and applied. The value of an organization lies within its information -- its security is critical for business operations, as well as retaining credibility and earning the trust of clients.

Threats to sensitive and private information come in many different forms, such as malware and phishing attacks, identity theft and ransom-ware. To deter attackers and mitigate vulnerabilities at various points, multiple security controls are implemented and coordinated as part of a layered defense in depth strategy. This should minimize the impact of an attack. To be prepared for a security breach, security groups should have an incident response plan (IRP) in place. This should allow them to contain and limit the damage, remove the cause and apply updated defense controls.

Information security processes and policies typically involve physical and digital security measures to protect data from unauthorized access, use, replication or destruction. These measures can include mantraps, encryption key management, network intrusion detection systems, password policies and regulatory compliance. A security audit may be conducted to evaluate the organization's ability to maintain secure systems against a set of established criteria.

IT security

Sometimes referred to as computer security, information technology security (IT security) is information security applied to technology (most often some form of computer system). It is worthwhile to note that a computer does not necessarily mean a home desktop. A computer is any device with a processor and some memory. Such devices can range from non-networked standalone devices as simple as calculators, to networked mobile computing devices such as smartphones and tablet computers. IT security specialists are almost always found in any major enterprise/establishment due to the nature and value of the data within larger businesses. They are responsible for keeping all of the technology within the company secure from malicious cyber attacks that often attempt to breach into critical private information or gain control of the internal systems.

Information assurance

The act of providing trust of the information, that the Confidentiality, Integrity and Availability (CIA) of the information are not violated, e.g. ensuring that data is not lost when critical issues arise. These issues include, but are not limited to: natural disasters, computer/server malfunction or physical theft. Since most information is stored on computers in our modern era, information assurance is typically

dealt with by IT security specialists. A common method of providing information assurance is to have an off-site backup of the data in case one of the mentioned issues arise.

Threats

Information security threats come in many different forms. Some of the most common threats today are software attacks, theft of intellectual property, identity theft, theft of equipment or information, sabotage, and information extortion. Most people have experienced software attacks of some sort. Viruses, worms, phishing attacks, and Trojan horses are a few common examples of software attacks. The theft of intellectual property has also been an extensive issue for many businesses in the IT field. Identity theft is the attempt to act as someone else usually to obtain that person's personal information or to take advantage of their access to vital information. Theft of equipment or information is becoming more prevalent today due to the fact that most devices today are mobile. Cell phones are prone to theft and have also become far more desirable as the amount of data capacity increases. Sabotage usually consists of the destruction of an organization′s website in an attempt to cause loss of confidence on the part of its customers. Information extortion consists of theft of a company′s property or information as an attempt to receive a payment in exchange for returning the information or property back to its owner, as with ransomware.

Responses to threats

Possible responses to a security threat or risk are:

- reduce/mitigate – implement safeguards and countermeasures to eliminate vulnerabilities or block threats
- assign/transfer – place the cost of the threat onto another entity or organization such as purchasing insurance or outsourcing
- accept – evaluate if cost of countermeasure outweighs the possible cost of loss due to threat
- ignore/reject – not a valid or prudent due-care response

How to Secure Information Systems:

To secure information system the main objective is to protect an organization’s information by reducing the risk of loss of confidentiality, integrity and availability of that information to an acceptable level. A good information security program involves two major elements, risk analysis and risk management. In the risk analysis phase, an inventory of all information systems is taken. For each system, its value to the organization is established and the degree to which the organization is exposed to risk is determined. Risk management, on the other hand, involves selecting the controls and security measures that reduce the organization’s exposure to risk to an acceptable level. To be effective, efficient and reflect common sense,

risk management must be done within a security framework where information security measures are complemented by computer, administrative, personnel and physical security measures (see Figure):

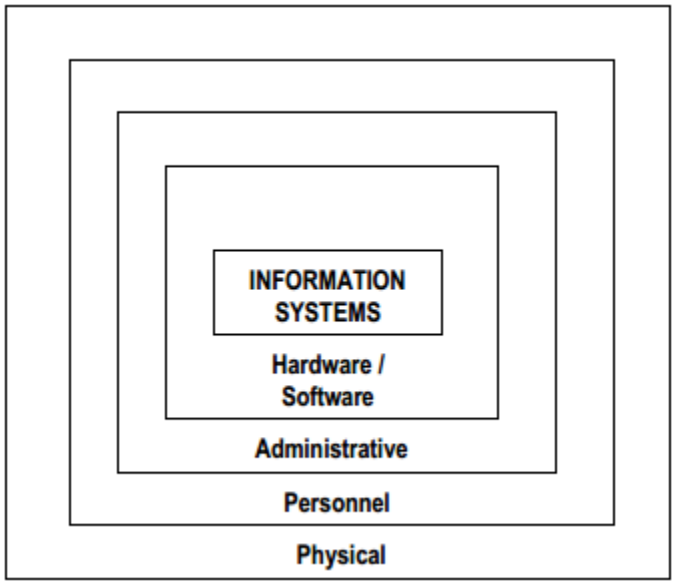


Figure: Complementary Layers of Information Security

Risk management becomes a senior management issue. A balance has to be reached between the value of the information to the organization on the one hand and the cost of the personnel, administrative and technological security measures on the other hand. The security measures put in place need to be less expensive than the potential damage caused by the loss of confidentiality, integrity and availability of the information. Many formal risk analysis methodologies on the market require technical expertise in the area of information technology and relevant controls and availability of precise threat frequencies that may be beyond the reach of many audit offices, at least initially. The objective is to build up over time the necessary expertise and resources.

Information Security Framework

Information security is one element of a security infrastructure and, as such, should not be examined in a vacuum. There should be a framework of security policies dealing with all aspects of physical security, personnel security and information security. There should be clear roles and responsibilities for users, security officers and the Information Systems Steering Committee. An information security program should include all aspects of the sensitivity of corporate information, including confidentiality, integrity and availability. A pro/gram of security awareness should be in place reminding all staff of the possible risks and exposures and of their responsibilities as custodians of corporate information.

Referring to Figure above, information security is a set of measures at the physical, personnel, administrative, computer and information system levels. They must all work together. Information security is good management control and shortcomings at any level can threaten the security at other levels. If personnel security policies, for instance, are not well designed and implemented, information security could become very costly or almost impossible to support. On the other hand, minimal measures at all levels should ensure a minimum of protection to the information, provided the security risk is reasonable and accepted by management. There are also situations where security measures at one level may compensate for security weaknesses elsewhere. Encryption, for instance, adds an extra layer of protection for data confidentiality and integrity even in cases where physical, personnel or administrative security measures may be weak. Encryption remains one last defense to help prevent a breach of confidentiality or of integrity.

In planning for information security, the value of the information to management and the volume of that information relative to other types of information have to be balanced against the basic security limitations of the medium. In many government departments, unless there are extreme requirements for carrying top secret information on a suitably protected laptop, the information should simply be created and carried otherwise. For those departments, the cost and the constraints of the appropriate security controls and measures may just not be acceptable, given the small volume of information that needs such protection.

7.2 IS vulnerability and computer crime

Information systems vulnerabilities cover more territory than just personal losses. Computer information systems are vulnerable to physical attacks, electronic hacking, and natural disasters. With computer information systems serving as the vital life blood of many organizations, managers must be aware of the both the risks and the opportunities to minimize the risks to information systems.
Discussion is divided into types of computer crime, information systems and technology vulnerabilities, and ways to manage the risks.

Types of Computer Crime

Typically, computer crime can be categorized by the type of activity which occurs. Four basic categories are utilized in describing computer crime. These are: theft, fraud, copyright infringement, and attacks.

Theft: Theft in computer crime may refer to either unauthorized removal of physical items such as hardware or unauthorized removal or copying of data or information. It is well known that laptop computers are targeted at airports and restaurants. The prize garnered with theft of a laptop is usually the data or information such as passwords for corporate systems contained on the laptops rather than the hardware.

Fraud: Fraud on the Internet may run the gamut from credit card offers which are utilized only to capture personal information, to investor postings which promote a stock or investment offer to encourage investment which will benefit the person posting the information, to medical and pharmaceutical -related sites which purport to provide correct medical advice or sell altered medications.

Copyright infringement: The Internet has provided a unique opportunity and environment for copyright infringement. This type of computer crime encompasses use of software, music, etc which is not appropriately acquired (purchased). Software piracy occurs more easily with the ability to post files for downloading all over the world. However, another more costly copyright infringement occurs when trademarks and logos of corporations are posted on non-authorized web sites. Some criminals utilize the trademarks and logos to appear to be a legitimate site to perpetrate fraud. Many corporations have employees or consulting contractors who constantly crawl the web to sniff out illegal usage of trademarks and logos.

Attacks on organizations and individuals: Attacks on organizational information systems may be either physical or logical. There are several instances of web sites, products, and individuals being libeled or attacked by individuals or groups. One of the classic examples was the attack on Proctor and Gamble as an occult organization. AOL and other ISPs cooperate fully with criminal justice systems to reveal identities of those deploying web sites of question**.**

Denial of Service Attacks (DoS) target specific web sites and associated servers. Some of the newsworthy examples of DoS during 2000 - 2001 have occurred at Microsoft.com, eBay.com, and Amazon.com. Web servers and connections can only handle so much traffic so Denial of Service (DoS) usually take the form of one of two ways:

– Coordinated attack (typically from unsuspecting desktops) to a particular IP address or URL requesting a page – overwhelms server and DoS occurs

– Attack sends incomplete packets so that traffic gets jammed with requests for re-send.

Information Systems and Technology Vulnerabilities

There are several classes of activities which may also harm information systems and supporting technology. These activities may result in criminal charges depending upon the circumstances and impact on information systems. Currently, these activities fall within classes of viruses, worms, Trojan Horse, time bomb, logic bomb, and trapdoors.

Viruses: A virus is a program with intent to harm or render a computer system useless. The virus method of attack is to attach itself to specific files such as data files. It is not a free standing program. It copies itself when the infected file is executed.

A virus can damage data, delete files, erase your hard drive, or just cause annoying screen displays or sounds. Viruses may hide within macros of Word or Excel documents. Some viruses are programmed to trigger execution on a particular date or time. Viruses do not cause hardware damage. Viruses spread from file to file. There are thousands of documented viruses!!!! Some recent examples of viruses include the Melissa, Chernobyl, and Michelangelo.

Most virus protection software provides monthly updates to ensure that the computer system is covered from recent virus discoveries. Two of the more popular versions of virus protection include Norton (Symantec) and McAfee.

Worms: Worms are another destructive program designed to create instability information systems and supporting technology. Worms differ from viruses in that a worm is a free standing program. A worm executes on its own functionality. Worms spread from computer system to computer system rather than from file to file.

Examples of notorious worms include the July and August, 2001 attack of CODE RED on IIS servers. IIS (Internet Information Services) is part of the Microsoft Windows Server operating system which provides internet connectivity. Servers including federal government web sites, Qwest DSL servers, and other corporate or governmental sites were hit.

A worm can reply to e-mails while attaching itself to the e-mail; can destroy File Allocation System (FAT) on Windows systems and other similar attacks on other files systems on hard drives. Because worms are free standing, they can spread on their own and do not require human intervention to spread. Thus, in some ways, worms are more lethal than viruses.

Trojan Horse: This software derives its name from the Greek mythology depicting war activity between the Greeks and Trojans of Troy. The Greeks pretended to depart the besieged Troy but left behind a giant wooden horse as a “gift”. The Trojans brought the horse within the gates of Troy and Greek warriors were hidden in the horse. The Greek warriors then captured Troy. Therefore, the Trojan Horse appears to have one function but in reality does something else.

Typically, a Trojan horse performs something destructive while the person at the keyboard thinks they are downloading an animation or some other file. The Trojan Horse commonly either loads a software program

to be utilized in a later Denial of Service attack or reads your passwords, credit card numbers, etc., saved within your system. This vital information is later used to make purchases or other criminal activities.

In August of 2001, a particularly damaging Trojan Horse named the Trojan Offensive has been reported. It damages the Registry of Windows operating system so that the system is trashed.

Time bomb: These are software attacks that are designed to occur at a predetermined time or date. The difference between a time bomb and a virus such as the Michelangelo is that technically the time bomb does not spread. It impacts on the system upon which it has been loaded.

Logic bomb: Logic bombs are software attacks that triggered by a predetermined event. The most common logic bombs occur when information technology employees are laid off from employment. Then, for example, billing systems go awry when an employee id number is no longer on the payroll database.

Trapdoor: Trapdoors are system entrances that circumvent security system. These are hidden logins or administrative user definitions added by system developers for unscrupulous reasons. Trapdoors allow an unauthorized or unknown user to control a computer system. Trapdoors are typically only aimed at servers or mainframe corporate systems.

ACTIVITIES

1. List down the threats to information systems and explain how they are existing?
2. How will you create a control environment for information systems?
3. Explain the various security hazards faced by an information system.

7.3 Protecting information system

As businesses and individual depend more and more on computers for storage of data and production of information, the capability to conduct business is vulnerable to attacks on computer systems. In addition to malicious attacks, there are always accidents and careless behaviors. Are backups really current? What happens if a pipe breaks causing flooding? These are examples of careless behaviors and accidents. There are electrical storms, fires, tornadoes, hurricanes, earthquakes, and other nature related events for which information systems must be prepared.

Managing the risk of computer crime: Protecting systems and data with passwords, encryption, auditing software, and access logs is vital. These logical protections must be reviewed and analyzed in order to ensure the system has not been penetrated.

Physical entry to computer systems must be protected. Locations of computer systems are often times hidden from the public knowledge in order to make the systems more difficult to find. Card key systems and login/logout of entry and exit to computer systems should be a regular business procedure.

Managing the risk of fraud: Various regulatory agencies and criminal justice units are using the web to locate fraud. Regulatory agencies such as the FDA review medical and pharmaceutical related sites. Also, professional organizations such as the AMA and credible health care organizations such as Mayo Clinic provide correct information on sponsored sites.

Managing the risk of copyright infringement: Businesses are learning to crawl the web for attacks and copyright infringements. Disney and Coca Cola are particularly vigilant. Bloomberg Financial was the victim of a spoof site which created quite a stir in the financial community.

Managing the risk of technology vulnerabilities: The major activity deployed by businesses to protect computer systems and data from electronic intrusion is the utilization of firewalls and virus protection software.

Firewalls are utilized to establish a barrier between the business computer systems and the outside world. Firewalls may be a combination of hardware and software or it may be software only. A firewall filters or restricts access externally to enter system and access internally to exist system. The usual implementation of a firewall is to place a barrier between a computer system and the Internet. There are four basic actions firewall software can take when communication is attempted. These are:

Piece of information can be dropped

– Alert issued

– Returned to sender

– Log action & receipt of information.

Communications are classified as inbound and outbound. When an attempt is made from another system to connect to your system, this is considered an inbound connection. And, when your system is attempting to connect outside of local, this is considered to be an outbound connection. Once the connection is allowed, bi-directional communications can occur without filtering. Therefore, firewalls must be set too restrict access both inbound and outbound.

Virus protection software must be kept current as new viruses are developed on what seems like a daily basis. Some businesses have policies that limit acceptance of attachments to e-mails, prohibit the use of disks prepared on non-business systems, and restrict downloading files from the Internet from trusted sites only.

7.4 Disaster Recovery Planning

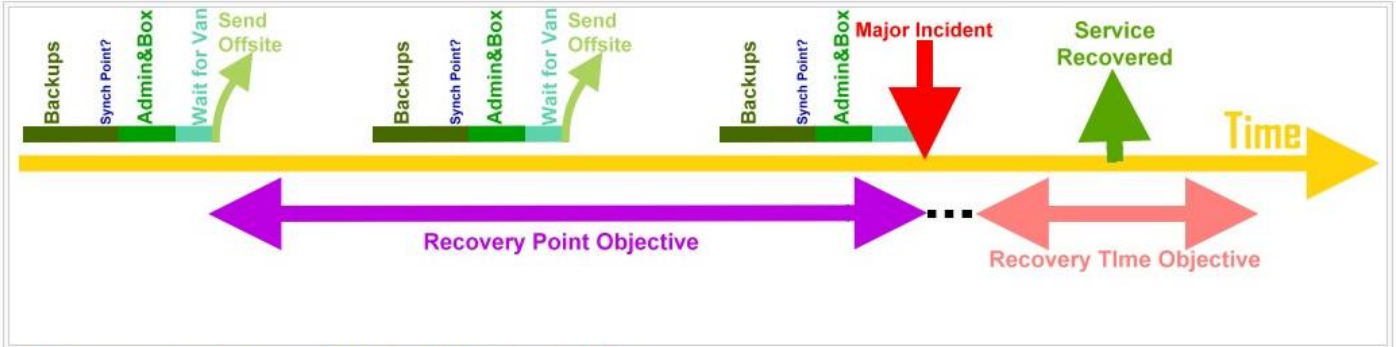
A **disaster recovery plan** (DRP) is a documented process or set of procedures to recover and protect a business IT infrastructure in the event of a disaster. Such plan, ordinarily documented in written form, specifies procedures an organization is to follow in the event of a disaster. It is "a comprehensive statement of consistent actions to be taken before, during and after a disaster." The disaster could be natural, environmental or man-made. Man-made disasters could be intentional (for example, an act of a terrorist) or unintentional (that is, accidental, such as the breakage of a man-made dam).

Given organizations' increasing dependency on information technology to run their operations, a disaster recovery plan, sometimes erroneously called a **continuity of operations plan (COOP)**, is increasingly associated with the recovery of information technology data, assets, and facilities.

Objective:

Organizations cannot always avoid disasters, but with careful planning the effects of a disaster can be minimized. The objective of a disaster recovery plan is to minimize downtime and data loss.[3] The primary objective is to protect the organization in the event that all or part of its operations and/or computer services is rendered unusable. The plan minimizes the disruption of operations and ensures that some level of organizational stability and an orderly recovery after a disaster will prevail. Minimizing downtime and data loss is measured in terms of two concepts: the Recovery Time Objective (RTO) and the Recovery Point Objective (RPO).

The recovery time objective is the time within which a business process must be restored, after a major incident (MI) has occurred, in order to avoid unacceptable consequences associated with a break in business continuity. The recovery point objective (RPO) is the age of files that must be recovered from backup storage for normal operations to resume if a computer, system, or network goes down as a result of a MI. The RPO is expressed backwards in time (that is, into the past) starting from the instant at which the MI occurs, and can be specified in seconds, minutes, hours, or days. The recovery point objective (RPO) is thus the maximum acceptable amount of data loss measured in time. It is the age of the files or data in backup storage required to resume normal operations after the MI.



A DR plan illustrating the chronology of the RPO and the RTO with respect to the MI.

Relationship to the Business Continuity Plan

According to the SANS institute, the Business Continuity Plan (BCP) is a comprehensive organizational plan that includes the disaster recovery plan. The Institute further states that a Business Continuity Plan (BCP) consists of the five component plans:

- Business Resumption Plan
- Occupant Emergency Plan
- Continuity of Operations Plan
- Incident Management Plan
- Disaster Recovery Plan

The Institute states that the first three plans (Business Resumption, Occupant Emergency, and Continuity of Operations Plans) do not deal with the IT infrastructure. They further state that the Incident Management Plan (IMP) does deal with the IT infrastructure, but since it establishes structure and procedures to address cyber attacks against an organization's IT systems, it generally does not represent an agent for activating the Disaster Recovery Plan, leaving The Disaster Recovery Plan as the only BCP component of interest to IT.

Disaster Recovery Institute International states that disaster recovery is the area of business continuity that deals with *technology* recovery as opposed to the recovery of business operations.

Benefits of Disaster Recovery Plan:

Types of plans

There is no one right type of disaster recovery plan, nor is there a one-size-fits-all disaster recovery plan. However, there are three basic strategies that feature in all disaster recovery plans: (1) preventive measures, (2) detective measures, and (3) corrective measures. Preventive measures will try to prevent a disaster from occurring. These measures seek to identify and reduce risks. They are designed to mitigate or

prevent an event from happening. These measures may include keeping data backed up and off site, using surge protectors, installing generators and conducting routine inspections. Detective measures are taken to discover the presence of any unwanted events within the IT infrastructure. Their aim is to uncover new potential threats. They may detect or uncover unwanted events. These measures include installing fire alarms, using up-to-date antivirus software, holding employee training sessions, and installing server and network monitoring software. Corrective measures are aimed to restore a system after a disaster or otherwise unwanted event takes place. These measures focus on fixing or restoring the systems after a disaster. Corrective measures may include keeping critical documents in the Disaster Recovery Plan or securing proper insurance policies, after a "lessons learned" brainstorming session.

A disaster recovery plan must answer at least three basic questions:

(1) What is its objective and purpose?

(2) Who will be the people or teams who will be responsible in case any disruptions happen?

(3) What will these people do (the procedures to be followed) when the disaster strikes?

The primary objectives of a Discovery Recovery Plan are to guide an organization in the event of a disaster and to effectively reestablish critical business operations within the shortest possible period of time with a minimal loss of data. The goals of the planning project are to assess current and anticipated vulnerabilities, define the requirements of the business and IT communities, design and implement risk mitigation procedures and provide the organization with a plan that will enable it to react quickly and efficiently at the time of a disaster. It must be remembered that disaster recovery planning is not limited to the IT community. It is equally a business issue, since there may be systems or process used that are not widely know outside the group using them. At a minimum, the DR Plan must address the processing needs of the business community; this means key members of the business community must be involved in the planning to insure that these needs are adequately documented and understood. IT personnel should not make assumptions as to whole systems are critical to the business community. Whether the DR is a stand alone effort or is being done in conjunction with an overall BC effort, the following are some of the points which need to be addressed:

- A project team must be selected that incorporates an adequate balance between IT and business community members to ensure that the resulting Plan will cover the requirements of both the IT and business communities.
- The recovery requirements of the business and IT communities must be defined and agreed upon. Furthermore, they should be posted somewhere accessible to everyone in the organization (such as the company intranet). This type of visibility helps ensure that people realize the importance in the effort, and their role in its success.

- Solutions to fit the requirements of the business and IT communities, including risk identification, analysis and mitigation, must be designed. For more information on risk management as it relates to BC/DR .
- The final Plan, which incorporates those solution, must be easy to understand(by people unfamiliar with the systems and under stress), put into practice, and easy to maintain.
- The final Plan needs to be integrated with any other existing plans – including other DR Plans, Emergency Management Plans, Evacuation Plans etc. this is usually part of an overall BC effort, but in case of a standalone DR effort this should still be addressed.
- A process needs to be developed to keep the plan up to date, representing the true business and computing environments at all times. It must also be understood that disaster recovery planning is a highly complex and time-consuming activity and requires a firm commitment from management to expend the man hours and funds necessary to achieve success. In addition, implementing solutions designed to mitigate risk often necessitates major expenditures.

SELF ASSESSMENT QUESTIONS

Please mark the correct option

A. The person generally responsible for the program design strategy, standards, and construction is called a(n):
   1. program librarian
   2. backup chief programmer
   3. network designer
   4. chief programmer
   5. systems analyst

B. Which one of the tests is performed on a subset of a program?
   1. subset test
   2. unit test
   3. stub test
   4. program test
   5. system test

C. Which of the following is the last phase of the Systems Construction phase? 1. build and test databases
   2. write and test new programs
   3. prepare conversion plan
   4. build and test networks
   5. none of the above

D. Which installation strategy is a variation on the abrupt and parallel conversion?
   1. location conversion
   2. staged conversion
   3. partial conversion
   4. hierarchical conversion
   5. none of the above

E. Which of the following is(are) the task(s) of the Systems Conversion phase? 1. prepare conversion plan
2. train users
3. convert to new system
4. build and test networks
5. none of the above

F. Which of the tests is a final system test performed by end users using real data over an extended period of time?
1. final test
2. complete test
3. systems acceptance test
4. parallel test
5. none of the above

7.5Auditing

The term audit is derived from the Latin term ‘audire,’ which means to hear. In early days an auditor used to listen to the accounts read over by an accountant in order to check them. The original objective of auditing was to detect and prevent errors and frauds. Auditing evolved and grew rapidly after the industrial revolution in the 18th century. With the growth of the joint stock companies the ownership and management became separate. The shareholders who were the owners needed a report from an independent expert on the accounts of the company managed by the board of directors who were the employees. The objective of audit shifted and audit was expected to ascertain whether the accounts were true and fair rather than detection of errors and frauds.

Spicer and Pegler defined "Auditing is such an examination of books of accounts and vouchers of business, as will enable the auditors to satisfy himself that the balance sheet is properly drawn up, so as to give a true and fair view of the state of affairs of the business and that the profit and loss account gives true and fair view of the profit/loss for the financial period, according to the best of information and explanation given to him and as shown by the books; and if not, in what respect he is not satisfied."

Prof. L.R.Dicksee defined "auditing is an examination of accounting records undertaken with a view to establish whether they correctly and completely reflect the transactions to which they relate.

FEATURES OF AUDITING

a. Audit is a systematic and scientific examination of the books of accounts of a business;

b. Audit is undertaken by an independent person or body of persons who are duly qualified for the job.

c Audit is a verification of the results shown by the profit and loss account and the state of affairs as shown by the balance sheet.

d. Audit is a critical review of the system of accounting and internal control.

e. Audit is done with the help of vouchers, documents, information and explanations received from the authorities.

f. The auditor has to satisfy himself with the authenticity of the financial statements and report that they exhibit a true and fair view of the state of affairs of the concern.

g The auditor has to inspect, compare, check, review, scrutinize the vouchers supporting the transactions and examine correspondence, minute books of share holders, directors, Memorandum of Association and Articles of association etc., in order to establish correctness of the books of accounts.

OBJECTIVES OF AUDITING

There are two main objectives of auditing. The primary objective and the secondary or incidental objective are:

a. **Primary objective –** as per Section 227 of the Companies Act 1956, the primary duty (objective) of the auditor is to report to the owners whether the balance sheet gives a true and fair view of the Company's state of affairs and the profit and loss A/c gives a correct figure of profit of loss for the financial year.

b. **Secondary objective –** it is also called the incidental objective as it is incidental to the satisfaction of the main objective. The incidental objectives of auditing are:

i. Detection and prevention of Frauds, and

ii. Detection and prevention of Errors.

Detection of material frauds and errors as an incidental objective of independent financial auditing flows from the main objective of determining whether or not the financial statements give a true and fair view. As the Statement on auditing Practices issued by the Institute of Chartered Accountants of India states, an auditor should bear in mind the possibility of the existence of frauds or errors in the accounts under audit since they may cause the financial position to be misstated.

Fraud refers to intentional misrepresentation of financial information with the intention to deceive. Frauds can take place in the form of manipulation of accounts, misappropriation of cash and misappropriation of goods. It is of great importance for the auditor to detect any frauds, and prevent their recurrence. Errors refer to unintentional mistake in the financial information arising on account of ignorance of accounting principles i.e. principle errors, or error arising out of negligence of accounting staff i.e. Clerical errors.

BASIC PRINCIPLES OF AUDIT

AAS-1 describes the basic principles, which govern the auditor's professional responsibilities and which should be complied with whenever an audit is carried out. These are:-

1. Integrity, objectivity and independence:

The auditor should be straightforward, honest and sincere in his approach to his professional work. He must be fair and must not allow prejudice or bias to override his objectivity. He should maintain an impartial attitude and appear to be free of any interest which might be regarded. Whatever it's actual effect, as being incompatible with integrity and objectivity.

2. Confidentiality:

The auditor should respect the confidentiality of information acquired in the course of his work and should not disclose any such information to a third party without specific authority or unless there is legal or professional duty to disclose. It is remarked that an auditor should keep his ears and eyes open but his mouth shut.

3. Skill and competence:

The audit should be performed and the report prepared with due professional care by persons who have adequate training, experience and competence. This can be acquired through a combination of general education, technical knowledge obtained through study and formal courses concluded by a qualifying examination recognized for this purpose and practical experience under proper supervision.

4. Work performed by others:

When the auditor delegates work to assistant* or uses work performed by other auditors or experts, he will continue to be responsible for forming and expressing his opinion on the financial information. At the same time he is entitled to rely on work performed by others provided he exercises adequate skills and care and is not aware of any reason to believe that he should not have relied. The auditor should carefully direct, supervise & review work delegated by assistants. He should obtain reasonable assurance that work performed by other auditors or experts is adequate for this purpose.

5. Documentation:

The auditor should document matters, which are important in providing evidence that the audit was carried out in accordance with the basic principles.

6. Planning:

The auditor should plan his work to enable him to conduct an effective audit in an efficient and timely manner. Plans should be based on knowledge of client's business. They should be further developed and revised, if required, during the course of audit.

7. Audit evidence:

The auditor should obtain sufficient appropriate audit evidence through the performance of compliance and substantive test procedure. It will enable him to draw reasonable Summarys there from on which he has to base his opinion on the financial information.

8. Accounting system & internal control:

The auditor should gain an understanding of the accounting system and related internal controls. He should study and evaluate the operation of those internal controls upon which he wishes to rely in determining the nature, timing and extent of other audit procedures.

9. Audit Summarys and reporting:

The auditor should review and assess the Summarys drawn from the audit evidence obtained and from his knowledge of business of the entity as the basis for the expression of his opinion on the financial information. The audit report should contain a written expression of opinion of the financial information. It should comply with the legal requirements. In case of a qualified opinion, adverse opinion or disclaimer of opinion is given or reservation on any matter is to be made reasons thereof.

AUDIT TYPES

MEANING:

Audit is not legally obligatory for all types of business organizations or institutions. On this basis audits may be of two broad categories i.e., audit required under law and voluntary audits.

(i) Audit required under law: The organizations which require audit under law are the following:

(a) Companies governed by the Companies Act, 1956;

(b) Banking companies governed by the Banking Regulation Act,1949;

(c) Electricity supply companies governed by the Electricity supply Act, 1948;

(d) Co-operative societies registered under the co-operative Societies Act, 1912;

(e) Public and charitable trusts registered under various Religious and Endowment Acts;

(f) Corporations set up under an Act of parliament or State Legislature such as the Life Insurance Corporation of India.

(g) Specified entities under various sections of the Income-tax Act, 1961.

(ii) In the voluntary category are the audits of the accounts of proprietary entities, partnership firms, Hindu undivided families, etc. in respect of such accounts, there is no basic legal requirement of audit. Many of such enterprises as a matter of internal rules require audit. Some may be required to get their accounts audited on the directives of Government for various purposes like sanction of grants, loans, etc. But the important motive for getting accounts audited lies in the advantages that follow from an independent

professional audit. This is perhaps the reason why large numbers of proprietary and partnership business get their accounts audited.

INTERIM AUDIT:

An audit that is taken up between two annual audits is called an Interim Audit. A specific date, as per the client's requirement is taken into account, e.g. 30th September, 31st December, etc. a trial balance is drawn and verified with a view to prepare financial statement. Financial statement are prepared and authenticated for the interim audit period. Assets and liabilities are verified for interim balance sheet purposes. Independence is considered less independent than the statutory Auditor; generally an employee of the enterprise will be the internal auditor. In the interim audit no format is prescribed. It depends on the nature of work, coverage and audit observations.

CONTINUOUS AUDIT:

A continuous audit is one in which the auditor's staff is engaged continuously in checking the accounts of the client, during the whole year round or when for the purpose, the staff attends at quite frequent intervals say weekly basis during the financial period. A continuous audit is preferred for the following reasons:

i. It makes it possible for the management to exercise a stricter control over the accounts in as much as one is able to check sooner the causes of any errors of frauds uncovered by such an audit.

ii. The frequent attendance by the staff deters persons so inclined, from committing a fraud.

iii. The accounting staff of the client is motivated to keep the books of account up-to-day.

Risk Management

Risk management is the process of identifying vulnerabilities and threats to the information resources used by an organization in achieving business objectives, and deciding what countermeasures, if any, to take in reducing risk to an acceptable level, based on the value of the information resource to the organization."

It is the process of identifying, assessing and controlling threats to an organization's capital and earnings. These threats, or risks, could stem from a wide variety of sources, including financial uncertainty, legal liabilities, strategic management errors, accidents and natural disasters. IT security threats and data-related risks, and the risk management strategies to alleviate them, have become a top priority for digitized companies. As a result, a risk management plan increasingly includes companies' processes for identifying and controlling threats to its digital assets, including proprietary corporate data, a customer's personally identifiable information and intellectual property.

Risk management strategies and processes

All risk management plans follow the same steps that combine to make up the overall risk management process:

- **Risk identification.** The company identifies and defines potential risks that may negatively influence a specific company process or project.
- **Risk analysis.** Once specific types of risk are identified, the company then determines the odds of it occurring, as well as its consequences. The goal of the analysis is to further understand each specific instance of risk, and how it could influence the company's projects and objectives.
- **Risk assessment and evaluation.** The risk is then further evaluated after determining the risk's overall likelihood of occurrence combined with its overall consequence. The company can then make decisions on whether the risk is acceptable and whether the company is willing to take it on based on its risk appetite.
- **Risk mitigation**. During this step, companies assess their highest-ranked risks and develop a plan to alleviate them using specific risk controls. These plans include risk mitigation processes, risk prevention tactics and contingency plans in the event the risk comes to fruition.
- **Risk monitoring**. Part of the mitigation plan includes following up on both the risks and the overall plan to continuously monitor and track new and existing risks. The overall risk management process should also be reviewed and updated accordingly.

7.6 Summary

Information security is one element of a security infrastructure and, as such, should not be examined in a vacuum. There should be a framework of security policies dealing with all aspects of physical security, personnel security and information security. There should be clear roles and responsibilities for users, security officers and the Information Systems Steering Committee. An information security program should include all aspects of the sensitivity of corporate information, including confidentiality, integrity and availability. A program of security awareness should be in place reminding all staff of the possible risks and exposures and of their responsibilities as custodians of corporate information.

computer crime can be categorized by the type of activity which occurs. Four basic categories are utilized in describing computer crime. These are: theft, fraud, copyright infringement, and attacks.

Theft: Theft in computer crime may refer to either unauthorized removal of physical items such as hardware or unauthorized removal or copying of data or information. It is well known that laptop computers are targeted at airports and restaurants. The prize garnered with theft of a laptop is usually the data or information such as passwords for corporate systems contained on the laptops rather than the hardware.

Fraud: Fraud on the Internet may run the gamut from credit card offers which are utilized only to capture personal information, to investor postings which promote a stock or investment offer to encourage investment which will benefit the person posting the information, to medical and pharmaceutical -related sites which purport to provide correct medical advice or sell altered medications.

Copyright infringement: The Internet has provided a unique opportunity and environment for copyright infringement. This type of computer crime encompasses use of software, music, etc which is not appropriately acquired (purchased). Software piracy occurs more easily with the ability to post files for downloading all over the world. However, another more costly copyright infringement occurs when trademarks and logos of corporations are posted on non-authorized web sites. Some criminals utilize the trademarks and logos to appear to be a legitimate site to perpetrate fraud. Many corporations have employees or consulting contractors who constantly crawl the web to sniff out illegal usage of trademarks and logos.

A **disaster recovery plan** (DRP) is a documented process or set of procedures to recover and protect a business IT infrastructure in the event of a disaster. Such plan, ordinarily documented in written form, specifies procedures an organization is to follow in the event of a disaster. It is "a comprehensive statement of consistent actions to be taken before, during and after a disaster." The disaster could be natural, environmental or man-made. Man-made disasters could be intentional (for example, an act of a terrorist) or unintentional (that is, accidental, such as the breakage of a man-made dam).Three basic strategies that feature in all disaster recovery plans: (1) preventive measures, (2) detective measures, and (3) corrective measures. Preventive measures will try to prevent a disaster from occurring.

A disaster recovery plan must answer at least three basic questions:

(1) What is its objective and purpose?

(2) Who will be the people or teams who will be responsible in case any disruptions happen?

(3) What will these people do (the procedures to be followed) when the disaster strikes?

Audit is not legally obligatory for all types of business organizations or institutions. On this basis audits may be of two broad categories i.e., audit required under law and voluntary audits.

(i) Audit required under law: The organizations which require audit under law are the following:

(a) Companies governed by the Companies Act, 1956;

(b) Banking companies governed by the Banking Regulation Act,1949;

(c) Electricity supply companies governed by the Electricity supply Act, 1948;

(d) Co-operative societies registered under the co-operative Societies Act, 1912;

(e) Public and charitable trusts registered under various Religious and Endowment Acts;

(f) Corporations set up under an Act of parliament or State Legislature such as the Life Insurance Corporation of India.

(g) Specified entities under various sections of the Income-tax Act, 1961.

7.7 Glossary

Viruses: A virus is a program with intent to harm or render a computer system useless. The virus method of attack is to attach itself to specific files such as data files. It is not a free standing program. It copies itself when the infected file is executed.

Worms: Worms are another destructive program designed to create instability information systems and supporting technology. Worms differ from viruses in that a worm is a free standing program. A worm executes on its own functionality. Worms spread from computer system to computer system rather than from file to file.

Trojan Horse: This software derives its name from the Greek mythology depicting war activity between the Greeks and Trojans of Troy. The Greeks pretended to depart the besieged Troy but left behind a giant wooden horse as a "gift". The Trojans brought the horse within the gates of Troy and Greek warriors were hidden in the horse. The Greek warriors then captured Troy. Therefore, the Trojan Horse appears to have one function but in reality does something else.

Time bomb: These are software attacks that are designed to occur at a predetermined time or date. The difference between a time bomb and a virus such as the Michelangelo is that technically the time bomb does not spread. It impacts on the system upon which it has been loaded.

Logic bomb: Logic bombs are software attacks that triggered by a predetermined event. The most common logic bombs occur when information technology employees are laid off from employment. Then, for example, billing systems go awry when an employee id number is no longer on the payroll database.

Trapdoor: Trapdoors are system entrances that circumvent security system. These are hidden logins or administrative user definitions added by system developers for unscrupulous reasons. Trapdoors allow an

unauthorized or unknown user to control a computer system. Trapdoors are typically only aimed at servers or mainframe corporate systems.

7.8 References

O'Brien, J (1999). *Management Information Systems – Managing Information Technology in the Internetworked Enterprise*. Boston: Irwin McGraw-Hill. ISBN 0-07-112373-3.

Laudon, Kenneth C.; Laudon, Jane P. (2009). *Management Information Systems: Managing the Digital Firm* (11 ed.). Prentice Hall/CourseSmart. p. 164.

Bidgoli, Hossein, (2004). The Internet Encyclopedia, Volume 1, John Wiley & Sons, Inc. p. 707.

Pant, S., Hsu, C., (1995), Strategic Information Systems Planning: A Review, Information Resources Management Association International Conference, May 21–24, Atlanta.

7.9 Further Readings

Taylor, Victoria. "Supply Chain Management: The Next Big Thing?". *Sept. 12, 2011*. Business Week. Retrieved 5 March 2014.

Lynn, Samara. "What is CRM?". PC Mag. Retrieved 5 March 2014.

Joshi, Girdhar (2013). *Management Information Systems*. New Delhi: Oxford University Press. p. 328. ISBN 9780198080992.

Laudon, K.,&Laudon, J. (2010). Management information systems: Managing the digital firm. (11th ed.). Upper Saddle River, NJ: Pearson Prentice Hall

7.10 Model Questions

Q:1 Explain in detail about IS security.

Q:2 Explain about the IS vulnerability and computer crime.

Q:3 Explain how you will protect information system.

Q:4 What do you understand by Disaster recovery plan. What is its importance?

ANSWERS TO SELF ASSESSMENT QUESTIONS

A. 4
B. 3
C. 2
D. 2
E. 4
F. 3

UNIT–IV

CHAPTER 8

Computer Communication Networks

8.0 Objectives

After studying this chapter you will be able to:

4. **Understand the Computer communication network**
5. **Understand various types of computer networks**
6. **Understand about the LAN, WAN, MAN**
7. **Understand about the communication media**

8.1Computer Communication Networks

A **computer network** or **data network** is a telecommunications network that allows computers to exchange data. In computer networks, networked computing devices pass data to each other along data connections. Data is transferred in the form of packets. The connections (network links) between nodes are established using either cable media or wireless media. The best-known computer network is the Internet.

Network computer devices that originate, route and terminate the data are called network nodes. Nodes can include hosts such as personal computers, phones, servers as well as networking hardware. Two such devices are said to be networked together when one device is able to exchange information with the other device, whether or not they have a direct connection to each other.

Computer networks differ in the physical media used to transmit their signals, the communications protocols to organize network traffic, the network's size, topology and organizational intent. In most cases, communications protocols are layered on (i.e. work using) other more specific or more general communications protocols, except for the physical layer that directly deals with the physical media.

Computer networks support applications such as access to the World Wide Web, shared use of application and storage servers, printers, and fax machines, and use of email and instant messaging applications

A computer network, or simply a network, is a collection of computers and other hardware components interconnected by communication channels that allow sharing of resources and information. Today, computer networks are the core of modern communication. All modern aspects of the public switched telephone network (PSTN) are computer-controlled. Telephony increasingly runs over the Internet Protocol, although not necessarily the public Internet. The scope of communication has increased significantly in the past decade. This boom in communications would not have been possible without the progressively advancing computer network. Computer networks, and the technologies that make communication between networked computers possible, continue to drive computer hardware, software, and peripherals industries. The expansion of related industries is mirrored by growth in the numbers and types of people using networks, from the researcher to the home user.

The following is a chronology of significant computer network developments:

- In the late 1950s, early networks of communicating computers included the military radar system Semi-Automatic Ground Environment (SAGE).
- In 1960, the commercial airline reservation system semi-automatic business research environment (SABRE) went online with two connected mainframes.
- In 1962, J.C.R. Licklider developed a working group he called the "Intergalactic Computer Network", a precursor to the ARPANET, at the Advanced Research Projects Agency (ARPA).
- In 1964, researchers at Dartmouth developed the Dartmouth Time Sharing System for distributed users of large computer systems. The same year, at Massachusetts Institute of Technology, a research group supported by General Electric and Bell Labs used a computer to route and manage telephone connections.
- Throughout the 1960s, Leonard Kleinrock, Paul Baran, and Donald Davies independently developed network systems that used packets to transfer information between computers over a network.
- In 1965, Thomas Marill and Lawrence G. Roberts created the first wide area network (WAN). This was an immediate precursor to the ARPANET, of which Roberts became program manager.

- Also in 1965, the first widely used telephone switch that implemented true computer control was introduced by Western Electric.
- In 1969, the University of California at Los Angeles, the Stanford Research Institute, the University of California at Santa Barbara, and the University of Utah were connected as the beginning of the ARPANET network using 50 kbit/s circuits.
- In 1972, commercial services using X.25 were deployed, and later used as an underlying infrastructure for expanding TCP/IP networks.
- In 1973, Robert Metcalfe wrote a formal memo at Xerox PARC describing Ethernet, a networking system that was based on the Aloha network, developed in the 1960s by Norman Abramson and colleagues at the University of Hawaii. In July 1976, Robert Metcalfe and David Boggs published their paper "Ethernet: Distributed Packet Switching for Local Computer Networks" and collaborated on several patents received in 1977 and 1978. In 1979, Robert Metcalfe pursued making Ethernet an open standard.
- In 1976, John Murphy of Data point Corporation created ARCNET, a token-passing network first used to share storage devices.
- In 1995, the transmission speed capacity for Ethernet was increased from 10 Mbit/s to 100 Mbit/s. By 1998, Ethernet supported transmission speeds of a Gigabit. The ability of Ethernet to scale easily (such as quickly adapting to support new fiber optic cable speeds) is a contributing factor to its continued use today.

Properties

Computer networking may be considered a branch of electrical engineering, telecommunications, computer science, information technology or computer engineering, since it relies upon the theoretical and practical application of the related disciplines.

A computer network facilitates interpersonal communications allowing people to communicate efficiently and easily via email, instant messaging, chat rooms, telephone, video telephone calls, and video conferencing. Providing access to information on shared storage devices is an important feature of many networks. A network allows sharing of files, data, and other types of information giving authorized users the ability to access information stored on other computers on the network. A network allows sharing of network and computing resources. Users may access and use resources provided by devices on the network, such as printing a document on a shared network printer. Distributed computing uses computing resources across networks, to accomplish various task. A computer network may be used by computer Crackers to deploy computer viruses or computer worms on devices connected to the network, or to prevent these devices from

accessing the network (denial of service). A complex computer network may be difficult to set up. It may be costly to set up an effective computer network in a large organization.

A network can be characterized by its physical capacity or its organizational purpose. Use of the network, including user authorization and access rights, differ accordingly.

ACTIVITIES

1. Describe in your own words a database and database management system. Discuss the objectives and advantages of a database.
2. Explain the drawbacks of the file processing system. What is the alternative to the files approach?

Nanoscale Network

A nanoscale communication network has key components implemented at the nanoscale including message carriers and leverages physical principles that differ from macroscale communication mechanisms. Nanoscale communication extends communication to very small sensors and actuators such as those found in biological systems and also tends to operate in environments that would be too harsh for classical communication.

Personal area network

A personal area network (PAN) is a computer network used for communication among computer and different information technological devices close to one person. Some examples of devices that are used in a PAN are personal computers, printers, fax machines, telephones, PDAs, scanners, and even video game consoles. A PAN may include wired and wireless devices. The reach of a PAN typically extends to 10 meters. A wired PAN is usually constructed with USB and FireWire connections while technologies such as Bluetooth and infrared communication typically form a wireless PAN.

Local area network

A local area network (LAN) is a network that connects computers and devices in a limited geographical area such as a home, school, office building, or closely positioned group of buildings. Each computer or device on the network is a node. Wired LANs are most likely based on Ethernet technology. Newer standards such as ITU-T G.hn also provide a way to create a wired LAN using existing wiring, such as coaxial cables, telephone lines, and power lines.

All interconnected devices use the network layer (layer 3) to handle multiple subnets (represented by different colors). Those inside the library have 10/100 Mbit/s Ethernet connections to the user device and a Gigabit Ethernet connection to the central router. They could be called *Layer 3 switches*, because they only have Ethernet interfaces and support the Internet Protocol. It might be more correct to call them access routers, where the router at the top is a distribution router that connects to the Internet and to the academic networks' customer access routers.

The defining characteristics of a LAN, in contrast to a wide area network (WAN), include higher data transfer rates, limited geographic range, and lack of reliance on leased lines to provide connectivity. Current Ethernet or other IEEE 802.3 LAN technologies operate at data transfer rates up to 10 Gbit/s. The IEEE investigates the standardization of 40 and 100 Gbit/s rates. A LAN can be connected to a WAN using a router.

Home area network

A home area network (HAN) is a residential LAN used for communication between digital devices typically deployed in the home, usually a small number of personal computers and accessories, such as printers and mobile computing devices. An important function is the sharing of Internet access, often a broadband service through a cable TV or digital subscriber line (DSL) provider.

Storage area network

A storage area network (SAN) is a dedicated network that provides access to consolidated, block level data storage. SANs are primarily used to make storage devices, such as disk arrays, tape libraries, and optical jukeboxes, accessible to servers so that the devices appear like locally attached devices to the operating system. A SAN typically has its own network of storage devices that are generally not accessible through the local area network by other devices. The cost and complexity of SANs dropped in the early 2000s to levels allowing wider adoption across both enterprise and small to medium sized business environments.

Campus area network

A campus area network (CAN) is made up of an interconnection of LANs within a limited geographical area. The networking equipment (switches, routers) and transmission media (optical fiber, copper plant, Cat5 cabling, etc.) are almost entirely owned by the campus tenant / owner (an enterprise, university, government, etc.).

For example, a university campus network is likely to link a variety of campus buildings to connect academic colleges or departments, the library, and student residence halls.

Backbone network

A backbone network is part of a computer network infrastructure that provides a path for the exchange of information between different LANs or sub-networks. A backbone can tie together diverse networks within the same building, across different buildings, or over a wide area.

For example, a large company might implement a backbone network to connect departments that are located around the world. The equipment that ties together the departmental networks constitutes the network backbone. When designing a network backbone, network performance and network congestion are critical factors to take into account. Normally, the backbone network's capacity is greater than that of the individual networks connected to it.

Another example of a backbone network is the Internet backbone, which is the set of wide area networks (WANs) and core routers that tie together all networks connected to the Internet.

Metropolitan area network

A Metropolitan area network (MAN) is a large computer network that usually spans a city or a large campus.

Wide area network

A wide area network (WAN) is a computer network that covers a large geographic area such as a city, country, or spans even intercontinental distances. A WAN uses a communications channel that combines many types of media such as telephone lines, cables, and air waves. A WAN often makes use of transmission facilities provided by common carriers, such as telephone companies. WAN technologies generally function at the lower three layers of the OSI reference model: the physical layer, the data link layer, and the network layer.

Enterprise private network

An enterprise private network is a network that a single organization builds to interconnect its office locations (e.g., production sites, head offices, remote offices, shops) so they can share computer resources.

Virtual private network

A virtual private network (VPN) is an overlay network in which some of the links between nodes are carried by open connections or virtual circuits in some larger network (e.g., the Internet) instead of by physical wires. The data link layer protocols of the virtual network are said to be tunneled through the larger network when this is the case. One common application is secure communications through the public Internet, but a

VPN need not have explicit security features, such as authentication or content encryption. VPNs, for example, can be used to separate the traffic of different user communities over an underlying network with strong security features.

VPN may have best-effort performance, or may have a defined service level agreement (SLA) between the VPN customer and the VPN service provider. Generally, a VPN has a topology more complex than point-to-point.

Global area network

A global area network (GAN) is a network used for supporting mobile across an arbitrary number of wireless LANs, satellite coverage areas, etc. The key challenge in mobile communications is handing off user communications from one local coverage area to the next. In IEEE Project 802, this involves a succession of terrestrial wireless LANs.

Organizational scope

Networks are typically managed by the organizations that own them. Private enterprise networks may use a combination of intranets and extranets. They may also provide network access to the Internet, which has no single owner and permits virtually unlimited global connectivity.

Intranets

An intranet is a set of networks that are under the control of a single administrative entity. The intranet uses the IP protocol and IP-based tools such as web browsers and file transfer applications. The administrative entity limits use of the intranet to its authorized users. Most commonly, an intranet is the internal LAN of an organization. A large intranet typically has at least one web server to provide users with organizational information. An intranet is also anything behind the router on a local area network.

Extranet

An extranet is a network that is also under the administrative control of a single organization, but supports a limited connection to a specific external network. For example, an organization may provide access to some aspects of its intranet to share data with its business partners or customers. These other entities are not necessarily trusted from a security standpoint. Network connection to an extranet is often, but not always, implemented via WAN technology.

Internetwork

An internetwork is the connection of multiple computer networks via a common routing technology using routers.

Internet

The Internet is the largest example of an internetwork. It is a global system of interconnected governmental, academic, corporate, public, and private computer networks. It is based on the networking technologies of the Internet Protocol Suite. It is the successor of the Advanced Research Projects Agency Network (ARPANET) developed by DARPA of the United States Department of Defense. The Internet is also the communications backbone underlying the World Wide Web (WWW).

Participants in the Internet use a diverse array of methods of several hundred documented, and often standardized, protocols compatible with the Internet Protocol Suite and an addressing system (IP addresses) administered by the Internet Assigned Numbers Authority and address registries. Service providers and large enterprises exchange information about the reachability of their address spaces through the Border Gateway Protocol (BGP), forming a redundant worldwide mesh of transmission paths.

Darknet

A Darknet is an overlay network, typically running on the internet that is only accessible through specialized software. A darknet is an anonymizing network where connections are made only between trusted peers — sometimes called "friends" (F2F) — using non-standard protocols and ports.

Darknets are distinct from other distributed peer-to-peer networks as sharing is anonymous (that is, IP addresses are not publicly shared), and therefore users can communicate with little fear of governmental or corporate interference.

Components of a Network

A computer network comprises the following components:

- A minimum of at least 2 computers
- Cables that connect the computers to each other, although wireless communication is becoming more common
- A network interface device on each computer (this is called a network interface card or NIC)
- A ‘Switch’ used to switch the data from one point to another. Hubs are outdated and are little used for new installations.

- Network operating system software

SELF ASSESSMENT QUESTIONS

Please select the correct option

A. Which of the following is part of a static view of information?
   1. Logical data model
   2. Meta data
   3. Data flow model
   4. Information process model

B. Contemporary Information Systems are interfacing with customers and suppliers using :
   1. BPR
   2. CRM
   3. SCM
   4. Both A and B
   5. Both B and C

C. Information systems that support the business functions that reach out to suppliers are known as:
   1. back office information systems
   2. decision support systems
   3. expert information systems
   4. front office information systems
   5. none of the above

D. Which of the following is not a class of information system applications?
   1. database management system
   2. decision support system
   3. expert system
   4. management information system
   5. office automation system

E. Who are the people that actually use the system to perform or support the work to be completed?
   1. system analysts
   2. system designers
   3. system owners

4. system builders
5. none of the above

F. Which is not a typical business function?
1. Sales
2. Service
3. Manufacturing
4. Accounting
5. Benefits and Compensation

8.2 Telecommunication and computer networks

A computer network consists of a collection of computers, printers and other equipment that is connected together so that they can communicate with each other. Figure below gives an example of a network in a school comprising of a local area network or LAN connecting computers with each other, the internet, and various servers.

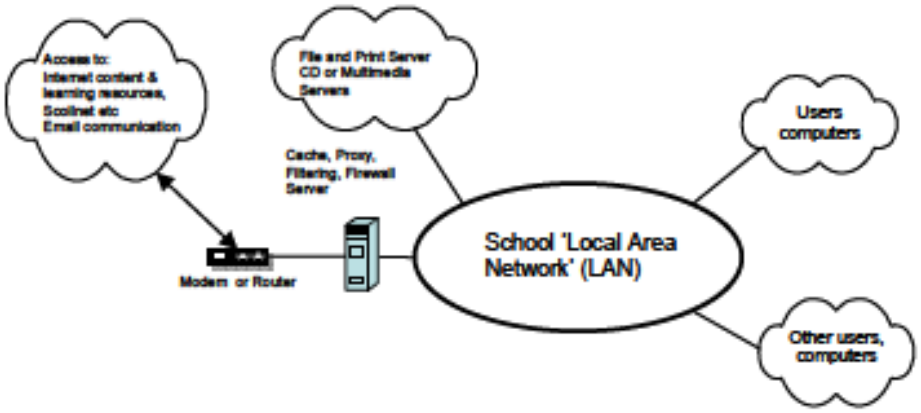


Figure: Representation of Network in a school

Broadly speaking, there are two types of network configuration, peer-to-peer networks and client/server networks.

Peer-to-peer networks are more commonly implemented where less then ten computers are involved and where strict security is not necessary. All computers have the same status, hence the term 'peer', and they communicate with each other on an equal footing. Files, such as word processing or spreadsheet documents, can be shared across the network and all the computers on the network can share devices, such as printers or scanners, which are connected to any one computer.

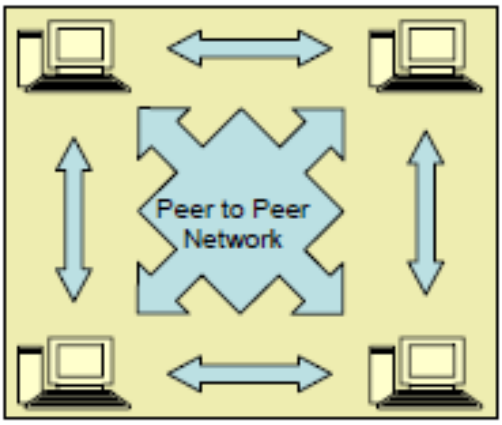


Figure: Peer to Peer Network

Client/server networks are more suitable for larger networks. A central computer, or 'server', acts as the storage location for files and applications shared on the network. Usually the server is a higher than average performance computer. The server also controls the network access of the other computers which are referred to as the 'client' computers. Typically, teachers and students in a school will use the client computers for their work and only the network administrator (usually a designated staff member) will have access rights to the server.

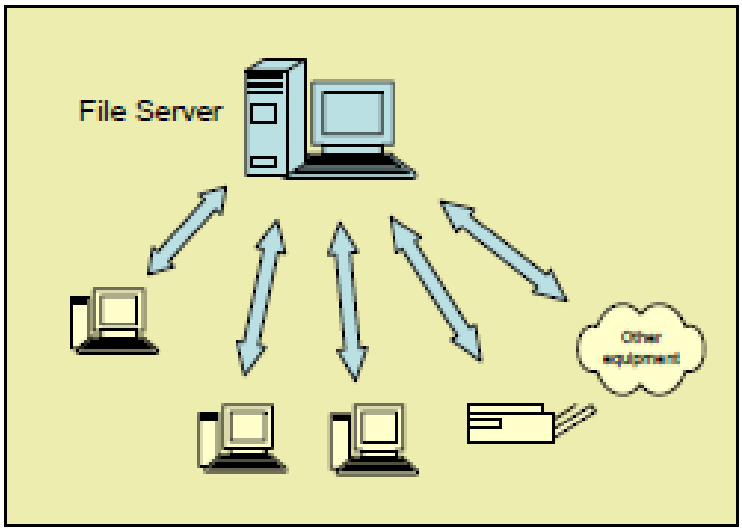


Figure: Client/server network

Table: provides a summary comparison between Peer-to-Peer and Client/Server Networks.

Peer-to-Peer Networks vs Client/Server Networks	
Peer-to-Peer Networks	**Client/Server Networks**
Easy to set up	More difficult to set up
Less expensive to install	More expensive to install
Can be implemented on a wide range of operating	A variety of operating systems can be supported on

systems	the client computers, but the server needs to run an operating system that supports networking
More time consuming to maintain the software being used (as computers must be managed individually)	Less time consuming to maintain the software being used (as most of the maintenance is managed from the server)
Very low levels of security supported or none at all. These can be very cumbersome to set up, depending on the operating system being used	High levels of security are supported, all of which are controlled from the server. Such measures prevent the deletion of essential system files or the changing of settings
Ideal for networks with less than 10 computers	No limit to the number of computers that can be supported by the network
Does not require a server	Requires a server running a server operating system
Demands a moderate level of skill to administer the network	Demands that the network administrator has a high level of IT skills with a good working knowledge of a server operating system

8.3 Network Types

Computer Networks are classified into many categories based on their respective attributes. These includes:

- Geographical span
- Inter-connectivity
- Administration
- Architecture

Geographical Span

Geographically a network can be seen in one of the following categories:

- It may be spanned across your table, among Bluetooth enabled devices. Ranging not more than few meters.
- It may be spanned across a whole building, including intermediate devices to connect all floors.
- It may be spanned across a whole city.
- It may be spanned across multiple cities or provinces.
- It may be one network covering whole world.

Inter-connectivity

Components of a network can be connected to each other differently in some fashion. By connectedness we mean either logically or physically or both ways.

- Every single device can be connected to every other device on network, making the network mesh.
- All devices can be connected to a single medium but geographically disconnected, created bus like structure.
- Each device is connected to its left and right peers only, creating linear structure.
- All devices connected together with a single device, creating star like structure.
- All devices connected arbitrarily using all previous ways to connect each other, resulting in a hybrid structure.

Administration

From an administrator's point of view, a network can be private network which belongs a single autonomous system and cannot access outside its physical or logical domain. Or a network can be a public network, which can be accessed by all.

Network Architecture

- There can be one or more systems acting as Server. Other being Client, request the Server to serve requests. Servers take and process request on behalf of Clients.
- Two systems can be connected Point-to-Point, or in other words back-to-back fashion. They both reside on same level and called peers.
- There can be hybrid network which involves network architecture of both the above types.

Network Applications

Computer systems and peripherals are connected to form a network provides bunch of advantages:

- Resource sharing such as printers and storage devices.
- Exchange of Information by means of eMails and FTP.
- Information sharing by using Web or Internet.
- Interaction with other users using dynamic web pages.
- IP phones
- Video Conferences
- Parallel computing
- Instant Messaging

8.4 PAN,LAN, WAN, MAN

Personal Area Network

A Personal Area Network or simply PAN, is smallest network which is very personal to a user. This may include Bluetooth enabled devices or infra-red enabled devices. PAN has connectivity range up to 10 meters. PAN may include wireless computer keyboard and mouse, Bluetooth enabled headphones, wireless printers and TV remotes for example.

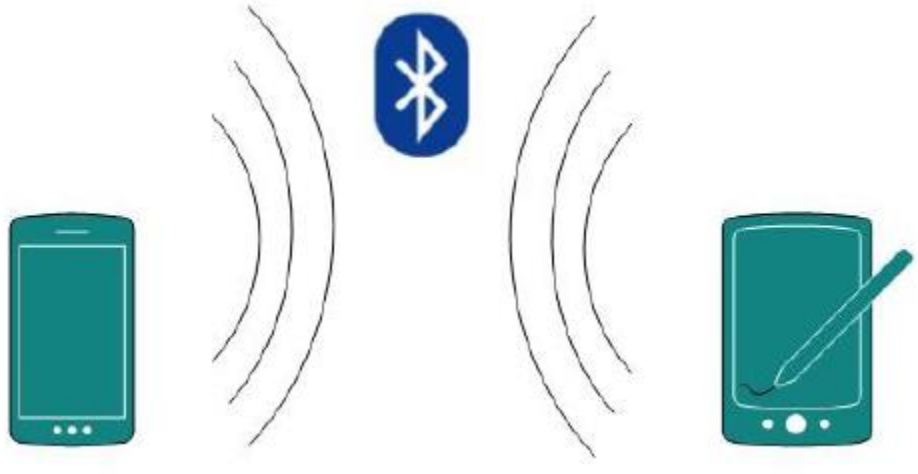


Figure: Personal Are Network

Piconet is an example Bluetooth enabled Personal Area Network which may contain up to 8 devices connected together in a master-slave fashion

A **wireless personal area network** (**WPAN**) is a PAN carried over wireless network technologies such as:

- INSTEON
- IrDA
- Wireless USB
- Bluetooth
- Z-Wave
- ZigBee
- Body Area Network

The reach of a WPAN varies from a few centimeters to a few meters. A PAN may also be carried over wired computer buses such as USB and FireWire

A wireless personal area network (WPAN) is a personal area network — a network for interconnecting devices centered on an individual person's workspace — in which the connections are wireless. Wireless PAN is based on the standard IEEE 802.15. The two kinds of wireless technologies used for WPAN are Bluetooth and Infrared Data Association.

A WPAN could serve to interconnect all the ordinary computing and communicating devices that many people have on their desk or carry with them today; or it could serve a more specialized purpose such as allowing the surgeon and other team members to communicate during an operation.

A key concept in WPAN technology is known as "plugging in". In the ideal scenario, when any two WPAN-equipped devices come into close proximity (within several meters of each other) or within a few kilometers of a central server, they can communicate as if connected by a cable. Another important feature is the ability of each device to lock out other devices selectively, preventing needless interference or unauthorized access to information.

The technology for WPANs is in its infancy and is undergoing rapid development. Proposed operating frequencies are around 2.4 GHz in digital modes. The objective is to facilitate seamless operation among home or business devices and systems. Every device in a WPAN will be able to plug into any other device in the same WPAN, provided they are within physical range of one another. In addition, WPANs worldwide will be interconnected. Thus, for example, an archeologist on site in Greece might use a PDA to directly access databases at the University of Minnesota in Minneapolis, and to transmit findings to that database.

Bluetooth

Bluetooth uses short-range radio waves over distances up to approximately 10 metres. For example, Bluetooth devices such as a keyboard, pointing devices, audio head sets, printers may connect to personal digital assistants (PDAs), cell phones, or computers wirelessly.

A Bluetooth PAN is also called a *piconet* (combination of the prefix "pico," meaning very small or one trillionth, and network), and is composed of up to 8 active devices in a master-slave relationship (a very large number of devices can be connected in "parked" mode). The first Bluetooth device in the piconet is the master, and all other devices are slaves that communicate with the master. A piconet typically has a range of 10 metres (33 ft), although ranges of up to 100 metres (330 ft) can be reached under ideal circumstances.

Infrared Data Association

Infrared Data Association (IrDA) uses infrared light, which has a frequency below the human eye's sensitivity. Infrared in general is used, for instance, in TV remotes. Typical WPAN devices that use IrDA include printers, keyboards, and other serial data interfaces.

Wi-Fi

Wi-Fi uses radio waves for connection over distances up to around 91 meters, usually in a local area network (LAN) environment. Wi-Fi can be used to connect local area networks, to connect cell phones to the Internet to download music and other multimedia, to allow PC multimedia content to be stream to the TV (Wireless Multimedia Adapter), and to connect video game consoles to their networks (Nintendo Wi-Fi Connection).

Body area network

A body area network is based on the IEEE 802.15.6 standard for transmission via the capacitive near field of human skin allowing near field communication of devices worn by and near the wearer.[3] The Skinplex implementation can detect and communicate up to 1 metre (3 ft 3 in) from a human body.[4] It is used for access control to door locks and jamming protection in convertible car roofs. Projects that implement Body Area Network include the work of RedTacton RT/Aswini.

Local Area Network

A computer network spanned inside a building and operated under single administrative system is generally termed as Local Area Network. Usually, Local Area Network covers an organization's offices, schools, college/universities etc. Number of systems may vary from as least as two to as much as 16 million LAN provides a useful way of sharing resources between end users. Resources like Printers, File Servers, Scanners and internet is easy sharable among computers.

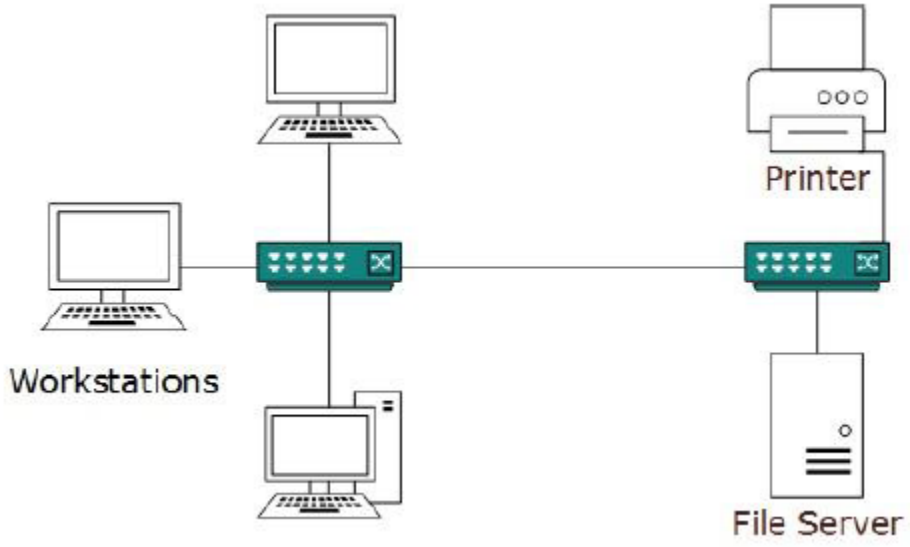


Figure: Local Are Network

Local Area Networks are composed of inexpensive networking and routing equipment. It may contains local servers serving file storage and other locally shared applications. It mostly operates on private IP addresses and generally do not involve heavy routing. LAN works under its own local domain and controlled centrally.

LAN uses either Ethernet or Token-ring technology. Ethernet is most widely employed LAN technology and uses Star topology while Token-ring is rarely seen. LAN can be wired or wireless or in both forms at once.

Cabling

Early LAN cabling had generally been based on various grades of coaxial cable. Shielded twisted pair was used in IBM's Token Ring LAN implementation, but in 1984, Star LAN showed the potential of simple *unshielded* twisted pair by using Cat3 cable—the same simple cable used for telephone systems. This led to the development of 10Base-T (and its successors) and structured cabling which is still the basis of most commercial LANs today.

Fiber-optic cabling is common for links between switches, but fiber to the desktop is uncommon.

Wireless

As well as traditional cabling, many LANs are now based partly or wholly on wireless technologies. Almost all of today's smart phones, tablets and laptops have wireless support built-in so a wireless local area network, or WLAN, gives users the ability to move around within a local coverage area and still be connected to the network. Wireless networks have become popular in domestic homes due to ease of installation, and in commercial complexes to offer easy network access to their staff. Visiting guest are often offered internet access via a hotspot service.

Technical Aspects

Network topology describes the layout of interconnections between devices and network segments. At the Data Link Layer and Physical Layer, a wide variety of LAN topologies have been used, including ring, bus, mesh and star, but the most common LAN topology in use today is switched Ethernet. At the higher layers, the Internet Protocol (TCP/IP) has become the standard, replacing NetBEUI, IPX/SPX, AppleTalk and others.

Simple LANs generally consist of one or more switches. A switch can be connected to a router, cable modem, or ADSL modem for Internet access. Complex LANs are characterized by their use of redundant links with switches using the spanning tree protocol to prevent loops, their ability to manage differing traffic types via quality of service(QoS), and to segregate traffic with VLANs. A LAN can include a wide variety of network devices such as switches, firewalls, routers, load balancers, and sensors.

LANs can maintain connections with other LANs via leased lines, leased services, or the Internet using virtual private network technologies. Depending on how the connections are established and secured in a LAN, and the distance involved, a LAN may also be classified as a metropolitan area network (MAN) or a wide area network (WAN).

Advantages of LAN

- Files can be stored on a central computer(the file server) allowing data to be shared throughout an organization
- Files can be backed up more easily when they are all on a central fileserver rather than when they are scattered across a number of independent workstations.
- Networking also allow security to be established ensuring that the network users may only have access to certain files and applications.
- Software and resources can be centrally managed
- Network versions of software often allow for their speedy installations on workstations from the file server
- Expensive devices such as laser printers or scanners can be shared
- Users can access their files from any workstation

Disadvantages of LAN

- Power - a good LAN is required to be on all the times
- Security- each computer and device become another point of entry for undesirables.
- Unkeep – when things g wrong or the software gets updated
- A lot of times a network shares one Internet connection – if all computers running at once can reduce speed for each
- Area covered is limited

Metropolitan Area Network

A **metropolitan area network** (**MAN**) is computer network larger than a local area network, covering an area of a few city blocks to the area of an entire city, possibly also including the surrounding areas.

> " A MAN is optimized for a larger geographical area than a LAN, ranging from several blocks of buildings to entire cities. MANs can also depend on communications channels of moderate-to-high data rates. A MAN might be owned and operated by a single organization, but it usually will be used by many individuals and organizations. MANs might also be owned and operated as public utilities. They will often provide means for inter networking of local networks. "

Kenneth C. Laudan and Jane P. Laudan define a metropolitan area network as:

> " A Metropolitan Area Network (MAN) is a large computer network that spans a metropolitan area or campus. Its geographic scope falls between a WAN and LAN. MANs provide Internet connectivity for LANs in a metropolitan region, and connect them to wider area networks like the Internet. "

Advantages

1. It is used for city networking.

2. More than 10 computers can be connected in this networking.

MAN, generally expands throughout a city such as cable TV network. It can be in form of Ethernet, Token-ring, ATM or FDDI. Metro Ethernet is a service which is provided by ISPs. This service enables its users to expand their Local Area Networks. For example, MAN can help an organization to connect all of its offices in a City.

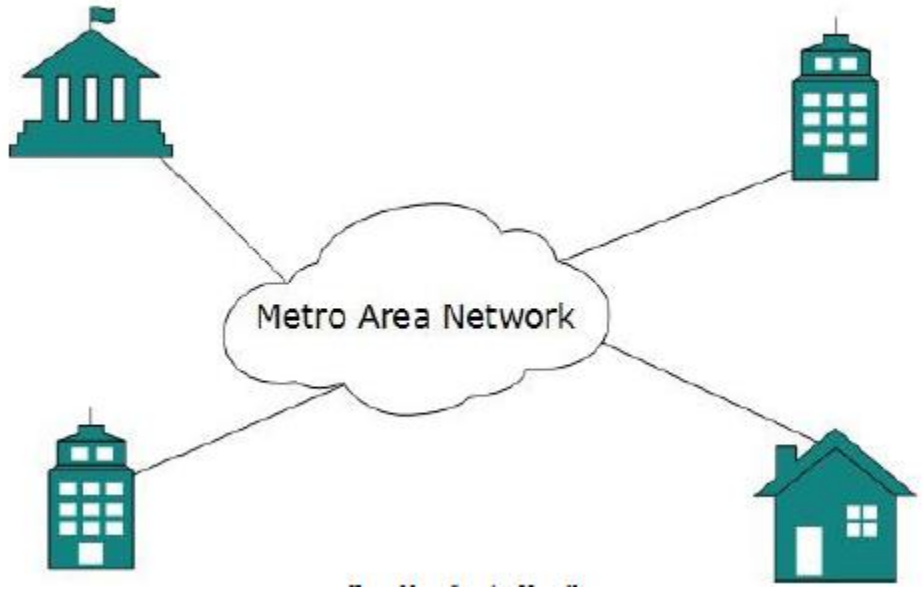


Figure: MAN

Backbone of MAN is high-capacity and high-speed fiber optics. MAN is works in between Local Area Network and Wide Area Network. MAN provides uplink for LANs to WANs or Internet.

Implementation

Also known as a Municipal Area Network, networking technologies used in municipal networks include Asynchronous Transfer Mode(ATM), FDDI, and SMDS. However, these technologies are increasingly being displaced by Ethernet-based connections (e.g., Metro Ethernet). MAN links between local area networks have been built with wireless links using either microwave, radio, or infra-red laser transmission. Most companies rent or lease circuits from common carriers because laying long stretches of cable is expensive.

Distributed-queue dual-bus (DQDB) refers to the metropolitan area network standard for data communication specified in the IEEE 802.6 standard. With DQDB, networks can extend up to 20 miles (30 km) long and operate at speeds of 34–155 Mbit/s.

MAN Advantages

MAN can cover a wider area than a LAN. MAN networks are usually operated at airports, or a combination of several pieces at a local school. By running a large network connectedness, information can be disseminated more widely, rapidly and significantly. Public libraries and government agencies typically use a MAN

MAN Disadvantages

MAN will only apply if the personal computer or a terminal can compete. If a personal computer is used as a terminal, move the file (file transfer software) allows users to retrieve files (downloaded) from the hose or hose to deliver the data (upload). Download files means open and retrieve data from a personal computer to another and deliver the data to the computer pertaining requested by the user

Wide Area Network

As name suggests, this network covers a wide area which may span across provinces and even a whole country. Generally, telecommunication networks are Wide Area Network. These networks provides connectivity to MANs and LANs. Equipped with very high speed backbone, WAN uses very expensive network equipment

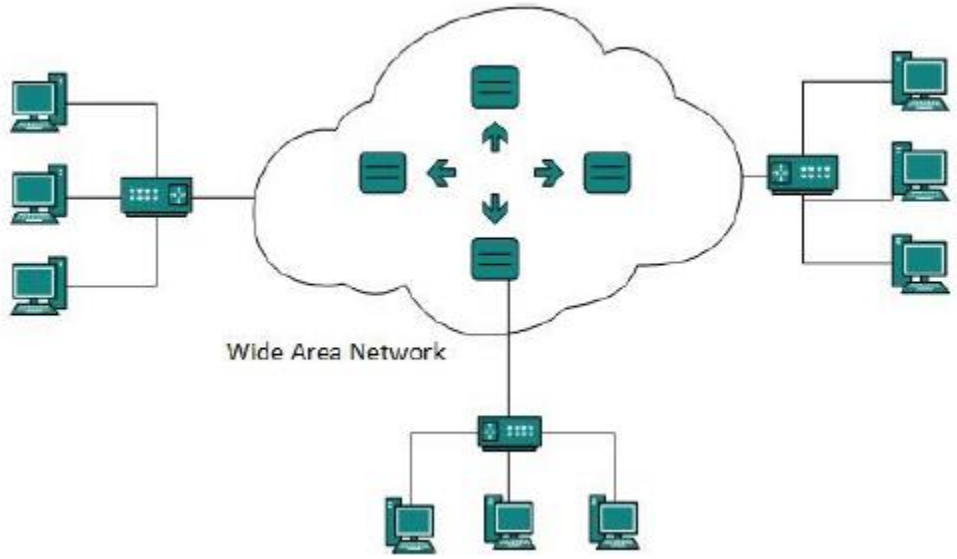


Figure: Wide Area Network

WAN may use advanced technologies like Asynchronous Transfer Mode (ATM), Frame Relay and SONET. WAN may be managed under by more than one administration.

8.5Communication Media

Communication media refers to the means of delivering and receiving data or information. In telecommunication, these means are transmission and storage tools or channels for data storage and transmission. The term is also commonly used in place of mass media or news media. Different media are employed for transmitting data from one computer terminal to the central computer or to other computer systems inside some kind of network. The communication media acts as a communication channel for linking various computing devices so that they may interact with each other

There are two forms of communication media:

- Analog: Includes the conventional radio, telephonic and television broadcasts
- Digital: Computer-mediated communication, computer networking and telegraphy

The most commonly used data communication media include:

- Wire pairs
- Coaxial cable
- Microwave transmission
- Communication satellites
- Fiber optics

Two Wire Open Line: this is the simplest of all the transmission media. It consists of simple pair of metallic wires made of copper or aluminum of between 0.4 and 1mm diameter. This is used for short distance upto 50 m and can transfer up to 19,200 bps.

Twisted pair cable: A twisted pair consists of insulated conductors that are twisted together. It is used for communication up to distance of 1 km and can achieve transfer rate 1-2 mbps. Twisted pair cable is widely used in telephone network.

Coaxial Cable: A coaxial cable consists of solid conductor running coaxial cable can be used over a distance about 1 KM and can achieve a transfer rate of up to 100 mbps. A coaxial cable is of two types: a 75 ohm cable which is used by cable TV operator and the 50 ohm cable which is used in high speed broadband.

Fiber optics cable: A fiber optics cable carries signals in the form of fluctuating light in a glass or plastic fiber. It has very high data transfer rates of about 1000mbps. These are known as guided media.

Radiowave, microwave and satellite: Radiowave. Microwave, satellite channeks use electromagnetic propagation in open space. It covers large geographical area. These are known as un-guided media.

Network: A network is a way or means of transmitting or receiving of information from one or more sources.

8.6 Communication Hardware

The transfer of data over a communications circuit requires appropriate hardware and software. Once these are working you can use various applications for the following services and operations:-

- Internet

Having registered with an Internet service provider (ISP), you can use the World Wide Web (WWW), send messages via electronic mail (e-mail), convey files using file transfer protocol (FTP) or use Internet telephony to speak to other Internet users for the cost of a local call, irrespective of distance. However, to talk to others that aren't on the Internet you must register with an internet telephone service provider (ITSP).

- Peer-to-Peer (Point-to-Point) Communications

This lets you transfer data between computers using a communications application, but without using the Internet. Unfortunately, the call charges are expensive over long distances.

- Terminal Emulation

This is similar to peer-to-peer operation, but lets you communicate with an old-fashioned mainframe computer or a more modern system that uses a compatible terminal.

Hardware Options

A wide range of hardware is available for interconnecting computers, most of which is designed for the Internet. As well as the hardware itself, you'll need to deal with the cost of *installation* and any associated *cables*. And for Internet access you'll need to *subscribe* to a *service provider*, although some companies can provide all that you need in a single package.

The choice of hardware is a compromise between cost and *speed*. The latter, measured in *kilobits per second (kbit/s)*, must be considered in two directions. First, there's the *upstream speed*, the rate at which you can send data or *upload* files to a remote computer. Secondly, there's the *downstream speed*, the rate a which you can receive data or *download* files from another machine. Of the two, the latter is more important, since most people receive more data than they send.

The following table shows the various options available in order of mass popularity. The speeds shown here are typical, although higher rates are possible.

Connection Method	Access	Downstream (kbit/s)	Upstream (kbit/s)
Standard modem	Dial-up	56	56
ADSL	Permanent	512	256
SDSL	Permanent	512	512
Cable	Permanent	128/512	256
ISDN	Dial-up	64/128	64/128
Leased line	Permanent	64/128	64/128
Satellite	Permanent	400	56
Wireless	Permanent	128/512	256/512

A *permanent connection* is essential if you want to operate your own Internet *server*.

Standard Modem (56 kbit/s)

This lets you convey digital data over a normal phone circuit, also known as the public switched telephone network (PSTN) or plain old telephone service (POTS). Both the caller and the recipient must have a modem connected to their computers. As with a normal telephone call, the link is operated as a dial-up service, only providing a connection when required.

The modem at the sending end modulates the data into an audio signal which the modem at the receiving end demodulates back into data. Speed is seriously limited, since the phone system is designed for speech signals of restricted bandwidth.

Some modems support multilink operation, sharing data over several phone circuits and giving increased speed.

ADSL (512 kbit/s)

An Asymmetric Digital Subscriber Line requires special wiring from your telephone company, as well as an ADSL terminal adaptor (ADSL TA) or an ADSL modem. ADSL is faster and more reliable than a dial-up circuit. When used with a router, it gives a permanent connection to the Internet, requiring you to use personal firewall software for protection against hackers.

SDSL (512 kbit/s)

A Symmetric Digital Subscriber Line, also known as G.SHDSL, is similar to ADSL, but offers the same speed in both directions, as required by some professional users. This makes it an ideal alternative to a leased line or ISDN connection.

Cable (128/512 kbit/s)

Cable-assisted television (CATV), also known as cable TV, employs coaxial cables to carry TV pictures, but can also provide a permanent connection to the Internet. This requires a CATV splitter box for feeding your TV and a special cable modem.

ISDN (64/128 kbit/s)

The Integrated Services Digital Network (ISDN) predates ADSL, but is well-established in professional circles. It uses one or more circuits from your local telephone exchange, normally operated as a dial-up service. However, instead of a modem, you'll need an ISDN terminal adaptor (ISDN TA), ISDN card or USB-to-ISDN adaptor. Such devices support peer-to-peer communication while some TAs also accommodate multilink operation.

Leased Line (64/128 kbit/s)

This option, sometimes known as a T1 connection, is similar to ISDN but provides a permanent connection, which means you'll need firewall software to protect yourself against hackers.

Mobile Phone (9.6 kbit/s - 2 Mbit/s)

You can use your mobile phone for sending and receiving data, preferably in conjunction with a portable computer. Although GSM and other current systems are very slow, the second generation of phones operates at a similar rate as PSTN, and the third generation (3G) is even faster.

Satellite (400 kbit/s)

Satellite communication offers rates of 400 kbit/s or higher, although early systems used a telephone modem for sending data 'upstream'. More recent offerings download at 512 kbit/s and upload at 128 kbit/s, both via satellite, with an option for downloading at 2 MB/s. Most parts of the UK can use a 890 mm satellite dish, although a 980 mm version is needed in Northern Ireland and Scotland.

•The time-lag introduced by satellite systems makes them unsuitable for real-time games.

Wireless (128/512 kbit/s)

This technology is still emerging, although it will undoubtedly become popular.

8.7 Summary

A **computer network** or **data network** is a telecommunications network that allows computers to exchange data. In computer networks, networked computing devices pass data to each other along data connections. Data is transferred in the form of packets. The connections (network links) between nodes are established using either cable media or wireless media. The best-known computer network is the Internet.

A nanoscale communication network has key components implemented at the nanoscale including message carriers and leverages physical principles that differ from macroscale communication mechanisms. Nanoscale communication extends communication to very small sensors and actuators such as those found in biological systems and also tends to operate in environments that would be too harsh for classical communication.

A personal area network (PAN) is a computer network used for communication among computer and different information technological devices close to one person.

A local area network (LAN) is a network that connects computers and devices in a limited geographical area such as a home, school, office building, or closely positioned group of buildings. Each computer or device on the network is a node.

A home area network (HAN) is a residential LAN used for communication between digital devices typically deployed in the home, usually a small number of personal computers and accessories, such as printers and mobile computing devices.

A campus area network (CAN) is made up of an interconnection of LANs within a limited geographical area. The networking equipment (switches, routers) and transmission media (optical fiber, copper plant, Cat5 cabling, etc.) are almost entirely owned by the campus tenant / owner (an enterprise, university, government, etc.).

A backbone network is part of a computer network infrastructure that provides a path for the exchange of information between different LANs or sub-networks. A backbone can tie together diverse networks within the same building, across different buildings, or over a wide area.

A wide area network (WAN) is a computer network that covers a large geographic area such as a city, country, or spans even intercontinental distances. A WAN uses a communications channel that combines many types of media such as telephone lines, cables, and air waves.

An enterprise private network is a network that a single organization builds to interconnect its office locations (e.g., production sites, head offices, remote offices, shops) so they can share computer resources.

A virtual private network (VPN) is an overlay network in which some of the links between nodes are carried by open connections or virtual circuits in some larger network (e.g., the Internet) instead of by physical wires.

An extranet is a network that is also under the administrative control of a single organization, but supports a limited connection to a specific external network. An intranet is a set of networks that are under the control of a single administrative entity. The intranet uses the IP protocol and IP-based tools such as web browsers and file transfer applications.

Computer Networks are classified into many categories based on their respective attributes. These includes:

- Geographical span
- Inter-connectivity
- Administration
- Architecture

A Personal Area Network or simply PAN, is smallest network which is very personal to a user. This may include Bluetooth enabled devices or infra-red enabled devices.

A computer network spanned inside a building and operated under single administrative system is generally termed as Local Area Network. Usually, Local Area Network covers an organization’s offices, schools, college/universities etc. Number of systems may vary from as least as two to as much as 16 million LAN provides a useful way of sharing resources between end users. Resources like Printers, File Servers, Scanners and internet is easy sharable among computers.

A MAN is optimized for a larger geographical area than a LAN, ranging from several blocks of buildings to entire cities. MANs can also depend on communications channels of moderate-to-high data rates. A MAN might be owned and operated by a single organization, but it usually will be used by many individuals and organizations. MANs might also be owned and operated as public utilities. They will often provide means for inter networking of local networks.

8.8 Glossary

PAN: A personal area network (PAN) is a computer network used for communication among computer and different information technological devices close to one person.

LAN: A local area network (LAN) is a network that connects computers and devices in a limited geographical area such as a home, school, office building, or closely positioned group of buildings. Each computer or device on the network is a node.

HAN: A home area network (HAN) is a residential LAN used for communication between digital devices typically deployed in the home, usually a small number of personal computers and accessories, such as printers and mobile computing devices.

CAN: A campus area network (CAN) is made up of an interconnection of LANs within a limited geographical area. The networking equipment (switches, routers) and transmission media (optical fiber, copper plant, Cat5 cabling, etc.) are almost entirely owned by the campus tenant / owner (an enterprise, university, government, etc.).

8.9 References

1. *Computer network definition*, retrieved 2011-11-12
2. Chris Sutton. "Internet Began 35 Years Ago at UCLA with First Message Ever Sent Between Two Computers". *UCLA*. Archived from the original on March 8, 2008.
3. Ethernet: Distributed Packet Switching for Local Computer Networks, Robert M. Metcalfe and David R. Boggs, Communications of the ACM (pp 395–404, Vol. 19, No. 5), July 1976.
4. Spurgeon, Charles E. (2000). *Ethernet The Definitive Guide*. O'Reilly & Associates. ISBN 1-56592-660-9.
5. The Disadvantages of Wired Technology, Laura Acevedo, Demand Media.
6. "Bergen Linux User Group's CPIP Implementation". Blug.linux.no. Retrieved 2014-03-01.
7. A. Hooke (September 2000), *Interplanetary Internet*, Third Annual International Symposium on Advanced Radio Technologies, retrieved 2011-11-12
8. "Define switch.". WWW.Wikipedia.com. Retrieved April 8, 2008.
9. http://compnetworking.about.com/cs/internetworking/g/bldef_bridge.htm
10. D. Andersen; H. Balakrishnan; M. Kaashoek; R. Morris (October 2001), *Resilient Overlay Networks*, Association for Computing Machinery, retrieved 2011-11-12

8.10 Further Readings

11. "End System Multicast". *project web site*. Carnegie Mellon University. Retrieved May 25, 2013.
12. Wakeman, I (Jan 1992). "Layering considered harmful". *IEEE Network*: pp. 20–24.
13. Kurose, James; Ross, Kieth (2005). *Computer Networking: A Top-Down Approach*. Pearson.
14. Martin, Thomas. "Design Principles for DSL-Based Access Solutions". Retrieved 18 June 2011.
15. Nanoscale Communication Networks, Bush, S. F., ISBN 978-1-60807-003-9, Artech House, 2010. [2]
16. "personal area network (PAN)". Retrieved January 29, 2011.
17. *New global standard for fully networked home*, ITU-T, 2008-12-12, retrieved 2011-11-12
18. *IEEE P802.3ba 40Gb/s and 100Gb/s Ethernet Task Force*, retrieved 2011-11-12
19. "Mobile Broadband Wireless connections (MBWA)". Retrieved 2011-11-12.
20. Mansfield-Devine, Steve (December 2009). "Darknets". *Computer Fraud & Security* 2009 (12): 4–6. doi:10.1016/S1361-3723(09)70150-2.
21. Wood, Jessica (2010). "The Darknet: A Digital Copyright Revolution". *Richmond Journal of Law and Technology* 16 (4). Retrieved 25 October 2011.
22. RFC 1035, *Domain names - Implementation and Specification*, P. Mockapetris (November 1987)
23. Peterson LL, Davie BS. (2011). Computer Networks: A Systems Approach.
24. *Teletraffic Engineering Handbook*, ITU-T Study Group 2, archived from the original on 2007-01-11
25. Jump up ^ Telecommunications Magazine Online, Americas January 2003, Issue Highlights, Online Exclusive: Broadband Access Maximum Performance, Retrieved on February 13, 2005.
26. "State Transition Diagrams". Retrieved July 13, 2003.
27. "Definitions: Resilience". ResiliNets Research Initiative. Retrieved 2011-11-12.
28. Simmonds, A; Sandilands, P; van Ekert, L (2004). "An Ontology for Network Security Attack". *Lecture Notes in Computer Science*. Lecture Notes in Computer Science 3285: 317–323. doi:10.1007/978-3-540-30176-9_41. ISBN 978-3-540-23659-7.
29. "Is the U.S. Turning Into a Surveillance Society?". *American Civil Liberties Union*. Retrieved March 13, 2009.
30. "Bigger Monster, Weaker Chains: The Growth of an American Surveillance Society". *American Civil Liberties Union*. January 15, 2003. Retrieved March 13, 2009.
31. "Anonymous hacks UK government sites over 'draconian surveillance' ", Emil Protalinski, ZDNet, 7 April 2012, retrieved 12 March 2013
32. Hacktivists in the frontline battle for the internet retrieved 17 June 2012

8.11 **Model Questions**

Q:1. describe in detail different computer networks.

Q:2 Explain the terms LAN,MAN,PAN,WAN.

Q:3 What do you understand by Peer to peer and client server network.

Q:4 Explain in details the following: Intranets, Extranet, Internetwork and Internet

ANSWERS TO SELF ASSESSMENT QUESTIONS

A. 3

B. 5

C. 1

D. 1

E. 5

F. 5

CHAPTER 9

NETWORK TOPOLOGIES

STRUCTURE

9.0 Objectives

After studying this chapter you will be able to answer:

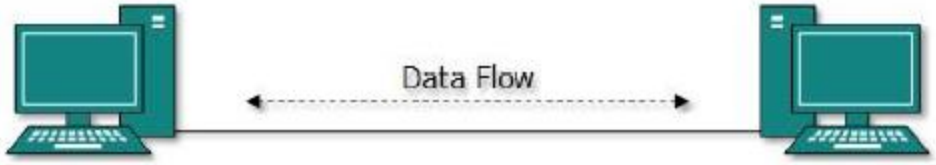
2. Understand about various network topologies.
3. Understand about Internet, Extranet and Intranet.
4. Understand about the Search Engines
5. Understand about the concept of E-Commerce

9.1 Network Topologies

A Network Topology is the way computer systems or network equipment connected to each other. Topologies may define both physical and logical aspect of the network. Both logical and physical topologies could be same or different in a same network.

Point-to-point

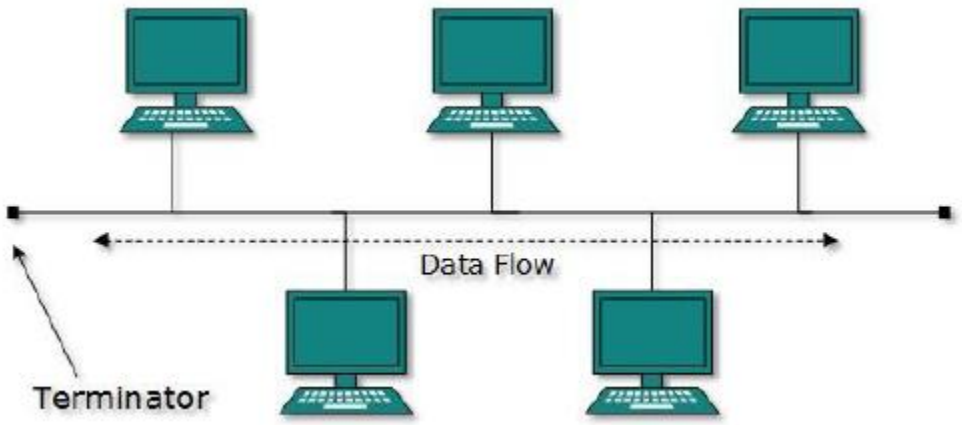
Point-to-point networks contains exactly two hosts (computer or switches or routers or servers) connected back to back using a single piece of cable. Often, the receiving end of one host is connected to sending end of the other end and vice-versa.



If the hosts are connected point-to-point logically, then may have multiple intermediate devices. But the end hosts are unaware of underlying network and see each other as if they are connected directly.

Bus Topology

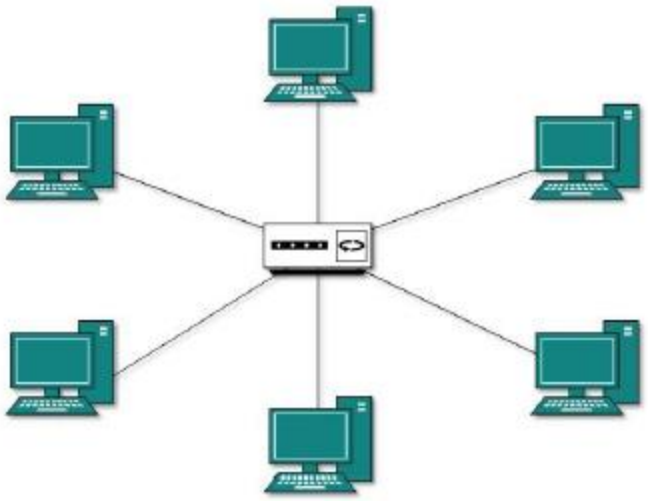
In contrast to point-to-point, in bus topology all device share single communication line or cable. All devices are connected to this shared line. Bus topology may have problem while more than one hosts sending data at the same time. Therefore, the bus topology either uses CSMA/CD technology or recognizes one host has Bus Master to solve the issue. It is one of the simple forms of networking where a failure of a device does not affect the others. But failure of the shared communication line make all other devices fail.



Both ends of the shared channel have line terminator. The data is sent in only one direction and as soon as it reaches the extreme end, the terminator removes the data from the line.

Star Topology

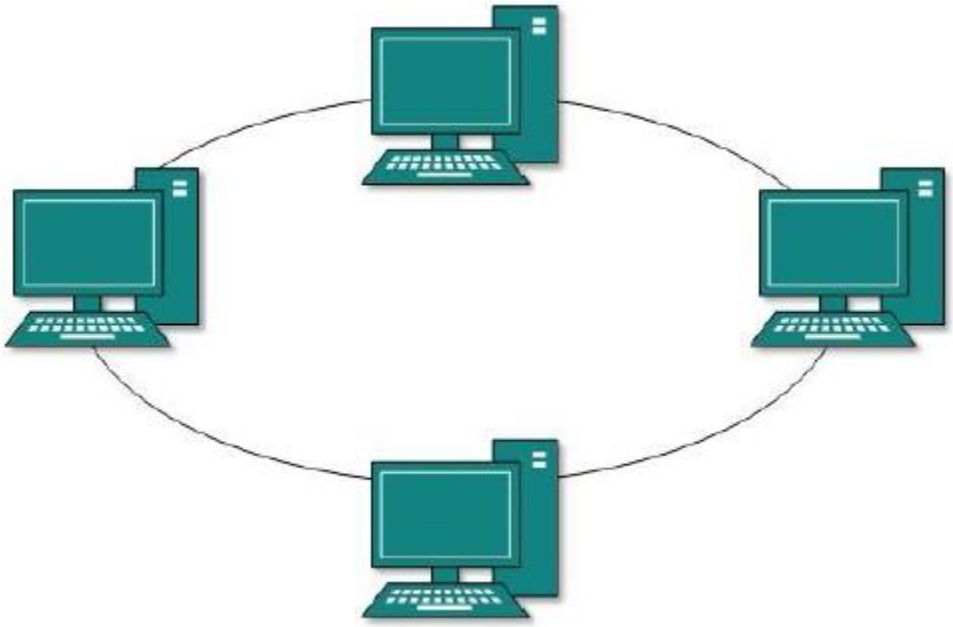
All hosts in star topology are connected to a central device, known as Hub device, using a point-to-point connection. That is, there exists a point to point connection between hosts and Hub. The hub device can be Layer-1 device (Hub / repeater) or Layer-2 device (Switch / Bridge) or Layer-3 device (Router / Gateway).



As in bus topology, hub acts as single point of failure. If hub fails, connectivity of all hosts to all other hosts fails. Every communication happens between hosts, goes through Hub only. Star topology is not expensive as to connect one more host, only one cable is required and configuration is simple.

Ring Topology

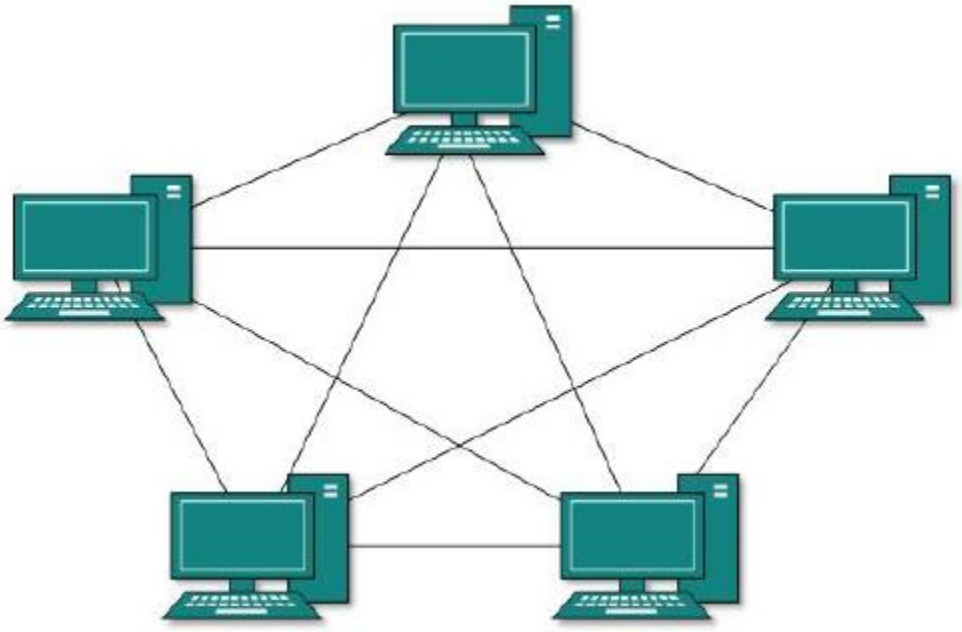
In ring topology, each host machine connects to exactly two other machines, creating a circular network structure. When one host tries to communicate or send message to a host which is not adjacent to it, the data travels through all intermediate hosts. To connect one more host in the existing structure administrator may need only one more extra cable.



Failure of any host results in failure of the whole ring. Thus every connection in the ring is point of failure. There exists a method which employs one more backup ring.

Mesh Topology

In this type of topology, a host is connected to one or two or more than two hosts. This topology may have hosts having point-to-point connection to every other hosts or may also have hosts which are having point to point connection to few hosts only

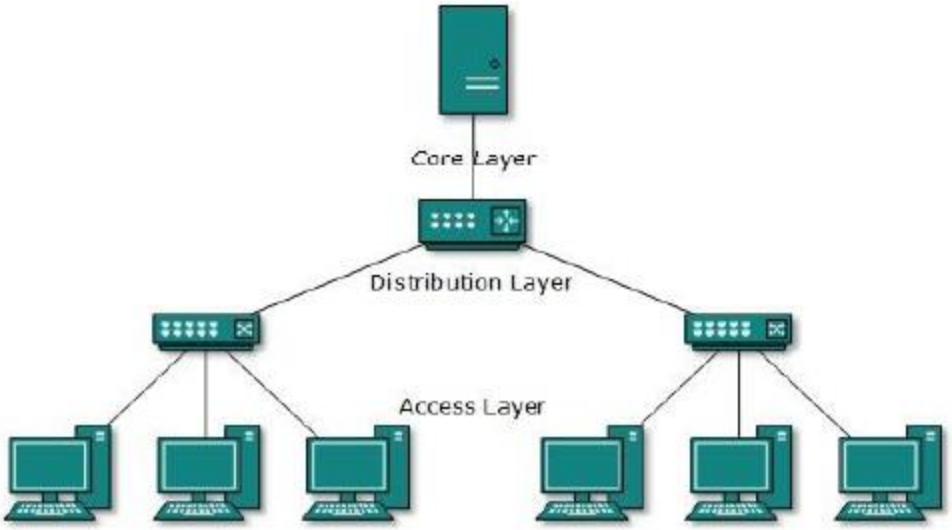


Hosts in Mesh topology also work as relay for other hosts which do not have direct point-to-point links. Mesh technology comes into two flavors:

- **Full Mesh**: All hosts have a point-to-point connection to every other host in the network. Thus for every new host n(n-1)/2 cables (connection) are required. It provides the most reliable network structure among all network topologies.
- **Partially Mesh**: Not all hosts have point-to-point connection to every other host. Hosts connect to each other in some arbitrarily fashion. This topology exists where we need to provide reliability to some host whereas others are not as such necessary.

Tree Topology

Also known as Hierarchical Topology is the most common form of network topology in use present day. This topology imitates as extended Star Topology and inherits properties of Bus topology. This topology divides the network in to multiple levels/layers of network. Mainly in LANs, a network is bifurcated into three types of network devices. The lowest most is access-layer where user‟s computer are attached. The middle layer is known as distribution layer, which works as mediator between upper layer and lower layer. The highest most layer is known as Core layer, and is central point of the network, i.e. root of the tree from which all nodes fork.



All neighboring hosts have point-to-point connection between them. Like bus topology, if the root goes down, the entire network suffers. Though it is not the single point of failure. Every connection serves as point of failure, failing of which divides the network into unreachable segment and so on.

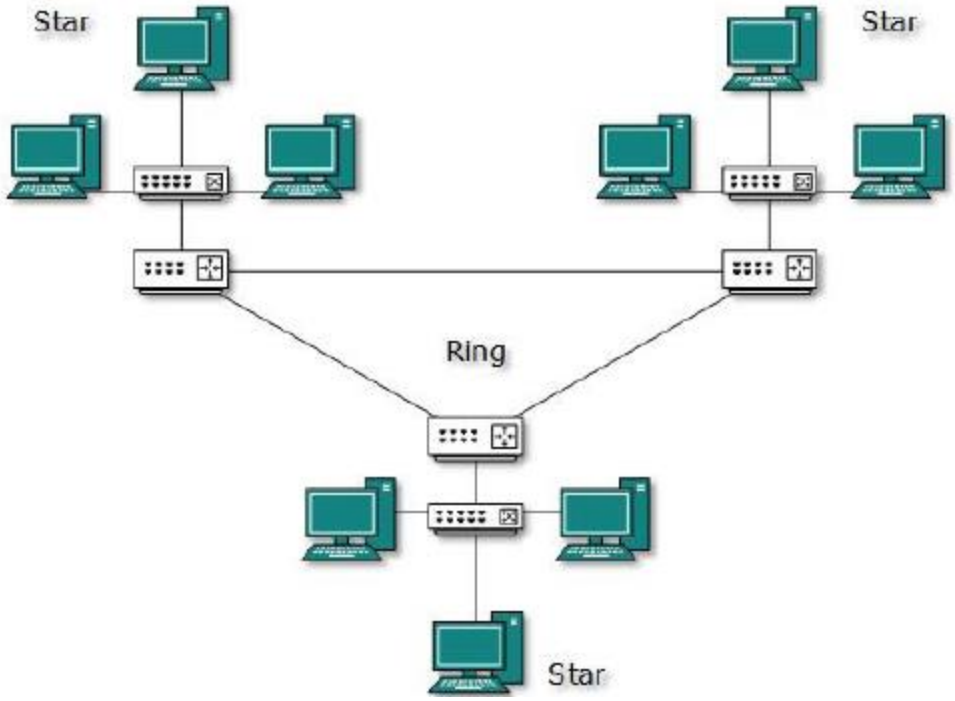
Daisy Chain

This topology connects all its hosts in a linear fashion. Similar to Ring topology, all hosts in this topology are connected to two hosts only, except the end hosts. That is if the end hosts in Daisy Chain are connected then it represents Ring topology.

Each link in Daisy chain topology represents single point of failure. Every link failure splits the network into two segments. Every intermediate host works as relay for its immediate hosts.

Hybrid Topology

A network structure whose design contains more than one topology is said to be Hybrid Topology. Hybrid topology inherits merits and demerits of all the incorporating topologies.



The above picture represents an arbitrarily Hybrid topology. The combining topologies may contain attributes of Star, Ring, Bus and Daisy-chain topologies. Most WANs are connected by means of dual Ring topology and networks connected to them are mostly Star topology networks. Internet is the best example of largest Hybrid topology

ACTIVITIES

1. Describe the concepts used in constructing a Network type. Use an example of your own to illustrate.
2. What are the 3 most important factors you would use in evaluating computer hardware and software? Explain. Why?

9.2 Internet

The **Internet** is a global system of interconnected computer networks that use the standard Internet protocol suite (TCP/IP) to link several billion devices worldwide. It is an international *network of networks* that consists of millions of private, public, academic, business, and government packet switched networks, linked by a broad array of electronic, wireless, and optical networking technologies. The Internet carries an extensive range of information resources and services, such as the inter-linked hypertext documents and applications of the World Wide Web (WWW), the infrastructure to support email, and peer-to-peer networks for file sharing and telephony.

The origins of the Internet date back to research commissioned by the United States government in the 1960s to build robust, fault-tolerant communication via computer networks. While this work, together with work in the United Kingdom and France, led to important precursor networks, they were not the Internet. There is no consensus on the exact date when the modern Internet came into being, but sometime in the early to mid-

1980s is considered reasonable. From that point, the network experienced decades of sustained exponential growth as generations of institutional, personal, and mobile computers were connected to it.

The funding of a new U.S. backbone by the National Science Foundation in the 1980s, as well as private funding for other commercial backbones, led to worldwide participation in the development of new networking technologies, and the merger of many networks. Though the Internet has been widely used by academia since the 1980s, the commercialization of what was by the 1990s an international network resulted in its popularization and incorporation into virtually every aspect of modern human life. As of June 2012, more than 2.4 billion people—over a third of the world's human population—have used the services of the Internet; approximately 100 times more people than were using it in 1995. Internet use grew rapidly in the West from the mid-1990s to early 2000s and from the late 1990s to present in the developing world. In 1994 only 3% of American classrooms had access to the Internet while by 2002 92% did.

Most traditional communications media including telephone, music, film, and television are being reshaped or redefined by the Internet, giving birth to new services such as voice over Internet Protocol (VoIP) and Internet Protocol television (IPTV). Newspaper, book, and other print publishing are adapting to website technology, or are reshaped into blogging and web feeds. The Internet has enabled and accelerated new forms of human interactions through instant messaging, Internet forums, and social networking. Online shopping has boomed both for major retail outlets and small artisans and traders. Business-to-business and financial services on the Internet affect supply chains across entire industries.

The Internet has no centralized governance in either technological implementation or policies for access and usage; each constituent network sets its own policies. Only the overreaching definitions of the two principal name spaces in the Internet, the Internet Protocol address space and the Domain Name System (DNS), are directed by a maintainer organization, the Internet Corporation for Assigned Names and Numbers (ICANN). The technical underpinning and standardization of the core protocols (IPv4 and IPv6) is an activity of the Internet Engineering Task Force (IETF), a non-profit organization of loosely affiliated international participants that anyone may associate with by contributing technical expertise.

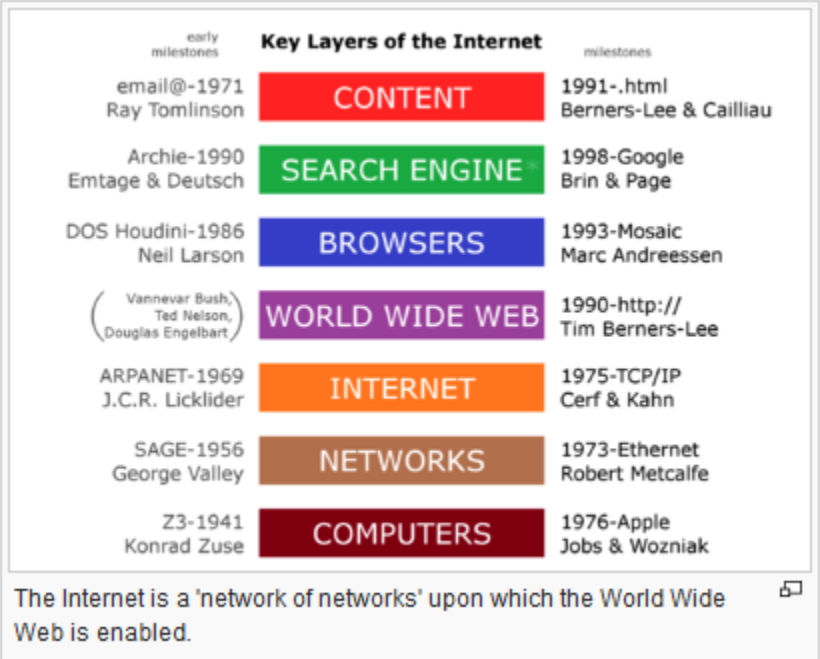
The *Internet*, referring to the specific global system of interconnected IP networks, is a proper noun and written with an initial capital letter. In the media and common use it is often not capitalized, viz. *the internet.* Some guides specify that the word should be capitalized when used as a noun, but not capitalized when used as a verb or an adjective.[9] The Internet is also often referred to as *the Net*.

Historically the word *internet* was used, un-capitalized, as early as 1883 as a verb and adjective to refer to interconnected motions. Starting in the early 1970s the term *internet* was used as a shorthand form of the

technical term internetwork, the result of interconnecting computer networks with special gateways or routers. It was also used as a verb meaning to connect together, especially for networks.

The terms *Internet* and *World Wide Web* are often used interchangeably in everyday speech; it is common to speak of "going on the Internet" when invoking a web browser to view web pages. However, the World Wide Web or *the Web* is just one of a very large number of services running on the Internet. The Web is a collection of interconnected documents (web pages) and other web resources, linked by hyperlinks and URLs. As another point of comparison, Hypertext Transfer Protocol, or HTTP, is the language used on the Web for information transfer, yet it is just one of many languages or protocols that can be used for communication on the Internet.[13] In addition to the Web, a multitude of other services are implemented over the Internet, including e-mail, file transfer, remote computer control, newsgroups, and online games. All of these services can be implemented on any intranet, accessible to network users.

The term *Interweb* is a portmanteau of *Internet* and *World Wide Web* typically used sarcastically to parody a technically un savvy user.



The Internet is a 'network of networks' upon which the World Wide Web is enabled.

Advantages:

1) Information on almost every subject imaginable.

2) Powerful search engines

3) Ability to do research from your home versus research libraries.

4) Information at various levels of study. Everything from scholarly articles to ones directed at children.

5) Message boards where people can discuss ideas on any topic. Ability to get wide range of opinions. People can find others that have a similar interest in whatever they are interested in.

6) The internet provides the ability of emails. Free mail service to anyone in the country.

7) Platform for products like SKYPE, which allow for holding a video conference with anyone in the world who also has access.

8) Friendships and love connections have been made over the internet by people involved in love/passion over similar interests.

9) Things such as Yahoo Answers and other sites where kids can have readily available help for homework.

10) News, of all kinds is available almost instantaneously. Commentary, on that news, from every conceivable viewpoint is also available.

Disadvantages:

1) There is a lot of wrong information on the internet. Anyone can post anything, and much of it is garbage.

2) There are predators that hang out on the internet waiting to get unsuspecting people in dangerous situations.

3) Some people are getting addicted to the internet and thus causing problems with their interactions of friends and loved ones.

4) Pornography that can get in the hands of young children too easily.

5) Easy to waste a lot of time on the internet. You can start surfing, and then realize far more time has passed than you realized. Internet and television together of added to the more sedentary lifestyles of people which further exacerbate the obesity problem.

6) Internet has a lot of "cheater" sites. People can buy essays and pass them off as their own far more easily than they used to be able to do.

7) There are a lot of unscrupulous businesses that have sprung up on the internet to take advantage of people.

8) Hackers can create viruses that can get into your personal computer and ruin valuable data.

9) Hackers can use the internet for identity theft.

10) It can be quite depressing to be on the internet and realize just how uneducated so many people have become in today's society.

Other advantages

1. Easy and cheap communication

Communicating with your friends and loved ones has been easy through e-mail and social communication sites like Facebook and MySpace. You don’t have to pay even a single cent just to chat with them because these services are free of charge!

2. Send small or big files with others easily!

If you have to send a file, for example, a video to your friend who’s living in other country, it isn’t practical nowadays to send him a package with the video cd. Instead, you can send him the video from your e-mail, or upload it in YouTube or other video sharing sites.

3. Loads of information

As I have mentioned earlier, internet has a lot of information that is very essential for the students so they don’t have to buy books or go to the library anymore. Search engines like Google and Yahoo! are always available when you need them.

4. Entertainment

Entertainment is one of the most popular reasons why many people prefer to surf the internet. There are a lot of games to play, videos to watch, and etc.

5. Services

Internet is making our life a lot easier by offering different services like online banking, online booking, hotel reservations, online shopping, and many more!

6. Earn money

Aside from entertainment, internet also lets you earn money while at the same time, enjoying what you’re doing! Like me, I’m a blogger and I love what I’m doing, and at the same time, I earn money.

7. Promote your product

Internet is one of the best and cheapest ways to promote your business or product. Starting from 10$, you can already have your own website and start reaching your potential customers. You can also use Multiply or Blogger if you don't want to spend money for your website.

DISADVANTAGES of INTERNET

1. Virus Threat

Most of the viruses came from the internet so be very careful when visiting or downloading from a site. If you are using Firefox, you can install Web of Trust add-on so you can get warnings about online scams, sites with adult content, and spam. With this, you can help keep your computer safe from online threats like spyware, adware, and viruses.

2. Theft of Personal Information

If you use the internet, there is a great risk of stealing your personal information such as name, address, credit card no., by those culprits.

3. Spamming

4. PORNOGRAPHY

This is probably the worst disadvantage of the internet especially for the parents who have kids.

SELF ASSESSMENT QUESTIONS

Fill in the blanks:

1. **If a computer on the network shares resources for others to use, it is called ____**
   a. Server
   b. Client
   c. Mainframe

2. **Terminators are used in _____ topology.**
   a. Bus
   b. Star

3. **. In _____ topology, if a computer's network cable is broken, whole network goes down.**
   a. Bus

b. Star

4. **For large networks, _____ topology is used.**
   a. Bus
   b. Star
   c. Ring

5. **ISO stands for**
   a. International Standard Organization
   b. International Student Organization
   c. Integrated Services Organization

9.3Extranet

An **extranet** is a computer network that allows controlled access from outside of an organization's intranet. Extranets are used for specific use cases including business-to-business (B2B). In a business-to-business context, an extranet can be viewed as an extension of an organization's intranet that is extended to users outside the organization, usually partners, vendors and suppliers, in isolation from all other Internet users. It is in context of that isolation that an extranet is different from an intranet or internet. In contrast, business-to-consumer (B2C) models involve known servers of one or more companies, communicating with previously unknown consumer users. An extranet is similar to a DMZ in that it provides access to needed services for channel partners, without granting access to an organization's entire network.

An extranet could be understood as an intranet mapped onto the public Internet or some other transmission system not accessible to the general public, but managed by more than one company's administrator(s). For example, military networks of different security levels may map onto a common military radio transmission system that never connects to the Internet. Any private network mapped onto a public one is a virtual private network (VPN), often using special security protocols.

For decades, institutions have been interconnecting to each other to create private networks for sharing information. One of the differences that characterizes an extranet, however, is that its interconnections are over a shared network rather than through dedicated physical lines. With respect to Internet Protocol networks, RFC 4364 states "If all the sites in a VPN are owned by the same enterprise, the VPN is a corporate intranet. If the various sites in a VPN are owned by different enterprises, the VPN is an extranet. A site can be in more than one VPN; e.g. in an intranet and several extranets. We regard both intranets and extranets as VPNs. In general, when we use the term VPN we will not be distinguishing between intranets and extranets. Even if this argument is valid, the term "extranet" is still applied and can be used to eliminate the use of the above description.

In the quote above from RFC 4364, the term "site" refers to a distinct networked environment. Two *sites* connected to each other across the public Internet backbone comprise a VPN. The term "site" does not mean

"website." Thus, a small company in a single building can have an "intranet," but to have a VPN, they would need to provide tunneled access to that network for geographically distributed employees.

Similarly, for smaller, geographically united organizations, "extranet" is a useful term to describe selective access to intranet systems granted to suppliers, customers, or other companies. Such access does not involve tunneling, but rather simply an authentication mechanism to a web server. In this sense, an "extranet" designates the "private part" of a website, where "registered users" can navigate, enabled by authentication mechanisms on a "login page".

An extranet requires network security. These can include firewalls, server management, the issuance and use of digital certificates or similar means of user authentication, encryption of messages and the use of virtual private networks (VPNs) that tunnel through the public network.

Many technical specifications describe methods of implementing extranets, but often never explicitly define an extranet. RFC 3457 presents requirements for remote access to extranets. RFC 2709 [2] discusses extranet implementation using IPsec and advanced network address translation (NAT).

Advantages:

- Exchange large volumes of data using Electronic Data Interchange (EDI)
- Share product catalogs exclusively with trade partners
- Collaborate with other companies on joint development efforts
- Jointly develop and use training programs with other companies
- Provide or access services provided by one company to a group of other companies, such as an online banking application managed by one company on behalf of affiliated banks

Disadvantages:

- Extranets can be expensive to implement and maintain within an organization (e.g., hardware, software, employee training costs), if hosted internally rather than by an application service provider.
- Security of extranets can be a concern when hosting valuable or proprietary information

9.4 Intranet

An **intranet** is a computer network that uses Internet Protocol technology to share information, operational systems, or computing services within an organization. This term is used in contrast to *extranet*, a network between organizations, and instead refers to a network within an organization. Sometimes, the term refers

only to the organization's internal website, but may be a more extensive part of the organization's information technology infrastructure, and may be composed of multiple local area networks. The objective is to organize each individual's desktop with minimal cost, time and effort to be more productive, cost efficient, timely, and competitive.

An intranet may host multiple private websites and constitute an important component and focal point of internal communication and collaboration. Any of the well known Internet protocols may be found in an intranet, such as HTTP (web services), SMTP (e-mail), and FTP (file transfer protocol). Internet technologies are often deployed to provide modern interfaces to legacy information systems hosting corporate data.

An intranet can be understood as a private analog of the Internet, or as a private extension of the Internet confined to an organization. The first intranet websites and home pages were published in 1991, and began to appear in non-educational organizations in 1994.

Intranets are sometimes contrasted to extranets. While intranets are generally restricted to employees of the organization, extranets may also be accessed by customers, suppliers, or other approved parties. Extranets extend a private network onto the Internet with special provisions for authentication, authorization and accounting (AAA protocol).

In many organizations, intranets are protected from unauthorized external access by means of a network gateway and firewall. For smaller companies, intranets may be created simply by using private IP address ranges. In these cases, the intranet can only be directly accessed from a computer in the local network; however, companies may provide access to off-site employees by using a virtual private network, or by other access methods, requiring user authentication and encryption.

Uses

Increasingly, intranets are being used to deliver tools, e.g. *collaboration* (to facilitate working in groups and teleconferencing) or sophisticated corporate directories, sales and customer relationship management tools, project management etc., to advance productivity.

Intranets are also being used as corporate culture-change platforms. For example, large numbers of employees discussing key issues in an intranet forum application could lead to new ideas in management, productivity, quality, and other corporate issues.

In large intranets, website traffic is often similar to public website traffic and can be better understood by using web metrics software to track overall activity. User surveys also improve intranet website

effectiveness. Larger businesses allow users within their intranet to access public internet through firewall servers. They have the ability to screen messages coming and going keeping security intact.

When a part of an intranet is made accessible to customers and others outside the business, that part becomes part of an extranet. Businesses can send private messages through the public network, using special encryption/decryption and other security safeguards to connect one part of their intranet to another.

Intranet user-experience, editorial, and technology teams work together to produce in-house sites. Most commonly, intranets are managed by the communications, HR or CIO departments of large organizations, or some combination of these.

Because of the scope and variety of content and the number of system interfaces, intranets of many organizations are much more complex than their respective public websites. Intranets and their use are growing rapidly. According to the Intranet design annual 2007 from Nielsen Norman Group, the number of pages on participants' intranets averaged 200,000 over the years 2001 to 2003 and has grown to an average of 6 million pages over 2005–2007.[

Benefits

- **Workforce productivity**: Intranets can help users to locate and view information faster and use applications relevant to their roles and responsibilities. With the help of a web browser interface, users can access data held in any database the organization wants to make available, anytime and — subject to security provisions — from anywhere within the company workstations, increasing employees' ability to perform their jobs faster, more accurately, and with confidence that they have the right information. It also helps to improve the services provided to the users.
- **Time**: Intranets allow organizations to distribute information to employees on an *as-needed* basis; Employees may link to relevant information at their convenience, rather than being distracted indiscriminately by email.
- **Communication**: Intranets can serve as powerful tools for communication within an organization, vertically strategic initiatives that have a global reach throughout the organization. The type of information that can easily be conveyed is the purpose of the initiative and what the initiative is aiming to achieve, who is driving the initiative, results achieved to date, and who to speak to for more information. By providing this information on the intranet, staffs have the opportunity to keep up-to-date with the strategic focus of the organization. Some examples of communication would be chat, email, and/or blogs. A great real world example of where an intranet helped a company communicate is when Nestle had a number of food processing plants in Scandinavia. Their central support system had to deal with a number of queries every day. When Nestle decided to invest in an intranet, they

quickly realized the savings. McGovern says the savings from the reduction in query calls was substantially greater than the investment in the intranet.

- **Web publishing** allows cumbersome corporate knowledge to be maintained and easily accessed throughout the company using hypermedia and Web technologies. Examples include: employee manuals, benefits documents, company policies, business standards, news feeds, and even training, can be accessed using common Internet standards (Acrobat files, Flash files, CGI applications). Because each business unit can update the online copy of a document, the most recent version is usually available to employees using the intranet.
- **Business operations and management**: Intranets are also being used as a platform for developing and deploying applications to support business operations and decisions across the internetworked enterprise.
- **Cost-effective**: Users can view information and data via web-browser rather than maintaining physical documents such as procedure manuals, internal phone list and requisition forms. This can potentially save the business money on printing, duplicating documents, and the environment as well as document maintenance overhead. For example, the HRM company PeopleSoft "derived significant cost savings by shifting HR processes to the intranet". McGovern goes on to say the manual cost of enrolling in benefits was found to be USD109.48 per enrollment. "Shifting this process to the intranet reduced the cost per enrollment to $21.79; a saving of 80 percent". Another company that saved money on expense reports was Cisco. "In 1996, Cisco processed 54,000 reports and the amount of dollars processed was USD19 million".[6]
- **Enhance collaboration**: Information is easily accessible by all authorized users, which enables teamwork.
- **Cross-platform capability**: Standards-compliant web browsers are available for Windows, Mac, and UNIX.
- **Built for one audience**: Many companies dictate computer specifications which, in turn, may allow Intranet developers to write applications that only have to work on one browser (no cross-browser compatibility issues). Being able to specifically address your "viewer" is a great advantage. Since Intranets are user-specific (requiring database/network authentication prior to access), you know exactly who you are interfacing with and can personalize your Intranet based on role (job title, department) or individual ("Congratulations Jane, on your 3rd year with our company!").
- **Promote common corporate culture**: Every user has the ability to view the same information within the Intranet.
- **Immediate updates**: When dealing with the public in any capacity, laws, specifications, and parameters can change. Intranets make it possible to provide your audience with "live" changes so they are kept up-to-date, which can limit a company's liability.

- **Supports a distributed computing architecture**: The intranet can also be linked to a company's management information system, for example a time keeping system

Intranet software

Microsoft SharePoint is the dominant software used for creating intranets. Estimates indicate that around 50% of all intranets are developed using SharePoint, however there are many alternatives. Other intranet software includes

• Autonomy Corporation	• Joomla
• Atlassian Confluence	• Liferay
• Bitrix24	• Lotus Notes
• Drupal	• OpenText
• eXo Platform	• Oracle Fusion Middleware
• Google Sites	• Plone (software)
• Igloo Software	• SAP NetWeaver Portal
• IBM Websphere	• Sitecore
• Interact Intranet	• ThoughtFarmer
• Hyperoffice	• WordPress
• Jive Software	• Yammer

9.5World Wide Web

"WWW" and "The web" redirect here. For other uses of WWW, see WWW (disambiguation). For other uses of web, see Web (disambiguation).

The **World Wide Web** (abbreviated as **WWW** or **W3**, commonly known as **the Web**) is a system of interlinked hypertext documents that are accessed via the Internet. With a web browser, one can view web pages that may contain text, images, videos, and other multimedia and navigate between them via hyperlinks.

It is important to know that this is not a synonym for the Internet. The World Wide Web, or just "the Web," as ordinary people call it, is a subset of the Internet. The Web consists of pages that can be accessed using a Web browser. The Internet is the actual network of networks where all the information resides. Things like Telnet, FTP, Internet gaming, Internet Relay Chat (IRC), and e-mail are all part of the Internet, but are not part of the World Wide Web. The Hyper-Text Transfer Protocol (HTTP) is the method used to transfer Web pages to your computer. With hypertext, a word or phrase can contain a link to another Web site. All Web pages are written in the hyper-text markup language (HTML), which works in conjunction with HTTP.

Tim Berners-Lee, a British computer scientist and former CERN employee, is considered the inventor of the Web. On March 12, 1989, Berners-Lee wrote a proposal for what would eventually become the World Wide Web. The 1989 proposal was meant for a more effective CERN communication system but Berners-Lee eventually realized the concept could be implemented throughout the world. Berners-Lee and Belgian computer scientist Robert Cailliau proposed in 1990 to use hypertext "to link and access information of various kinds as a web of nodes in which the user can browse at will", and Berners-Lee finished the first website in December of that year. The first test was completed around 20 December 1990 and Berners-Lee reported about the project on the newsgroup *alt.hypertext* on 7 August 1991

Many hostnames used for the World Wide Web begin with *www* because of the long-standing practice of naming Internet hosts according to the services they provide. The hostname of a web server is often *www*, in the same way that it may be *ftp* for an FTP server, and *news* or *nntp* for a USENET news server. These host names appear as Domain Name System (DNS) or subdomain names, as in *www.example.com*. The use of *www* is not required by any technical or policy standard and many web sites do not use it; indeed, the first ever web server was called *nxoc01.cern.ch.* According to Paolo Palazzi, who worked at CERN along with Tim Berners-Lee, the popular use of *www* as sub domain was accidental; the World Wide Web project page was intended to be published at www.cern.ch while info.cern.ch was intended to be the CERN home page, however the DNS records were never switched, and the practice of pretending *www* to an institution's website domain name was subsequently copied. Many established websites still use the prefix, or they employ other subdomain names such as *www2*, *secure* or *en* for special purposes. Many such web servers are set up so that both the main domain name (e.g., example.com) and the *www* sub domain (e.g., www.example.com) refer to the same site; others require one form or the other, or they may map to different web sites.

The use of a sub domain name is useful for load balancing incoming web traffic by creating a CNAME record that points to a cluster of web servers. Since, currently, only a sub domain can be used in a CNAME, the same result cannot be achieved by using the bare domain root.

When a user submits an incomplete domain name to a web browser in its address bar input field, some web browsers automatically try adding the prefix "www" to the beginning of it and possibly ".com", ".org" and ".net" at the end, depending on what might be missing. For example, entering 'Microsoft' may be transformed to *http://www.microsoft.com/* and 'open office' to *http://www.openoffice.org*. This feature started appearing in early versions of Mozilla Firefox, when it still had the working title 'Firebird' in early 2003, from an earlier practice in browsers such as Lynx. It is reported that Microsoft was granted a US patent for the same idea in 2008, but only for mobile devices.

Use of the www prefix is declining as Web 2.0 web applications seek to brand their domain names and make them easily pronounceable. As the mobile web grows in popularity, services like Gmail.com, MySpace.com, Facebook.com and Twitter.com are most often mentioned without adding "www." (or, indeed, ".com") to the domain.

Function

The terms *Internet* and *World Wide Web* are often used without much distinction. However, the two things are not the same. The Internet is a global system of interconnected computer networks. In contrast, the World Wide Web is one of the services transferred over these networks. It is a collection of text documents and other resources, linked by hyperlinks and URLs, usually accessed by web browsers from web servers.

Viewing a web page on the World Wide Web normally begins either by typing the URL of the page into a web browser, or by following a hyperlink to that page or resource. The web browser then initiates a series of background communication messages to fetch and display the requested page. In the 1990s, using a browser to view web pages—and to move from one web page to another through hyperlinks—came to be known as 'browsing,' 'web surfing,' (after channel surfing), or 'navigating the Web'. Early studies of this new behavior investigated user patterns in using web browsers. One study, for example, found five user patterns: exploratory surfing, window surfing, evolved surfing, bounded navigation and targeted navigation.

The following example demonstrates the functioning of web browser when accessing a page at the URL http://example.org/wiki/World_Wide_Web. The browser resolves the server name of the URL (*example.org*) into an Internet Protocol address using the globally distributed Domain Name System (DNS). This lookup returns an IP address such as *203.0.113.4*. The browser then requests the resource by sending an HTTP request across the Internet to the computer at that address. It requests service from a specific TCP port number that is well known for the HTTP service, so that the receiving host can distinguish an HTTP request from other network protocols it may be servicing. The HTTP protocol normally uses port number 80. The content of the HTTP request can be as simple as two lines of text:

GET */wiki/World_Wide_Web* HTTP/1.1
Host: *example.org*
The computer receiving the HTTP request delivers it to web server software listening for requests on port 80. If the web server can fulfill the request it sends an HTTP response back to the browser indicating success:
HTTP/1.0 200 OK
Content-Type: text/html; charset=UTF-8

followed by the content of the requested page. The Hypertext Markup Language for a basic web page looks like <html> <head> <title>Example.org – The World Wide Web</title> </head> <body> The World Wide Web, abbreviated as WWW and commonly known ... </body> </html>

The web browser parses the HTML and interprets the markup (<title>, for paragraph, and such) that surrounds the words to format the text on the screen. Many web pages use HTML to reference the URLs of other resources such as images, other embedded media, scripts that affect page behavior, and Cascading Style Sheets that affect page layout. The browser makes additional HTTP requests to the web server for these other Internet media types. As it receives their content from the web server, the browser progressively renders the page onto the screen as specified by its HTML and these additional resources

9.6 Search engines

A **web search engine** is a software system that is designed to search for information on the World Wide Web. The search results are generally presented in a line of results often referred to as search engine results pages (SERPs). The information may be a mix of web pages, images, and other types of files. Some search engines also mine data available in databases or open directories. Unlike web directories, which are maintained only by human editors, search engines also maintain real-time information by running an algorithm on a web crawler.

Example: Google, Bing, Yahoo, Baidu, AOL, Ask, Excite etc.

How Web Search Engines work?

A search engine operates in the following order:

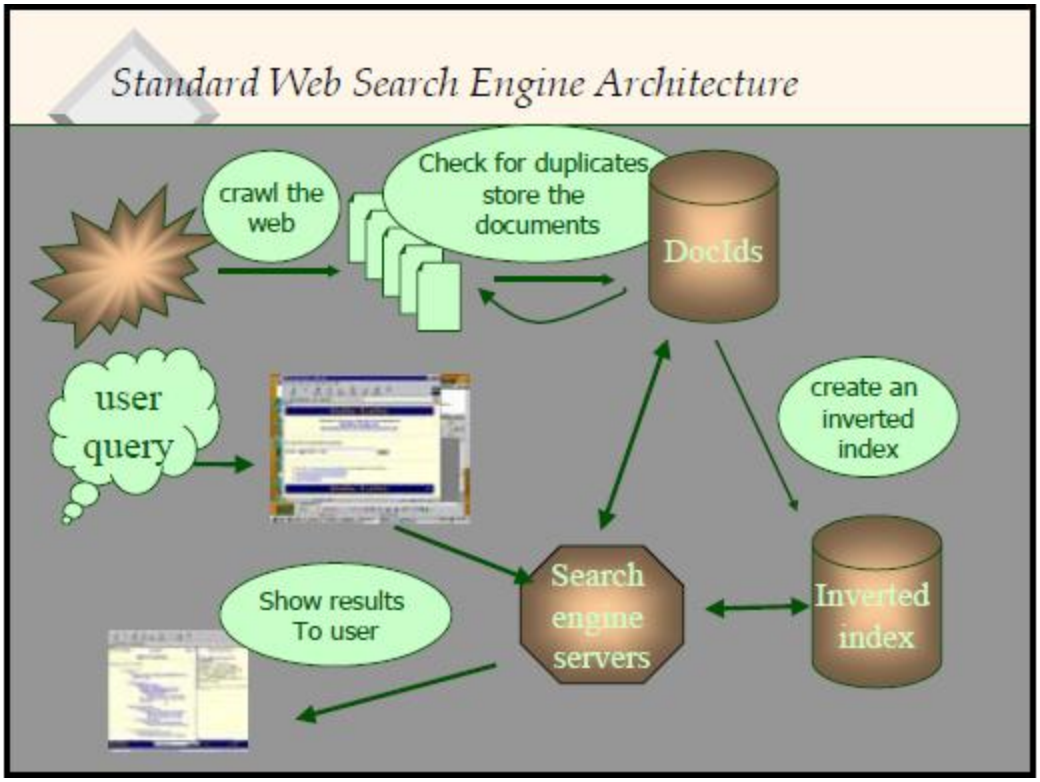
1. Web crawling
2. Indexing
3. Searching

Web search engines work by storing information about many web pages, which they retrieve from the HTML markup of the pages. These pages are retrieved by a Web crawler (sometimes also known as a spider) — an automated Web crawler which follows every link on the site. The site owner can exclude specific pages by using robots.txt.

The search engine then analyzes the contents of each page to determine how it should be indexed (for example, words can be extracted from the titles, page content, headings, or special fields called meta tags). Data about web pages are stored in an index database for use in later queries. A query from a user can be a single word. The index helps find information relating to the query as quickly as possible. Some search

engines, such as Google, store all or part of the source page (referred to as a cache) as well as information about the web pages, whereas others, such as AltaVista, store every word of every page they find. This cached page always holds the actual search text since it is the one that was actually indexed, so it can be very useful when the content of the current page has been updated and the search terms are no longer in it. This problem might be considered a mild form of linkrot, and Google's handling of it increases usability by satisfying user expectations that the search terms will be on the returned webpage. This satisfies the principle of least astonishment, since the user normally expects that the search terms will be on the returned pages. Increased search relevance makes these cached pages very useful as they may contain data that may no longer be available elsewhere.

When a user enters a query into a search engine (typically by using Glossary), the engine examines its index and provides a listing of best-matching web pages according to its criteria, usually with a short summary containing the document's title and sometimes parts of the text. The index is built from the information stored with the data and the method by which the information is indexed. From 2007 the Google.com search engine has allowed one to search by date by clicking "Show search tools" in the leftmost column of the initial search results page, and then selecting the desired date range. Most search engines support the use of the Boolean operators AND, OR and NOT to further specify the search query. Boolean operators are for literal searches that allow the user to refine and extend the terms of the search. The engine looks for the words or phrases exactly as entered. Some search engines provide an advanced feature called proximity search, which allows users to define the distance between Glossary. There is also concept-based searching where the research involves using statistical analysis on pages containing the words or phrases you search for. As well, natural language queries allow the user to type a question in the same form one would ask it to a human. A site like this would be ask.com.



The usefulness of a search engine depends on the relevance of the **result set** it gives back. While there may be millions of web pages that include a particular word or phrase, some pages may be more relevant, popular, or authoritative than others. Most search engines employ methods to rank the results to provide the "best" results first. How a search engine decides which pages are the best matches, and what order the results should be shown in, varies widely from one engine to another. The methods also change over time as Internet usage changes and new techniques evolve. There are two main types of search engine that have evolved: one is a system of predefined and hierarchically ordered Glossary that humans have programmed extensively. The other is a system that generates an "inverted index" by analyzing texts it locates. This first form relies much more heavily on the computer itself to do the bulk of the work.

Most Web search engines are commercial ventures supported by advertising revenue and thus some of them allow advertisers to have their listings ranked higher in search results for a fee. Search engines that do not accept money for their search results make money by running search related ads alongside the regular search engine results. The search engines make money every time someone clicks on one of these ads.

9.7 E-Commerce

E-commerce (electronic commerce or EC) is the buying and selling of goods and services, or the transmitting of funds or data, over an electronic network, primarily the Internet. These business transactions occur either business-to-business, business-to-consumer, consumer-to-consumer or consumer-to-business. The terms *e-*

commerce and *e-business* are often used interchangeably. The term *e-tail* is also sometimes used in reference to transactional processes around online retail.

Electronic Commerce, commonly known as **E-commerce** or **eCommerce**, is trading in products or services using computer networks, such as the Internet. Electronic commerce draws on technologies such as mobile commerce, electronic funds transfer, supply chain management, Internet marketing, online transaction processing, electronic data interchange (EDI), inventory management systems, and automated data collection systems. Modern electronic commerce typically uses the World Wide Web for at least one part of the transaction's life cycle, although it may also use other technologies such as e-mail.

E-commerce businesses usually employ some or all of the following practices:

- Provide Etail or virtual storefront on websites with online catalogs, sometimes gathered into a "virtual mall"
- Buy or sell on online marketplaces.
- Gather and use demographic data through web contacts and social media.
- Use electronic data interchange, the business-to-business exchange of data.
- Reach prospective and established customers by e-mail or fax (for example, with newsletters).
- Use business-to-business buying and selling.
- Provide secure business transactions.

E-commerce is conducted using a variety of applications, such as email, fax, online catalogs and shopping carts, Electronic Data Interchange (EDI), File Transfer Protocol, and Web services. Most of this is business-to-business, with some companies attempting to use email and fax for unsolicited ads (usually viewed as spam) to consumers and other business prospects, as well as to send out e-newsletters to subscribers.

The benefits of e-commerce include its around-the-clock availability, the speed of access, a wider selection of goods and services, accessibility, and international reach. It’s perceived downsides include sometimes-limited customer service, not being able to see or touch a product prior to purchase, and the necessitated wait time for product shipping.

To ensure the security, privacy and effectiveness of e-commerce, businesses should authenticate business transactions, control access to resources such as web pages for registered or selected users, encrypt communications and implement security technologies such as the Secure Sockets Layer.

Key Drivers of E-commerce

Key drivers	*Measurement criteria*
Technological factors	• Telecommunications infrastructure Backbone infrastructure and architecture Industry players and competition Pricing Internet service providers Range of services available (e.g. ADSL, ISDN) Ownership (private or public sector) • Access to new technology developments • Bandwidth • Speed of development and implementation of new technology by industry sector
Political factors	• Number and type of government incentives and programmes to support the use and development of new technology • Legislation – number and type of supportive or restrictive laws and policies that govern electronic data, contacts and financial transactions. For example, laws that recognise and enforce the validity of electronic documentation, contracts and transactions in a court of law; the validation of digital signatures; the legal usage of electronic security measures such as encryption • Public policies – whether government supports the growth of electronic transactions and processes. For example, filing tax returns to the Inland Revenue electronically, the national education curriculum and training
Social factors	• Skills of workforce • Number of users on-line • Penetration rate of PCs • Level of education; computer literacy and IT skills • Culture of technophilia – a willingness and ability to adopt new technology and the speed at which technology achieves critical mass as in Japan
Economic factors	• Economic growth – GDP • Average income • Cost of technology (hardware and software) • Cost of access to telecommunications infrastructure – pricing structures and rates • Commercial infrastructure – advancement of banking sector; payment systems • Innovative business models

Benefits of e-commerce to organizations

- *International marketplace*. What used to be a single physical marketplace located in a geographical area has now become a borderless marketplace including national and international markets. By becoming e-commerce enabled, businesses now have access to people all around the world. In effect all e-commerce businesses have become virtual multinational corporations.
- *Operational cost savings*. The cost of creating, processing, distributing, storing and retrieving paper-based information has decreased.
- *Mass customization*. E-commerce has revolutionized the way consumers buy goods and services. The pull-type processing allows for products and services to be customized to the customer's requirements. In the past when Ford first started making motor cars, customers could have any color so long as it was black. Now customers can configure a car according to their specifications within minutes on-line via the www.ford.com website.
- *Enables reduced inventories* and overheads by facilitating 'pull'-type supply chain management – this is based on collecting the customer order and then delivering through JIT (just-in-time) manufacturing. This is particularly beneficial for companies in the high technology sector, where

stocks of components held could quickly become obsolete within months. For example, companies like Motorola (mobile phones), and Dell (computers) gather customer orders for a product, transmit them electronically to the manufacturing plant where they are manufactured according to the customer's specifications (like color and features) and then sent to the customer within a few days.

- *Lower telecommunications cost*. The Internet is much cheaper than value added networks (VANs) which were based on leasing telephone lines for the sole use of the organization and its authorized partners. It is also cheaper to send a fax or e-mail via the Internet than direct dialing.
- *Digitization of products and processes*. Particularly in the case of software and music/video products, which can be downloaded or e-mailed directly to customers via the Internet in digital or electronic format?
- *No more 24-hour-time constraints*. Businesses can be contacted by or contact customers or suppliers at any time.

Benefits of e-commerce to consumers

- *24/7 access*. Enables customers to shop or conduct other transactions 24 hours a day, all year round from almost any location. For example, checking balances, making payments, obtaining travel and other information. In one case a pop star set up web cameras in every room in his house, so that he could check the status of his home by logging onto the Internet when he was away from home on tour.
- *More choices*. Customers not only have a whole range of products that they can choose from and customize, but also an international selection of suppliers.
- *Price comparisons*. Customers can 'shop' around the world and conduct comparisons either directly by visiting different sites, or by visiting a single site where prices are aggregated from a number of providers and compared (for example www.moneyextra.co.uk for financial products and services).
- *Improved delivery processes*. This can range from the immediate delivery of digitized or electronic goods such as software or audio-visual files by downloading via the Internet, to the on-line tracking of the progress of packages being delivered by mail or courier.
- *An environment of competition* where substantial discounts can be found or value added, as different retailers vie for customers. It also allows many individual customers to aggregate their orders together into a single order presented to wholesalers or manufacturers and obtain a more competitive price (aggregate buying), for example www.letsbuyit.com.

9.8 Summary

A Network Topology is the way computer systems or network equipment connected to each other. Topologies may define both physical and logical aspect of the network. Both logical and physical topologies could be same or different in a same network.

Point-to-point networks contains exactly two hosts (computer or switches or routers or servers) connected back to back using a single piece of cable. Often, the receiving end of one host is connected to sending end of the other end and vice-versa.

Bus topology may have problem while more than one hosts sending data at the same time. Therefore, the bus topology either uses CSMA/CD technology or recognizes one host has Bus Master to solve the issue.

In ring topology, each host machine connects to exactly two other machines, creating a circular network structure. When one host tries to communicate or send message to a host which is not adjacent to it, the data travels through all intermediate hosts. To connect one more host in the existing structure administrator may need only one more extra cable

In Mesh topology, a host is connected to one or two or more than two hosts. This topology may have hosts having point-to-point connection to every other hosts or may also have hosts which are having point to point connection to few hosts only.

Tree Topology is known as Hierarchical Topology is the most common form of network topology in use present day. This topology imitates as extended Star Topology and inherits properties of Bus topology. This topology divides the network in to multiple levels/layers of network. Mainly in LANs, a network is bifurcated into three types of network devices.

The **Internet** is a global system of interconnected computer networks that use the standard Internet protocol suite (TCP/IP) to link several billion devices worldwide. It is an international *network of networks* that consists of millions of private, public, academic, business, and government packet switched networks, linked by a broad array of electronic, wireless, and optical networking technologies. The Internet carries an extensive range of information resources and services, such as the inter-linked hypertext documents and applications of the World Wide Web (WWW), the infrastructure to support email, and peer-to-peer networks for file sharing and telephony.

An **extranet** is a computer network that allows controlled access from outside of an organization's intranet. Extranets are used for specific use cases including business-to-business (B2B). In a business-to-business context, an extranet can be viewed as an extension of an organization's intranet that is extended to users outside the organization, usually partners, vendors and suppliers, in isolation from all other Internet users. It is in context of that isolation that an extranet is different from an intranet or internet. In contrast, business-to-

consumer (B2C) models involve known servers of one or more companies, communicating with previously unknown consumer users. An extranet is similar to a DMZ in that it provides access to needed services for channel partners, without granting access to an organization's entire network.

An **intranet** is a computer network that uses Internet Protocol technology to share information, operational systems, or computing services within an organization. This term is used in contrast to *extranet*, a network between organizations, and instead refers to a network within an organization. Sometimes, the term refers only to the organization's internal website, but may be a more extensive part of the organization's information technology infrastructure, and may be composed of multiple local area networks. The objective is to organize each individual's desktop with minimal cost, time and effort to be more productive, cost efficient, timely, and competitive.

The **World Wide Web** (abbreviated as **WWW** or **W3**, commonly known as **the Web**) is a system of interlinked hypertext documents that are accessed via the Internet. With a web browser, one can view web pages that may contain text, images, videos, and other multimedia and navigate between them via hyperlinks.

A **web search engine** is a software system that is designed to search for information on the World Wide Web. The search results are generally presented in a line of results often referred to as search engine results pages (SERPs). The information may be a mix of web pages, images, and other types of files. Some search engines also mine data available in databases or open directories. Unlike web directories, which are maintained only by human editors, search engines also maintain real-time information by running an algorithm on a web crawler.

Electronic Commerce, commonly known as **E-commerce** or **eCommerce**, is trading in products or services using computer networks, such as the Internet. Electronic commerce draws on technologies such as mobile commerce, electronic funds transfer, supply chain management, Internet marketing, online transaction processing, electronic data interchange (EDI), inventory management systems, and automated data collection systems. Modern electronic commerce typically uses the World Wide Web for at least one part of the transaction's life cycle, although it may also use other technologies such as e-mail.

9.9 Glossary

Internet: A means of connecting a computer to any other computer anywhere in the world via dedicated routers and servers. When two computers are connected over the **Internet**, they can send and receive all kinds of information such as text, graphics, voice, video, and computer programs.

Extranet:

Intranet: a computer network with restricted access, as within a company, that uses software and protocols developed for the Internet.

World Wide Web: The **World Wide Web** (abbreviated as **WWW** or **W3**, commonly known as **the Web**) is a system of interlinked hypertext documents that are accessed via the Internet. With a web browser, one can view web pages that may contain text, images, videos, and other multimedia and navigate between them via hyperlinks.

Search Engine: a program that searches for and identifies items in a database that correspond to keywords or characters specified by the user, used especially for finding particular sites on the World Wide Web.

E-Commerce: The buying and selling of products and services by businesses and consumers through an electronic medium, without using any paper documents. E-commerce is widely considered the buying and selling of products over the internet, but any transaction that is completed solely through electronic measures can be considered e-commerce. E-commerce is subdivided into three categories: business to business or B2B (Cisco), business to consumer or B2C (Amazon), and consumer to consumer or C2C (eBay). **also called** electronic commerce.

EDI: Electronic data interchange is the computer-to-computer exchange of business documents in an electronic format between business partners.

9.10 References

O'Brien, J (1999). *Management Information Systems – Managing Information Technology in the Internetworked Enterprise*. Boston: Irwin McGraw-Hill. ISBN 0-07-112373-3.

Laudon, Kenneth C.; Laudon, Jane P. (2009). *Management Information Systems: Managing the Digital Firm* (11 ed.). Prentice Hall/CourseSmart. p. 164.

Transaction processing systems (TPS) collect and record the routine transactions of an organization. Examples of such systems are sales order entry, hotel reservations, payroll, employee record keeping, and shipping.

Bidgoli, Hossein, (2004). The Internet Encyclopedia, Volume 1, John Wiley & Sons, Inc. p. 707.

Pant, S., Hsu, C., (1995), Strategic Information Systems Planning: A Review, Information Resources Management Association International Conference, May 21–24, Atlanta.

9.10 Further Readings

Taylor, Victoria. "Supply Chain Management: The Next Big Thing?". *Sept. 12, 2011*. Business Week. Retrieved 5 March 2014.

Lynn, Samara. "What is CRM?". PC Mag. Retrieved 5 March 2014.

Joshi, Girdhar (2013). *Management Information Systems*. New Delhi: Oxford University Press. p. 328. ISBN 9780198080992.

Laudon, K.,&Laudon, J. (2010). Management information systems: Managing the digital firm. (11th ed.). Upper Saddle River, NJ: Pearson Prentice Hall

9.10 Model Questions

Q:1 Discuss in detail about various Network topologies.

Q:2 What do you understand by terms: Internet, Extranet and Intranet.

Q:3 Explain in detail about the World Wide Web.

Q:4 Explain about the concept of E-commerce.

ANSWERS TO SELF ASSESSMENT QUESTIONS

1. A
2. A
3. A
4. B
5. A

CHAPTER 10
WEB TECHNOLOGIES

Structure

10.0 Objectives

After studying this Chapter you will be able to:

1. Understand about the various tools for making web page operational web servers
2. Understand about Data warehouse and data marts
3. Understand about the concept of Data Mining
4. Understand about Internet 2 and mobile communication

10.1Tools and Technologies for making web page operational web servers

A **web page** (or **webpage**) is a web document that is suitable for the World Wide Web and the *web browser*.

A web browser displays a web page on a monitor or mobile device. The web page is what displays, but the term also refers to a computer file, usually written in HTML or comparable markup language. Web browsers coordinate the various web resource elements for the written web page, such as style sheets, scripts and images, to present the web page.

Typical web pages provide hypertext that include a navigation bar or a sidebar menu to *other* web pages via hyperlinks, often referred to as *links*.

On a network, a web browser can retrieve a web page from a remote web server. On a higher level, the web server may restrict access to only a private network such as a corporate intranet or it provides access to the World Wide Web. On a lower level, the web browser uses the Hypertext Transfer Protocol (HTTP) to make such requests.

A *static web page* is delivered exactly as stored, as web content in the web server's file system, while a *dynamic web page* is generated by a web application that is driven by server-side software or client-side scripting. Dynamic web pages help the browser (the client) to enhance the web page through user input to the server.

Creation of a web page

To create a web page, a text editor or a specialized HTML editor is needed. In order to upload the created web page to a web server, traditionally an FTP client is needed.

The design of a web page is highly personal. A design can be made according to one's own preference, or a premade web template can be used. Web templates let web page designers edit the content of a web page without having to worry about the overall aesthetics. Many people publish their own web pages using products like Tripod, or Angelfire. These web publishing tools offer free page creation and hosting up to a certain size limit. Other ways of making a web page is to download specialized software, like a Wiki, CMS, or forum. These options allow for quick and easy creation of a web page which is typically dynamic.

A **web server** is a computer system that processes requests via HTTP, the basic network protocol used to distribute information on the World Wide Web. The term can refer either to the entire system, or specifically to the software that accepts and supervises the HTTP requests.

A Web server is a program that, using the client/server model and the World Wide Web's Hypertext Transfer Protocol (HTTP), serves the files that form Web pages to Web users (whose computers contain HTTP clients that forward their requests). Every computer on the Internet that contains a Web site must have a Web server program. Two leading Web servers are Apache , the most widely-installed Web server, and Microsoft's Internet Information Server (IIS). Other Web servers include Novell's Web Server for users of its NetWare operating system and IBM's family of Lotus Domino servers, primarily for IBM's OS/390 and AS/400 customers. The most common use of web servers is to host websites, but there are other uses such as gaming, data storage, running enterprise applications, handling email, FTP, or other web uses.

The primary function of a web server is to store, process and deliver web pages to clients. The communication between client and server takes place using the Hypertext Transfer Protocol (HTTP). Pages delivered are most frequently HTML documents, which may include images, style sheets and scripts in addition to text content.

A user agent, commonly a web browser or web crawler, initiates communication by making a request for a specific resource using HTTP and the server responds with the content of that resource or an error message if unable to do so. The resource is typically a real file on the server's secondary storage, but this is not necessarily the case and depends on how the web server is implemented.

While the primary function is to serve content, a full implementation of HTTP also includes ways of receiving content from clients. This feature is used for submitting web forms, including uploading of files.

Many generic web servers also support server-side scripting using Active Server Pages (ASP), PHP, or other scripting languages. This means that the behavior of the web server can be scripted in separate files, while the actual server software remains unchanged. Usually, this function is used to generate HTML documents dynamically ("on-the-fly") as opposed to returning static documents. The former is primarily used for retrieving and/or modifying information from databases. The latter is typically much faster and more easily cached but cannot deliver dynamic content.

Web servers are not always used for serving the World Wide Web. They can also be found embedded in devices such as printers, routers, webcams and serving only a local network. The web server may then be used as a part of a system for monitoring and/or administering the device in question. This usually means that no additional software has to be installed on the client computer, since only a web browser is required (which now is included with most operating systems).

ACTIVITIES

1. List down the threats to information systems and explain how they are existing?
2. How will you create a control environment for information systems?
3. Explain the various security hazards faced by an information system.

10.2Concepts of Data ware housing

Definition of Data Warehouse

Data warehousing is a collection of methods, techniques, and tools used to support knowledge workers—senior managers, directors, managers, and analysts—to conduct data analyses that help with performing decision-making processes and improving information resources.

A data warehouse is a collection of data that supports decision-making processes. It provides the following features :

- It is subject-oriented.
- It is integrated and consistent.
- It shows its evolution over time and it is not volatile.

Data warehouses are subject-oriented because they hinge on enterprise-specific concepts, such as customers, products, sales, and orders. On the contrary, operational databases hinge on many different enterprise-specific applications. The data warehouses take advantage of multiple data sources, such as data extracted from production and then stored to enterprise databases, or even data from a third party's information systems. A data warehouse should provide a unified view of all the data. Generally speaking, we can state that creating a data warehouse system does not require that new information be added; rather, existing information needs rearranging. This implicitly means that an information system should be previously available.

Operational data usually covers a short period of time, because most transactions involve the latest data. A data warehouse should enable analyses that instead cover a few years. For this reason, data warehouses are regularly updated from operational data and keep on growing. If data were visually represented, it might progress like so: A photograph of operational data would be made at regular intervals. The sequence of photographs would be stored to a data warehouse, and results would be shown in a movie that reveals the status of an enterprise from its foundation until present.

Fundamentally, data is never deleted from data warehouses and updates are normally carried out when data warehouses are offline. This means that data warehouses can be essentially viewed as read-only databases. This satisfies the users' need for a short analysis query response time and has other important effects. First, it affects data warehouse–specific database management system (DBMS) technologies, because there is no need for advanced transaction management techniques required by operational applications. Second, data warehouses operate in read-only mode, so data warehouse–specific logical design solutions are completely different from those used for operational databases. For instance, the most obvious feature of data warehouse relational implementations is that table normalization can be given up to partially de normalize tables and improve performance.

Other differences between operational databases and data warehouses are connected with query types. Operational queries execute transactions that generally read/write a small number of tuples from/to many tables connected by simple relations. For example, this applies if you search for the data of a customer in

order to insert a new customer order. This kind of query is an OLTP query. On the contrary, the type of query required in data warehouses is OLAP. It features dynamic, multidimensional analyses that need to scan a huge amount of records to process a set of numeric data summing up the performance of an enterprise. It is important to note that OLTP systems have an essential workload core “frozen” in application programs, and ad hoc data queries are occasionally run for data maintenance. Conversely, data warehouse interactivity is an essential property for analysis sessions, so the actual workload constantly changes as time goes by.

The distinctive features of OLAP queries suggest adoption of a multidimensional representation for data warehouse data. Basically, data is viewed as points in space, whose dimensions correspond to many possible analysis dimensions. Each point represents an event that occurs in an enterprise and is described by a set of measures relevant to decision-making processes.

Data Warehouse Architectures

The following architecture properties are essential for a data warehouse system :

• **Separation** Analytical and transactional processing should be kept apart as much as possible.

• **Scalability** Hardware and software architectures should be easy to upgrade as the data volume, which has to be managed and processed, and the number of users’ requirements, which have to be met, progressively increase.

• **Extensibility** The architecture should be able to host new applications and technologies without redesigning the whole system.

• **Security** Monitoring accesses is essential because of the strategic data stored in data warehouses.

• **Administerability** Data warehouse management should not be overly difficult.

1.Single-Layer Architecture

A single-layer architecture is not frequently used in practice. Its goal is to minimize the amount of data stored; to reach this goal, it removes data redundancies. Figure shows the only layer physically available: the source layer. In this case, data warehouses are virtual.

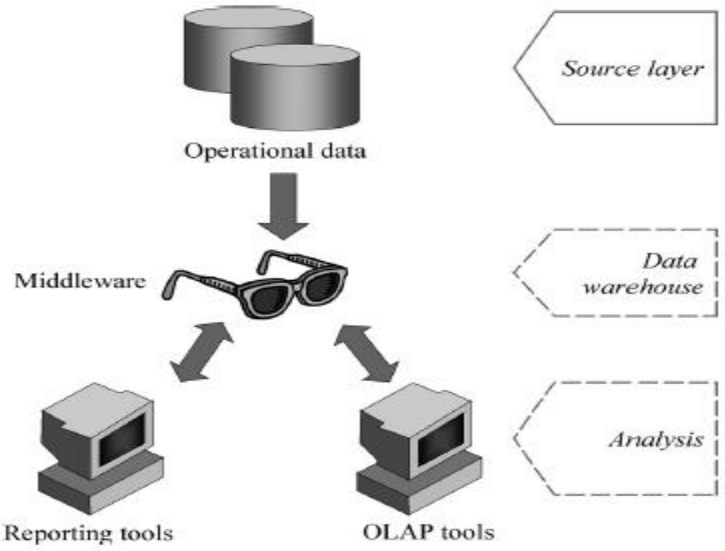


Figure: Single-layer architecture for a data warehouse system

This means that a data warehouse is implemented as a multidimensional view of operational data created by specific middleware, or an intermediate processing layer

The weakness of this architecture lies in its failure to meet the requirement for separation between analytical and transactional processing. Analysis queries are submitted to operational data after the middleware interprets them. It this way, the queries affect regular transactional workloads. In addition, although this architecture can meet the requirement for integration and correctness of data, it cannot log more data than sources do. For these reasons, a virtual approach to data warehouses can be successful only if analysis needs are particularly restricted and the data volume to analyze is huge.

2. **Two-Layer Architecture**

The requirement for separation plays a fundamental role in defining the typical architecture for a data warehouse system, as shown in Figure. Although it is typically called a two-layer architecture to highlight a separation between physically available sources and data warehouses, it actually consists of four subsequent data flow stages:

1. **Source layer** A data warehouse system uses heterogeneous sources of data. That data is originally stored to corporate relational databases or legacy1 databases, or it may come from information systems outside the corporate walls.
2. **Data staging** The data stored to sources should be extracted, cleansed to remove inconsistencies and fill gaps, and integrated to merge heterogeneous sources into one common schema. The so-called Extraction, Transformation, and Loading tools (ETL) can merge heterogeneous schemata, extract, transform, cleanse, validate, filter, and load source data into a data warehouse. Technologically speaking, this stage deals with problems that are typical for distributed information systems, such as inconsistent data management and incompatible data structures.

3. **Data warehouse layer** Information is stored to one logically centralized single repository: a data warehouse. The data warehouse can be directly accessed, but it can also be used as a source for creating data marts, which partially replicate data warehouse contents and are designed for specific enterprise departments. Meta-data repositories store information on sources, access procedures, data staging, users, data mart schemata, and so on.
4. **Analysis** In this layer, integrated data is efficiently and flexibly accessed to issue reports, dynamically analyze information, and simulate hypothetical business scenarios. Technologically speaking, it should feature aggregate data navigators, complex query optimizers, and user-friendly GUIs.

The architectural difference between data warehouses and data marts needs to be studied closer. The component marked as a data warehouse in Figure is also often called the primary data warehouse or corporate data warehouse. It acts as a centralized storage system for all the data being summed up. Data marts can be viewed as small, local data warehouses replicating (and summing up as much as possible) the part of a primary data warehouse required for a specific application domain.

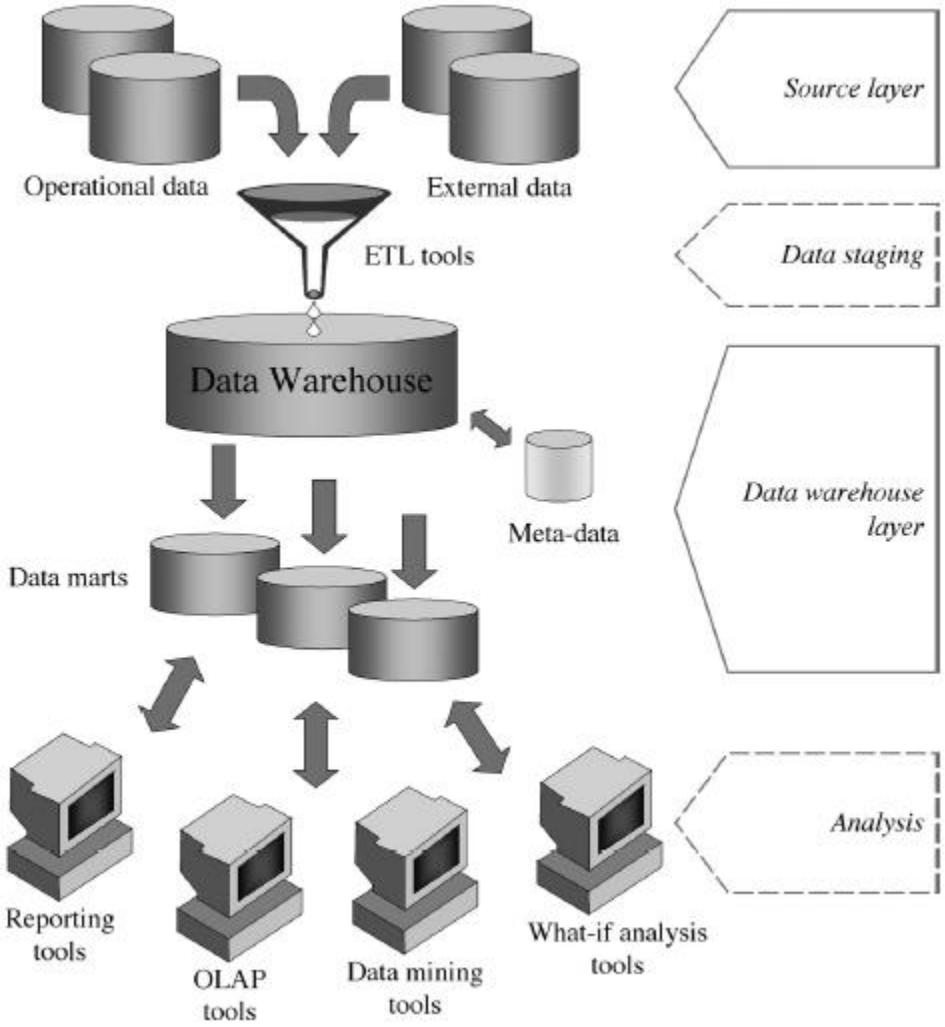


Figure: Two-layer architecture for a data warehouse system

Data Marts

A data mart is a subset or an aggregation of the data stored to a primary data warehouse. It includes a set of information pieces relevant to a specific business area, corporate department, or category of users.

The data marts populated from a primary data warehouse are often called dependent. Although data marts are not strictly necessary, they are very useful for data warehouse systems in midsize to large enterprises because

- they are used as building blocks while incrementally developing data warehouses;
- they mark out the information required by a specific group of users to solve queries;
- they can deliver better performance because they are smaller than primary data warehouses.

Sometimes, mainly for organization and policy purposes, you should use a different architecture in which sources are used to directly populate data marts. These data marts are called independent. If there is no primary data warehouse, this streamlines the design process, but it leads to the risk of inconsistencies between data marts. To avoid these problems, you can create a primary data warehouse and still have independent data marts. In comparison with the standard two-layer architecture of Figure, the roles of data marts and data warehouses are actually inverted. In this case, the data warehouse is populated from its data marts, and it can be directly queried to make access patterns as easy as possible.

The following list sums up all the benefits of a two-layer architecture, in which a data warehouse separates sources from analysis applications :

- In data warehouse systems, good quality information is always available, even when access to sources is denied temporarily for technical or organizational reasons.
- Data warehouse analysis queries do not affect the management of transactions, the reliability of which is vital for enterprises to work properly at an operational level.
- Data warehouses are logically structured according to the multidimensional model, while operational sources are generally based on relational or semi-structured models.
- A mismatch in terms of time and granularity occurs between OLTP systems, which manage current data at a maximum level of detail, and OLAP systems, which manage historical and summarized data.
- Data warehouses can use specific design solutions aimed at performance optimization of analysis and report applications.

3. **Three-Layer Architecture**

In this architecture, the third layer is the reconciled data layer or operational data store. This layer materializes operational data obtained after integrating and cleansing source data. As a result, those data are

integrated, consistent, correct, current, and detailed. Figure shows a data warehouse that is not populated from its sources directly, but from reconciled data.

The main advantage of the reconciled data layer is that it creates a common reference data model for a whole enterprise. At the same time, it sharply separates the problems of source data extraction and integration from those of data warehouse population. Remarkably, in some cases, the reconciled layer is also directly used to better accomplish some operational tasks, such as producing daily reports that cannot be satisfactorily prepared using the corporate applications, or generating data flows to feed external processes periodically so as to benefit from cleaning and integration. However, reconciled data leads to more redundancy of operational source data. Note that we may assume that even two-layer architectures can have a reconciled layer that is not specifically materialized, but only virtual, because it is defined as a consistent integrated view of operational source data.

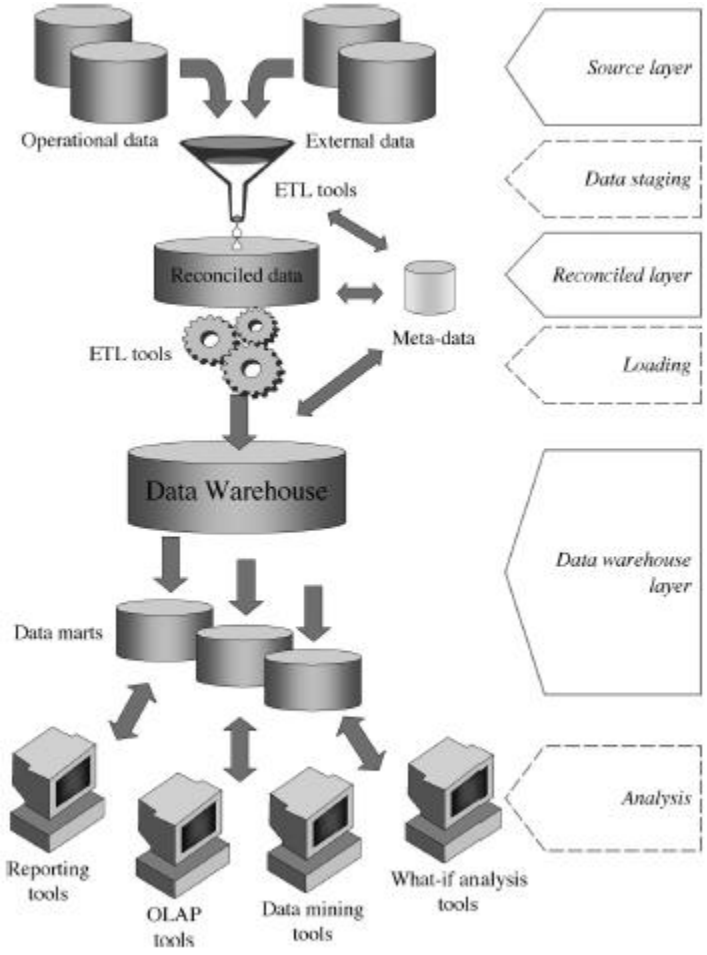


Figure: Three-layer architecture for a data warehouse system

Finally, let's consider a supplementary architectural approach, which provides a comprehensive picture. This approach can be described as a hybrid solution between the single-layer architecture and the two/three-layer architecture. This approach assumes that although a data warehouse is available, it is unable to solve all the

queries formulated. This means that users may be interested in directly accessing source data from aggregate data (drill-through). To reach this goal, some queries have to be rewritten on the basis of source data (or reconciled data if it is available). This type of architecture is implemented in a prototype and it needs to be able to go dynamically back to the source data required for queries to be solved (lineage).

4.**An Additional Architecture Classification**

The scientific literature often distinguishes five types of architecture for data warehouse systems, in which the same basic layers mentioned in the preceding paragraphs are combined in different ways.

In independent data marts architecture, different data marts are separately designed and built in a nonintegrated fashion. This architecture can be initially adopted in the absence of a strong sponsorship toward an enterprise-wide warehousing project, or when the organizational divisions that make up the company are loosely coupled. However, it tends to be soon replaced by other architectures that better achieve data integration and cross-reporting.

The bus architecture, recommended by Ralph Kimball, is apparently similar to the preceding architecture, with one important difference. A basic set of conformed dimensions (that is, analysis dimensions that preserve the same meaning throughout all the facts they belong to), derived by a careful analysis of the main enterprise processes, is adopted and shared as a common design guideline. This ensures logical integration of data marts and an enterprise-wide view of information. In the hub-and-spoke architecture, one of the most used in medium to large contexts, there is much attention to scalability and extensibility, and to achieving an enterprise-wide view of information.

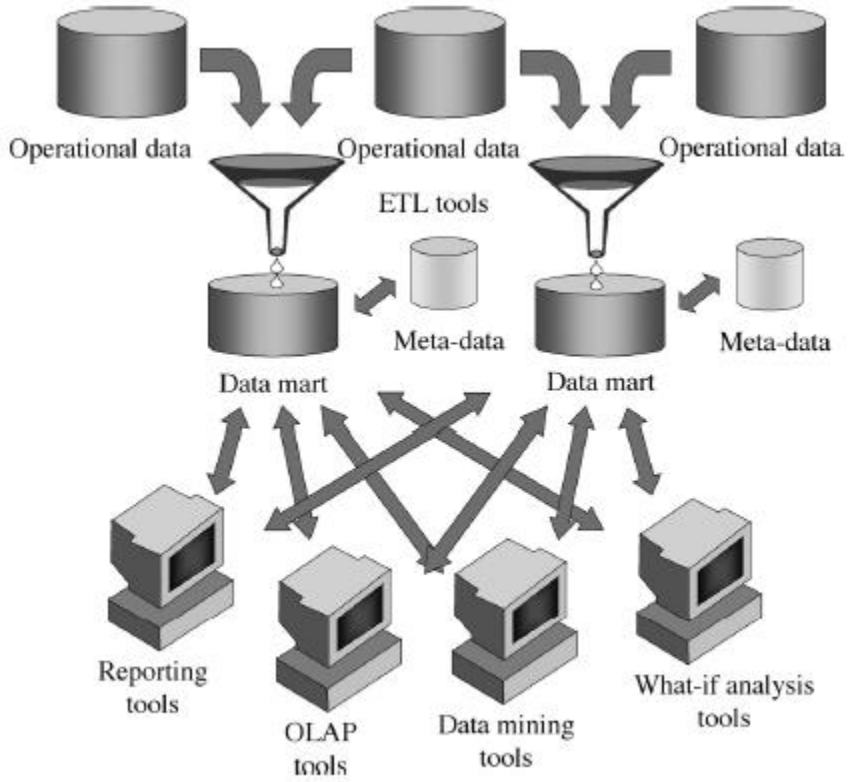


Figure: Independent data marts architecture

Atomic, normalized data is stored in a reconciled layer that feeds a set of data marts containing summarized data in multidimensional form (Figure , Hub-and-spoke architecture). Users mainly access the data marts, but they may occasionally query the reconciled layer.

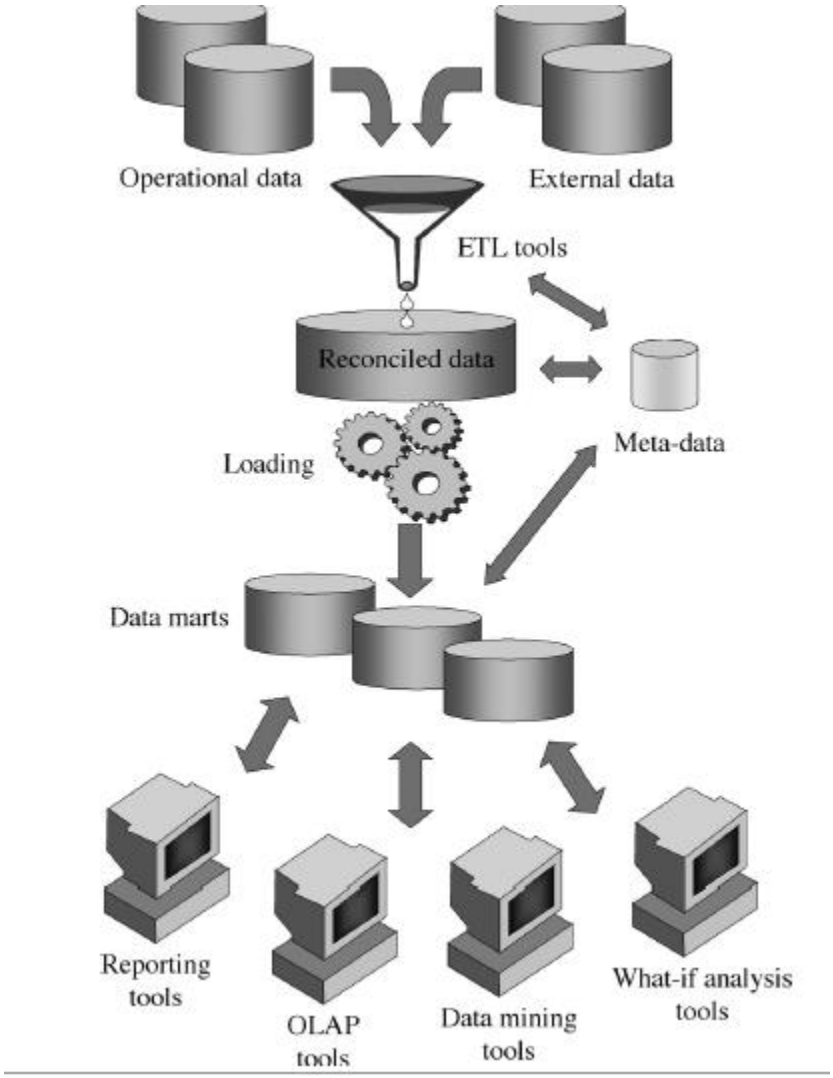


Figure: Hub-and-spoke architecture

The centralized architecture, recommended by Bill Inmon, can be seen as a particular implementation of the hub-and-spoke architecture, where the reconciled layer and the data marts are collapsed into a single physical repository.

The federated architecture is sometimes adopted in dynamic contexts where preexisting data warehouses/data marts are to be noninvasively integrated to provide a single, cross-organization decision support environment (for instance, in the case of mergers and acquisitions). Each data warehouse/data mart is either virtually or

physically integrated with the others, leaning on a variety of advanced techniques such as distributed querying, on topologies, and meta-data interoperability (Figure, Federated architecture).

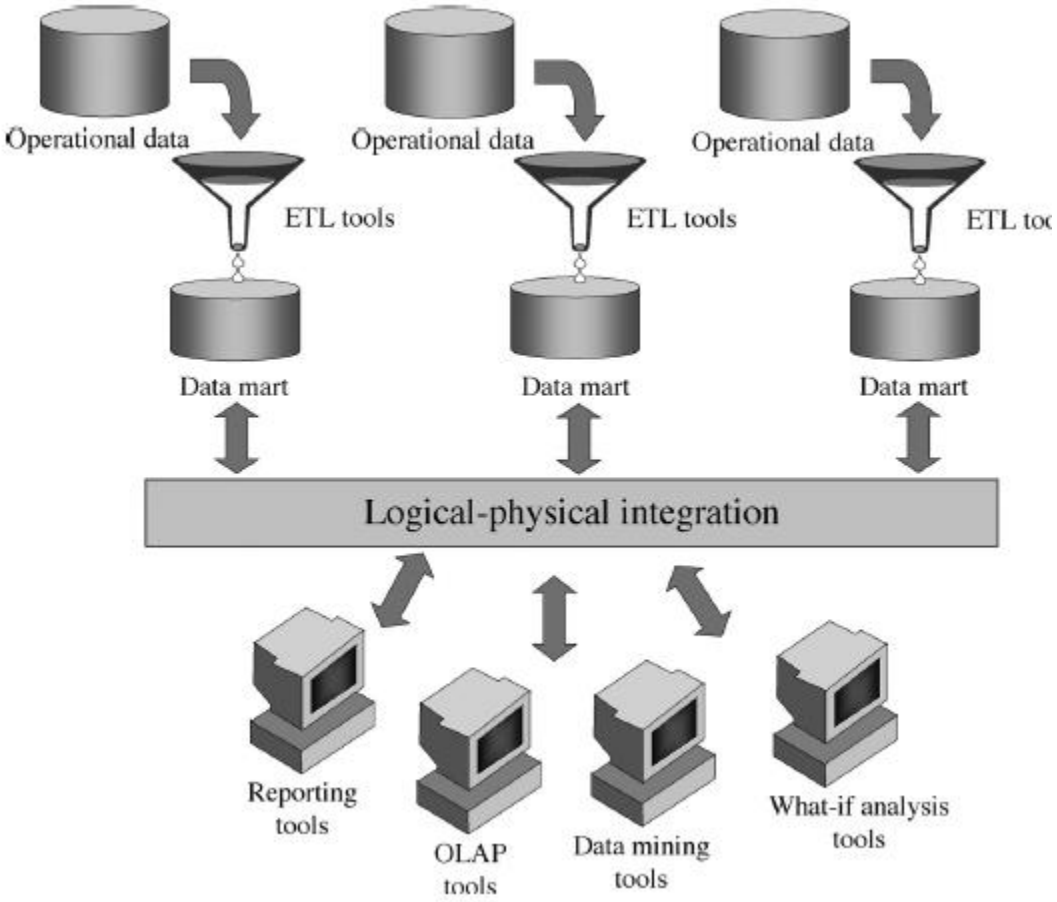


Figure: Federated architecture

The following list includes the factors that are particularly influential when it comes to choosing one of these architectures:

• The amount of interdependent information exchanged between organizational units in an enterprise and the organizational role played by the data warehouse project sponsor may lead to the implementation of enterprise-wide architectures, such as bus architectures, or department-specific architectures, such as independent data marts.

• An urgent need for a data warehouse project, restrictions on economic and human resources, as well as poor IT staff skills may suggest that a type of “quick” architecture, such as independent data marts, should be implemented.

• The minor role played by a data warehouse project in enterprise strategies can make you prefer an architecture type based on independent data marts over a hub-and-spoke architecture type.

• The frequent need for integrating preexisting data warehouses, possibly deployed on heterogeneous platforms, and the pressing demand for uniformly accessing their data can require a federated architecture type.

Data Warehouse Tools and Utilities Functions

The following are the functions of Data Warehouse tools and Utilities:

Data Extraction - Data Extraction involves gathering the data from multiple heterogeneous sources.

Data Cleaning - Data Cleaning involves finding and correcting the errors in data.

Data Transformation - Data Transformation involves converting data from legacy format to warehouse format.

Data Loading - Data Loading involves sorting, summarizing, consolidating, checking integrity and building indices and partitions.

Refreshing - Refreshing involves updating from data sources to warehouse.

*****Data Cleaning and Data Transformation are important steps in improving the quality of data and data mining results.

SELF ASSEMENT QUESTIONS

Fill in the blanks:

1.____ layer decides which physical pathway the data should take.

a. Application

b. Network

c. Physical

2. ISDN is an example of _____ network

a. Circuit switched

b. Packet switched

3. X.25 is an example of _____ network

a. Circuit switched

b. Packet switched

4. _____ allows LAN users to share computer programs and data.

a. Communication server

b. Print server

c. File server

5. Print server uses _____ which is a buffer that holds data before it is send to the printer.

a. Queue

b. Spool

c. Node

6. A standalone program that has been modified to work on a LAN by including concurrency controls such as file and record locking is an example of____

a. LAN intrinsic software

b. LAN aware software

c. Groupware

d. LAN ignorant software

10.3 Data Mining

Data mining (the analysis step of the "Knowledge Discovery in Databases" process, or KDD), an interdisciplinary subfield of computer science, is the computational process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use.[2] Aside from the raw analysis step, it involves database and data management aspects, data pre-processing, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, visualization, and online updating.

The term is a misnomer, because the goal is the extraction of patterns and knowledge from large amount of data, not the extraction of data itself. It also is a buzzword, and is also frequently applied to any form of large-scale data or information processing (collection, extraction, warehousing, analysis, and statistics) as well as any application of computer decision support system, including artificial intelligence, machine learning, and business intelligence. The popular book "Data mining: Practical machine learning tools and techniques with Java" (which covers mostly machine learning material) was originally to be named just "Practical machine learning", and the term "data mining" was only added for marketing reasons. Often the more general terms "(large scale) data analysis", or "analytics" – or when referring to actual methods, artificial intelligence and machine learning – are more appropriate.

The actual data mining task is the automatic or semi-automatic analysis of large quantities of data to extract previously unknown interesting patterns such as groups of data records (cluster analysis), unusual records (anomaly detection) and dependencies (association rule mining). This usually involves using database techniques such as spatial indices. These patterns can then be seen as a kind of summary of the input data, and may be used in further analysis or, for example, in machine learning and predictive analytics. For example, the data mining step might identify multiple groups in the data, which can then be used to obtain more accurate prediction results by a decision support system. Neither the data collection, data preparation, nor result interpretation and reporting are part of the data mining step, but do belong to the overall KDD process as additional steps.

Data mining involves six common classes of tasks:

- Anomaly detection (Outlier/change/deviation detection) – The identification of unusual data records, that might be interesting or data errors that require further investigation.

- Association rule learning (Dependency modeling) – Searches for relationships between variables. For example a supermarket might gather data on customer purchasing habits. Using association rule learning, the supermarket can determine which products are frequently bought together and use this information for marketing purposes. This is sometimes referred to as market basket analysis.

- Clustering – is the task of discovering groups and structures in the data that are in some way or another "similar", without using known structures in the data.

- Classification – is the task of generalizing known structure to apply to new data. For example, an e-mail program might attempt to classify an e-mail as "legitimate" or as "spam".

- Regression – attempts to find a function which models the data with the least error.

- Summarization – providing a more compact representation of the data set, including visualization and report generation

10.4 Internet 2

Internet2 is a not-for-profit United States computer networking consortium led by members from the research and education communities, industry, and government. The Internet2 consortium administrative headquarters are located in Ann Arbor, Michigan, with offices in Washington, D.C. and Emeryville, California.

As of November 2013, Internet2 has over 500 members including 251 institutions of higher education, 9 partners and 76 members from industry, over 100 research and education networks or connector organizations, and 67 affiliate members.

Internet2 operates the Internet2 Network, an Internet Protocol network using optical fiber that delivers network services for research and education, and provides a secure network testing and research environment. In late 2007, Internet2 began operating its newest dynamic circuit network, the Internet2 DCN, an advanced technology that allows user-based allocation of data circuits over the fiber-optic network.

The Internet2 Network, through its regional network and connector members, connects over 60,000 U.S. educational, research, government and "community anchor" institutions, from primary and secondary schools to community colleges and universities, public libraries and museums to health care organizations.

The Internet2 community develops and deploys network technologies for the future of the Internet. These technologies include large-scale network performance measurement and management tools, secure identity and access management tools and capabilities such as scheduling high-bandwidth, high-performance circuits.

Internet2 members serve on several advisory councils, collaborate in a variety of working groups and special interest groups, gather at spring and fall member meetings, and are encouraged to participate in the strategic planning process.

Objective

Internet2 provides the U.S. research and education community with a network that satisfies their bandwidth-intensive requirements. The network itself is a dynamic, robust and cost-effective hybrid optical and packet network. It furnishes a 100 Gbit/s network backbone to more than 210 U.S. educational institutions, 70 corporations and 45 non-profit and government agencies.

The objectives of the Internet2 consortium are:

- Developing and maintaining a leading-edge network.
- Fully exploiting the capabilities of broadband connections through the use of new-generation applications.
- Transferring new network services and applications to all levels of educational use, and eventually the broader Internet community.

The uses of the network span from collaborative applications, distributed research experiments, grid-based data analysis to social networking. Some of these applications are in varying levels of commercialization, such as IPv6, open-source middleware for secure network access, Layer 2 VPNs and dynamic circuit networks.

10.5 Mobile communication

Communication is one of the integral parts of science that has always been a focus point for exchanging information among parties at locations physically apart. After its discovery, telephones have replaced the telegrams and letters. Similarly, the term ‘mobile’ has completely revolutionized the communication by opening up innovative applications that are limited to one’s imagination. Today, mobile communication has become the backbone of the society. All the mobile system technologies have improved the way of living. Its main plus point is that it has privileged a common mass of society.

At the initial phase, mobile communication was restricted to certain official users and the cellular concept was never even dreamt of being made commercially available. Moreover, even the growth in the cellular networks was very slow. However, with the development of newer and better technologies starting from the 1970s and with the mobile users now connected to the PSTN, there has been a remarkable growth in the cellular radio. However, the spread of mobile communication was very fast in the 1990s when the government throughout the world provided radio spectrum licenses for Personal Communication Service (PCS) in 1.8 - 2 GHz frequency band

5. 1G: First Generation Networks

The first mobile phone system in the market was AMPS. It was the first U.S. cellular telephone system, deployed in Chicago in 1983. The main technology of this first generation mobile system was FDMA/FDD and analog FM.

2. 2G: Second Generation Networks

2G (or **2-G**) is short for second-generation wireless telephone technology. Second generation 2G cellular telecom networks were commercially launched on the GSM standard in Finland by Radiolinja (now part of Elisa Oyj) in 1991. Three primary benefits of 2G networks over their predecessors were that phone conversations were digitally encrypted; 2G systems were significantly more efficient on the spectrum allowing for far greater mobile phone penetration levels; and 2G introduced data services for mobile, starting with SMS text messages. 2G technologies enabled the various mobile phone networks to provide the services such as text messages, picture messages and MMS (multi media messages). All text messages sent over 2G are digitally encrypted, allowing for the transfer of data in such a way that only the intended receiver can receive and read it. After 2G was launched, the previous mobile telephone systems were retrospectively dubbed 1G. While radio signals on 1G networks are analog, radio signals on 2G networks are digital. Both systems use digital signaling to connect the radio towers (which listen to the handsets) to the rest of the telephone system. 2G has been superseded by newer technologies such as 2.5G, 2.75G, 3G, and 4G; however, 2G networks are still used in many parts of the world.

Capacity

Using digital signals between the handsets and the towers increases system capacity in two key ways:

- Digital voice data can be compressed and multiplexed much more effectively than analog voice encodings through the use of various codes, allowing more calls to be transmitted in same amount of radio bandwidth.
- The digital systems were designed to emit less radio power from the handsets. This meant that cells had to be smaller, so more cells had to be placed in the same amount of space. This was possible because cell towers and related equipment had become less expensive.

2G Data Transmission Capacity:

- With GPRS (General Packet Radio Service), you have a theoretical transfer speed of max. 50 kbit/s (40 kbit/s in practice).
- With EDGE (Enhanced Data Rates for GSM Evolution), you have a theoretical transfer speed of max. 1 mbit/s (500 kbit/s in practice).

Disadvantages

- In less populous areas, the weaker digital signal transmitted by a cellular phone may not be sufficient to reach a cell tower. This tends to be a particular problem on 2G systems deployed on higher frequencies, but is mostly not a problem on 2G systems deployed on lower frequencies. National regulations differ greatly among countries which dictate where 2G can be deployed.

- Analog has a smooth decay curve, but digital has a jagged steppy one. This can be both an advantage and a disadvantage. Under good conditions, digital will sound better. Under slightly worse conditions, analog will experience static, while digital has occasional dropouts. As conditions worsen, though, digital will start to completely fail, by dropping calls or being unintelligible, while analog slowly gets worse, generally holding a call longer and allowing at least some of the audio transmitted to be understood.

Advantage

- While digital calls tend to be free of static and background noise, the lossy compression they use reduces their quality, meaning that the range of sound that they convey is reduced. Talking on a digital cell phone, a caller hears less of the tonality of someone's voice.

3. 3G: Third Generation Networks

3G, short form of **third generation**, is the **third generation of mobile telecommunications technology**. This is based on a set of standards used for mobile devices and mobile telecommunications use services and networks that comply with the **International Mobile Telecommunications-2000 (IMT-2000)** specifications by the International Telecommunication Union. **3G** finds application in wireless voice telephony, mobile Internet access, fixed wireless Internet access, video calls and mobile TV. 3G telecommunication networks support services that provide an information transfer rate of at least 200 kbit/s. Later 3G releases, often denoted 3.5G and 3.75G, also provide mobile broadband access of several Mbit/s to smart phones and mobile modems in laptop computers. This ensures it can be applied to wireless voice telephony, mobile Internet access, fixed wireless Internet access, video calls and mobile TV technologies. A new generation of cellular standards has appeared approximately every tenth year since 1G systems were introduced in 1981/1982. Each generation is characterized by new frequency bands, higher data rates and non–backward-compatible transmission technology. The first 3G networks were introduced in 1998 and fourth generation "4G" networks in 2008.

The bandwidth and location information available to 3G devices gives rise to applications not previously available to mobile phone users. Some of the applications are:

- Global Positioning System (GPS)
- Location-based services
- Mobile TV
- Telemedicine
- Video Conferencing
- Video on demand

4. 4G: Four Generation Networks

4G, short for **fourth generation**, is the fourth generation of mobile telecommunications technology, succeeding 3G and preceding 5G. A 4G system, in addition to the usual voice and other services of 3G, provides mobile broadband Internet access, for example to laptops with wireless modems, to smart phones, and to other mobile devices. Potential and current applications include amended mobile web access, IP telephony, gaming services, high-definition mobile TV, video conferencing, 3D television, and cloud computing.

Two 4G candidate systems are commercially deployed: the Mobile WiMAX standard (first used in South Korea in 2007), and the first-release Long Term Evolution (LTE) standard (in Oslo, Norway and Stockholm, Sweden since 2009).

10.6 Summary

Data warehouse is subject Oriented, Integrated, Time-Variant and nonvolatile collection of data that support of management's decision making process. Let's explore this Definition of data warehouse.

Subject Oriented - The Data warehouse is subject oriented because it provide us the information around a subject rather the organization's ongoing operations. These subjects can be product, customers, suppliers, sales, revenue etc. The data warehouse does not focus on the ongoing operations rather it focuses on modeling and analysis of data for decision making.

Integrated - Data Warehouse is constructed by integration of data from heterogeneous sources such as relational databases, flat files etc. This integration enhance the effective analysis of data.

Time-Variant - The Data in Data Warehouse is identified with a particular time period. The data in data warehouse provide information from historical point of view.

Non Volatile - Non volatile means that the previous data is not removed when new data is added to it. The data warehouse is kept separate from the operational database therefore frequent changes in operational database are not reflected in data warehouse.

Metadata - Metadata is simply defined as data about data. The data that are used to represent other data is known as metadata. For example the index of a book serves as metadata for the contents in the book. In other words we can say that metadata is the summarized data that lead us to the detailed data.

In terms of data warehouse we can define metadata as following:

Metadata is a road map to data warehouse. Metadata in data warehouse define the warehouse objects. The metadata act as a directory. This directory helps the decision support system to locate the contents of data warehouse.

Data mart

Data mart contains the subset of organization-wide data. This subset of data is valuable to specific group of an organization. in other words we can say that data mart contains only that data which is specific to a particular group. For example the marketing data mart may contain only data related to item, customers and sales. The data mart is confined to subjects.

Generally the data warehouses adopt the three-tier architecture. Following are the three tiers of data warehouse architecture.

Bottom Tier - The bottom tier of the architecture is the data warehouse database server.It is the relational database system.We use the back end tools and utilities to feed data into bottom tier.these back end tools and utilities performs the Extract, Clean, Load, and refresh functions.

Middle Tier - In the middle tier we have OLAp Server. the OLAP Server can be implemented in either of the following ways.

o By relational OLAP (ROLAP), which is an extended relational database management system? The ROLAP maps the operations on multidimensional data to standard relational operations.

o By Multidimensional OLAP (MOLAP) model, which directly implements multidimensional data and operations?

Top-Tier - This tier is the front-end client layer. This layer holds the query tools and reporting tool, analysis tools and data mining tools.

Warehouse Manager

Warehouse manager is responsible for the warehouse management process. The warehouse manager consists of third party system software, C programs and shell scripts. The size and complexity of warehouse manager varies between specific solutions.

A **web page** (or **webpage**) is a web document that is suitable for the World Wide Web and the *web browser*. A web browser displays a web page on a monitor or mobile device. The web page is what displays, but the term also refers to a computer file, usually written in HTML or comparable markup language. Web browsers coordinate the various web resource elements for the written web page, such as style sheets, scripts and images, to present the web page.

A *static web page* is delivered exactly as stored, as web content in the web server's file system, while a *dynamic web page* is generated by a web application that is driven by server-side software or client-side scripting. Dynamic web pages help the browser (the client) to enhance the web page through user input to the server.

A **web server** is a computer system that processes requests via HTTP, the basic network protocol used to distribute information on the World Wide Web. The term can refer either to the entire system, or specifically to the software that accepts and supervises the HTTP requests

Data warehousing is a collection of methods, techniques, and tools used to support knowledge workers—senior managers, directors, managers, and analysts—to conduct data analyses that help with performing decision-making processes and improving information resources.

A data warehouse is a collection of data that supports decision-making processes. It provides the following features :

- It is subject-oriented.
- It is integrated and consistent.
- It shows its evolution over time and it is not volatile.

Architectural properties are essential for a data warehouse system :

- **Separation** Analytical and transactional processing should be kept apart as much as possible.
- **Scalability** Hardware and software architectures should be easy to upgrade as the data volume, which has to be managed and processed, and the number of users' requirements, which have to be met, progressively increase.
- **Extensibility** The architecture should be able to host new applications and technologies without redesigning the whole system.
- **Security** Monitoring accesses is essential because of the strategic data stored in data warehouses.
- **Administer ability** Data warehouse management should not be overly difficult.

Data Marts

A data mart is a subset or an aggregation of the data stored to a primary data warehouse. It includes a set of information pieces relevant to a specific business area, corporate department, or category of users.

Data Warehouse Tools and Utilities Functions

The following are the functions of Data Warehouse tools and Utilities:

Data Extraction - Data Extraction involves gathering the data from multiple heterogeneous sources.

Data Cleaning - Data Cleaning involves finding and correcting the errors in data.

Data Transformation - Data Transformation involves converting data from legacy format to warehouse format.

Data Loading - Data Loading involves sorting, summarizing, consolidating, checking integrity and building indices and partitions.

Refreshing - Refreshing involves updating from data sources to warehouse.

Internet2 is a not-for-profit United States computer networking consortium led by members from the research and education communities, industry, and government. The Internet2 consortium administrative headquarters are located in Ann Arbor, Michigan, with offices in Washington, D.C. and Emeryville, California.

10.7 Glossary

Web page : A **web** page (also written as webpage) is a document that is suitable for the World Wide **Web** and **web** browsers. A **web** browser displays a **web** page on a monitor or mobile device.

Web server: A **web server** is a computer that runs websites. It's a computer program that distributes **web** pages as they are requisitioned. The basic objective of the **web server** is to store, process and deliver **web** pages to the users. This intercommunication is done using Hypertext Transfer Protocol

Data warehousing : In computing, a **data warehouse** (**DW** or **DWH**), also known as an **enterprise data warehouse** (**EDW**), is a system used for reporting and data analysis, and is considered a core component of business intelligence. DWs are central repositories of integrated data from one or more disparate sources. They store current and historical data in one single place that are used for creating analytical reports for workers throughout the enterprise.

Online analytical processing (OLAP) is characterized by a relatively low volume of transactions. Queries are often very complex and involve aggregations. For OLAP systems, response time is an effectiveness measure. OLAP applications are widely used by Data Mining techniques. OLAP databases store aggregated, historical data in multi-dimensional schemas (usually star schemas). OLAP systems typically have data latency of a few hours, as opposed to data marts, where latency is expected to be closer to one day. The OLAP approach is used to analyze multidimensional data from multiple sources and perspectives. The three basic operations in OLAP are : Roll-up (Consolidation), Drill-down and Slicing & Dicing.

Online transaction processing (OLTP) is characterized by a large number of short on-line transactions (INSERT, UPDATE, DELETE). OLTP systems emphasize very fast query processing and maintaining data integrity in multi-access environments. For OLTP systems, effectiveness is measured by the number of transactions per second. OLTP databases contain detailed and current data. The schema used to store transactional databases is the entity model (usually 3NF).[10] Normalization is the norm for data modeling techniques in this system.

10.8 References

O'Brien, J (1999). *Management Information Systems – Managing Information Technology in the Internetworked Enterprise*. Boston: Irwin McGraw-Hill. ISBN 0-07-112373-3.

Laudon, Kenneth C.; Laudon, Jane P. (2009). *Management Information Systems: Managing the Digital Firm* (11 ed.). Prentice Hall/CourseSmart. p. 164.

Transaction processing systems (TPS) collect and record the routine transactions of an organization. Examples of such systems are sales order entry, hotel reservations, payroll, employee record keeping, and shipping.

Bidgoli, Hossein, (2004). The Internet Encyclopedia, Volume 1, John Wiley & Sons, Inc. p. 707.

10.9 **Further Readings**

Pant, S., Hsu, C., (1995), Strategic Information Systems Planning: A Review, Information Resources Management Association International Conference, May 21–24, Atlanta.

Taylor, Victoria. "Supply Chain Management: The Next Big Thing?". *Sept. 12, 2011*. Business Week. Retrieved 5 March 2014.

Lynn, Samara. "What is CRM?". PC Mag. Retrieved 5 March 2014.

Joshi, Girdhar (2013). *Management Information Systems*. New Delhi: Oxford University Press. p. 328. ISBN 9780198080992.

Laudon, K.,&Laudon, J. (2010). Management information systems: Managing the digital firm. (11th ed.). Upper Saddle River, NJ: Pearson Prentice Hall

Reference Books:

1. Leonard Jessup, Information Systems Today.
2. James A O' Brien, PHI, MIS: Managing Information Technology in the Networked Enterprise.
s3. J. A. Brady, EF Mark, Advanced cases MIS.
4. K. C. Lauden, J. P. Lauden, MIS: Managing the Digital Firm.
5. Curtis Frye, Microsoft Access; Plain and Simple.

10.10 Model Questions

Q:1 What do you understand by the terms web page and web servers.

Q:2 Explain in detail about data warehousing.

Q:3 explain about data marts and its different types

Q:4 What do you understand by the term Data mining and internet2.

ANSWERS TO SELF ASSESSMENT QUESTIONS

1. **c**
2. **a**
3. **b**
4. **c**
5. **b**
6. **a**