Data and Information

A/L ICT – Lesson 01

Information is defined as knowledge gained through study, communication, research, or instruction. Essentially, information is the result of analyzing and interpreting pieces of data. Whereas data is the individual figures, numbers, or graphs, information is the perception of those pieces of knowledge.

<u>Explores the basic concepts of ICT together with its role and applicability in today's knowledge based society.</u>

Competency Level 1.1: Investigates the basic building blocks of information and their characteristics.

The term data is used to refer raw facts and figures which are input into a system in order to process and create meaningful information. **Data consists of numbers, words, sounds and images which neither bear a particular meaning nor have organized in order to make decisions.** Lifecycle of data consists of data creation, management of data and removal of obsolete data.

For the purpose of creating meaningful information, data should be introduced into the system or create inside the system. Keeping data securely and making them accessible necessarily are done while management of data. Whenever data will be outdated and found to be unnecessary anymore, it's the time to remove that data from the system. This is known as the life cycle of data.

Data can be identified as the inputs into a system that intents to create meaningful information. **Information can be defined as the data organized meaningfully. Creation of information can be identified as processing of data.** For the purpose of processing there should be instructions to follow. After processing data according to the instructions given, information is produced which is the outcome the processed data. Although information is an output of a process the same information would be a data for another process. Thus, data and information can be identified distinctly through examining inputs and outputs of a particular process.
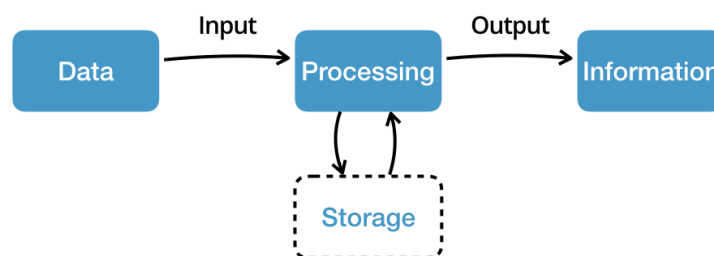
Individual pieces of data are rarely useful alone. For data to become information, data needs to be put into context. Data can be classified into two types as **quantitative and qualitative**. Quantitative data

is numerical and acquired through counting or measuring. Quantitative data can be presented with numbers and which can be arranged in a numerical order or can be subjected to arithmetic process.

Qualitative data is not measurable, but can be identified their properties. Examples: Qualitative data – **Goodwill of an organization**. Quantitative data – **Marks obtained for ICT paper**

Valuable information should have characteristics such as **relevance, timeliness, accuracy, completeness, and understandability.** Valuable information assists to make smart decisions.

- Relevance–Value of the information will depend on its relevance for a particular person.
- Timeliness–An information should be updated. Timely information is valuable for making effective decisions.
- Accuracy – Accurate information is helpful for making correct decisions.
- Completeness – Information presented with the context are useful where incomplete information is not sufficient to make decisions.
- Understandability – Information should be clear and unambiguous.

The golden rule of information graphically represents the value of information in its maximum level at the moment the information is created or the information is released. The value of the information is reduced gradually with the time and become less valuable.
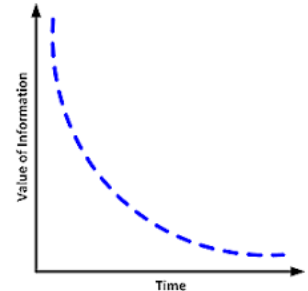
Big Data Analytics :- Organizations such as private companies and research institutions capture terabytes of data about their users' interactions in businesses and social media. The volume of data that one has to deal has been exploded to unimaginable levels with the time. The challenge of this era is to make sense of this large amount of data. This is where the big data analytics is needed to manage the data life cycle. Big Data Analytics largely involve collecting data from different sources, manage it in a way that it becomes available to be consumed by analysts and finally deliver data products to fulfil the organization's data requirements.

Competency Level 1.2:Investigates the need of technology to create, disseminate and manage data and information

Information is useful for making decisions of day to day life. To make plans, schedules, and policies information are needed. Today's information would helpful to predict how it would be in the future. For an example, present information on currency exchange would be helpful to predict that for the following month or for the following year.

Data processing and information creation can be done manually. Some drawbacks occur such as consuming more time in entering and organizing data manually. Also, human errors in data processing may lead to inaccurate information. Manual methods are inefficient in sharing information and hard to provide a quality service to the customer. Data inconsistency may occur in manual methods of manipulating data and information. Data inconsistency is, occurring the same data in two different files in different formats or values. When same data exist in different files, updates cannot be done easily.

Automatic data processing in contrast to manual data processing, utilizes technology to store, process and retrieve data. Automatic data processing handles data more efficiently than manual data processing and requires considerably less human interaction than in manual data processing. Some everyday applications in which automatic data processing outperforms manual data processing are emergency broadcast signals, security updates and weather advisories. In situations such as bad weather, harmful radiation and in some scientific or space research and also in war areas in which manual data processing is not possible, automatic methods save human lives.

With the dawn of Information and Communication Era, it has turned the world into a global village. Endless connectivity, interactive technologies, information sharing and infinite access of data by means of networking eliminate drawbacks of manual data manipulation and make available a fast, easy, convenient and accurate information creation and dissemination. ICT has emerged as a key tool for influencing the process of organizations and people around the world catching up with the technological advancement such as the Internet, the WWW, mobile communication, mobile computing and cloud computing discussed below.

The Internet

The Internet is defined as a network of networks in which users at any one computer can, if they have permission, get information from any other. The U.S. Department of Defense laid the foundation of the Internet with a network called ARPANET. However, the general public did not use the Internet much until after the development of the World Wide Web in the early 1990s.

In 1957, the U.S. government formed the Advanced Research Projects Agency (ARPA), a segment of the Department of Defense charged with ensuring U.S. leadership in science and technology with military applications. In 1969, ARPA established ARPANET, the forerunner of the Internet. ARPANET was a network that connected major computers at the Universities and other educational and research institutions later.

In response to the threat of nuclear attack, ARPANET was designed to allow continued communication if one or more sites were destroyed. In the beginning ARPANET served only computer professionals, engineers, and scientists who knew their way around its complex workings. Today, the Internet is a public, cooperative, and self-sustaining facility accessible to millions of people worldwide. Physically, the Internet uses a portion of the total resources of the currently existing public telecommunication networks. The Internet uses a set of protocols called TCP/IP (Transmission Control Protocol/Internet Protocol). **A protocol is an agreement or law behind the data transmission over networks**. TCP is a reliable transmission protocol.

Services provided by the Internet are World Wide Web (WWW), File Transfer (FTP), email, Video conferencing, Telnet, File sharing, Internet Relay chat, IP Telephone and IPTV. Electronic mail (e-mail) has replaced the postal letters. Live conversations can be carried out with other computer users, using the Internet Relay Chat (IRC) and more recently, the Internet telephony hardware and software allows real-time voice conversations. Real time voice and video can be transmitted over network with the use of video conferencing technologies. WWW allows users to view web pages distributed over the Internet and reach to the information they wish to find from anywhere anytime. File transfer allows users to transfer files over the Internet and Telnet facilitates remote login.

World Wide Web

The World Wide Web came into being in 1991. **Tim Berners-Lee is the developer of the WWW** and at the moment who directs the World Wide Web Consortium (W3C), a group of industry and university representatives that oversees the standards of Web technology.

HTTP is the set of rules for exchanging files (text, graphic images, sound, video, and other multimedia files) on the World Wide Web. Hypertext is the documents that make content on the web and it is possible to connect hypertext documents by means of hyperlinks. To view files on the Web, it is needed a web browsing software. This software is used to view different documents on the Web, which are known as Web pages. A collection of one or more web pages is a Web site. The first page of a Web site is often called as the home page. Each web site in the www has a unique address which is called the Uniform Resource Locator (URL).

Mobile communication and mobile computing

Mobile communication involves the transmission of information over a distance without the help of wires, cables or any other forms of electrical conductors. Initially the mobile communication was limited between one pair of users on single channel pair. To accommodate multiple users Time Division Multiple Access (TDMA), Code Division Multiple Access (CDMA), Frequency Division Multiple Access (FDMA) and their hybrids are used.

Mobile Computing is a technology that allows transmission of data, voice and video via a computer or any other wireless enabled device without having to be connected to a fixed physical link

Cloud Computing

The term Cloud refers to a Network or Internet. A Cloud is present at a remote location and provides services over a network. Applications such as e-mail, web conferencing, customer relationship management (CRM) run in clouds. Cloud Computing refers to manipulating, configuring, and accessing the applications online. It offers online data storage, infrastructure and application. Cloud Computing provides access to applications as utilities, over the Internet.

There are reference models on which the Cloud Computing is based. These can be categorized into three basic service models as listed below:

1. **Infrastructure as a Service** (IaaS) - provides access to fundamental resources such as physical machines, virtual machines, virtual storage, etc. Ex - Customer Relationship Management, games, virtual desktop applications

2. **Platform as a Service** (PaaS) - provides the runtime environment for applications, development & deployment tools, etc. Ex - Database, web server, deployment tools

3. **Software as a Service** (SaaS) - allows to use software applications as a service to end users. Ex - Virtual machines, servers, storage, networks
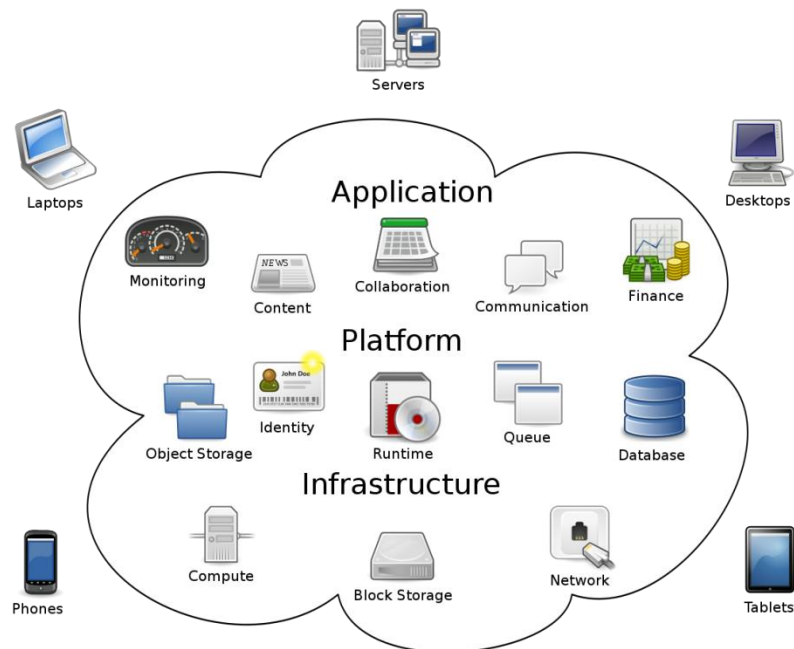
Competency Level 1.3: Formulates an Abstract model of information creation and evaluates its compliance with ICT

A system can be defined as an interrelated components work together to achieve a common objective. A system is made up of inputs, processing and output. A school, a vehicle or a bicycle can be introduced as a system. A system consists of interrelated components / sub systems and system should be analyzed to identify inputs, outputs, subsystems and their individual responsibilities. In an information system data input into the system is processed using processing instructions given and information is produced as output. **This is called an abstract model of information.**

1. A digital computer carries out the following five functions:
2. Takes data as input.
3. Stores the data/instructions in its memory and uses them as required.
4. Processes the data and converts it into useful information.
5. Generates the output.
6. Controls all the above four steps.

Though a computer system follows data input, processing data into information and output results, the basic functions of a computer system are well overlapped on the Abstract Model of Information.

Competency Level 1.4: Selects and classifies the basic components of a computer system.

Computer System consists of four major components called **Hardware, Software, Firmware and Live ware**.

Hardware components can be categorised into Input Devices, Output Devices, Processing Devices, Networking Devices and Memory Devices. In addition to those types of hardware there is a circuit board located inside a computer system called mother board by means of which every hardware component are connected to each other.

Input Devices correspond to input data into the computer system. Mouse, Trackball, Joystick, Touch pad, Touch screen and Light pen can be considered as pointing devices which are related to point out and select pixels on the screen. Keyboard is the most common and very popular input device which is used to input data to the computer. The layout of the keyboard is similar to that of traditional

typewriter, although there are some additional keys provided for performing additional functions. Scanner is an input device that captures images from the source which are then converted into a digital form that can be stored on the disk. These images can be edited before they are printed. Microphone is an input device to input sound that is then stored in a digital form. In addition to the devices mentioned above MICR, OCR, Barcode Reader also can be considered as input devices.

Output devices are the devices support to present information out from the computer system. Monitors, commonly called as Visual Display Unit (VDU), are the main output device of a computer. It forms images from tiny dots, called pixels that are arranged in a rectangular form. The resolution of the image depends upon the number of pixels.

There are two kinds of viewing screens used for monitors named as **Cathode-Ray Tube (CRT) and Flat-Panel** Display. The flat-panel display is divided into two categories as follows.

Light Emitting Diode (LED) – These are emissive displays that convert electrical energy into light.

Liquid Crystal Display (LCD)–These are non-emissive displays which use optical effects to convert light into graphics patterns.

Printer is an output device, which is used to print information on paper which is called a hardcopy. There are two types of printers,

Impact Printers - print the characters by striking them on the ribbon, which is then pressed on the paper. Dot Matrix Printer is an example for impact printers.

Non-impact printers do not physically strike the page. These printers print a complete page at a time, thus they are also called as Page Printers. Laser printers and Inkjet printers are laid on this type of printers.

Memory Devices

A memory is used to store data and instructions. Computer memory is the storage space in the computer system, where data is to be processed and instructions required for processing are stored. The memory is divided into large number of small parts called cells. Each location or cell has a unique address.

Memory is classified into three types

1. Cache Memory - Cache memory is a very high speed semiconductor memory which can speed up the CPU. It acts as a buffer between the CPU and the main memory. It is used to hold those parts of data and program which are most frequently used by the CPU.

2. Primary Memory/Main Memory - Primary memory holds only those data and instructions on which the computer is currently working. It has a limited capacity and data is lost when power is switched off. It is generally made up of semiconductor device.

3. Secondary Memory - This type of memory is also known as external memory or non-volatile. It is slower than the main memory. These are used for storing data/information permanently. CPU does not access these memories directly instead they are accessed via input-output routines. The contents of secondary memories are first transferred to the main memory, and then the CPU can access it.

The Memory Hierarchy
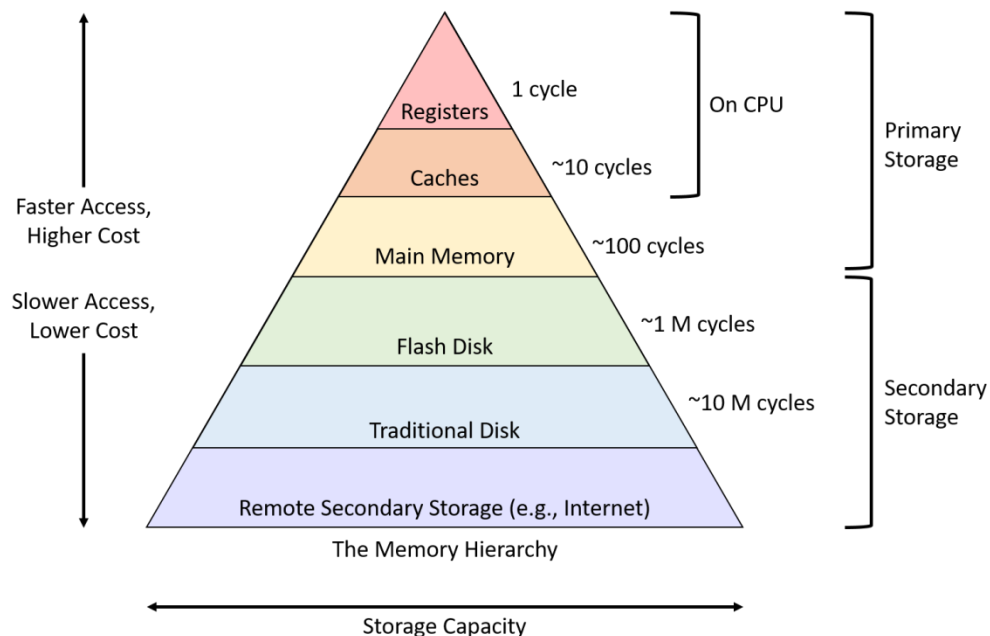
Processing Devices

Central Processing Unit (CPU) is the most common processing device of a computer which has the control of all other components. Arithmetic and logical operations take place inside the CPU. Graphics Processing Unit (GPU) is a specialized electronic circuit designed to rapidly manipulate and alter memory to accelerate the creation of images in a frame buffer intended for output to a display. GPUs

are used in embedded systems, mobile phones, personal computers, workstations, and game consoles.

Networking Devices

Computer networking devices are physical devices which are required for communication and interaction between devices on a computer network. Specifically, they mediate data in a computer network. Ex - Hubs, Switches, Bridges, Routes, Gateways, Network Interface Cards (NICs).

Computer software is classified as system software and application software. System software (systems software) is the computer software designed to operate and control the computer hardware and to provide a platform for running application software. System software can be separated into three different categories as follows.

1. Operating Systems –Windows, Linux, Mac OS

2. Utility Software– Virus Guard, Device drivers

3. Language Translators – Compiler, Interpreter

Application software products are designed to satisfy a particular need of a particular environment. Application software may consist of a single program, such as Microsoft's notepad for writing and editing a simple text. It may also consist of a collection of programs, often called a software package, which work together to accomplish a task.

Software can also be classified as **'open source software' and 'proprietary software'.**

- ✓ Open Source Software –source code is available and it is free to use, modify or redistribute these type of software.
- ✓ Proprietary Software – source code is not available and user should have the permission given by owner of the software to use this type of software

Firmware is a computer program that is "embedded" in a hardware device, that is, an essential part of the hardware. It is sometimes called embedded software. In computers firmware embedded in ROM and which handles booting up process of computers.

Liveware refers computer users. According to user commands generally computers are functioning.

Hardware, System Software, Application Software
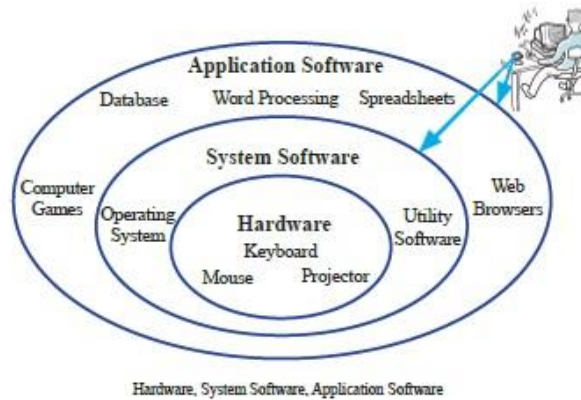
Competency Level 1.5: Analyses the activities of data processing

Data gathering

Data gathering methods vary as manual methods, Semi-automated and automated methods. Tools used in automated methods. Interviews, questionnaires and observation can be done in order to collect data manually. Using interviews detailed questions and answers may helpful for making the discussion fair enough to capture deserved data. It's a fast method to use of questionnaires which is most suitable for gathering quantitative data and answering anonymously, may helpful to gather fair and confident data.

In semi-automatic and automatic methods OMR, OCR, MICR, card/tape readers, magnetic strip readers, bar code readers, sensors and loggers are popular.

An Optical Mark Reader (OMR) is a scanning device that reads carefully placed pencil marks on a specially designed form or document. Most popular use of OMR is marking of MCQ answer sheets.

Optical Character Recognition (OCR) is done by Optical Character Reader which allows us to identify written or printed characters and extract editable text from it.

Magnetic Ink Character Recognition (MICR) is done by Magnetic Ink Character Reader which supports to input data written in special magnetic ink under cheque.
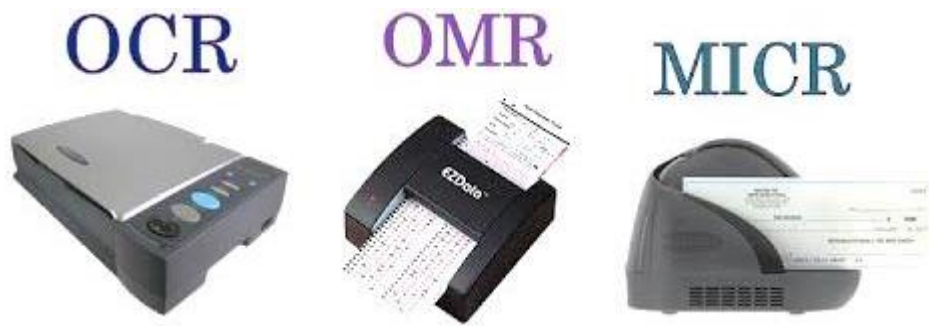
Magnetic strip readers use to read data stored in a magnetic strip attached on credit cards and bank ATM cards.

Bar code readers use visible light to scan and read barcodes which contain data to be input into computer systems.

Sensors may helpful for collecting data from different places in to a single data center fast. Sensor networks are consisting of variety of electronic sensors. Each node in a wireless sensor network is equipped with one or more sensors in addition to a microcontroller, wireless transceiver, and energy source.

Data validation methods

   ✓ Data type check – Check whether correct data type (numeric/text/date/currency etc) is input into the system
   ✓ Presence check – Presence of data (i.e. data entry field is not blank) is checked. It is compulsory to input important data
   ✓ Range check – Check whether the data is in allowed range (i.e. Marks have to be between 0 and 100) commonly used when working with data which consists of numbers, currency or dates/times

Modes of data input

- ✓ Direct and remote - In direct data input, data is input into the system directly without direct user interference. In remote entry, system is designed specially to store data. Data storing and transactions are taking place in different geographical places.
- ✓ Online and offline – In online data input, the data is input at the same time transaction taking place, where data of different transactions input after a specific time period as a batch in offline method.

Data processing

Data input and processing occur as a batch in **batch processing**. It is easy to handle huge amount of data as a batch. So, data stored until the system comes online to process the data in one 'batch'. Batch processing is done in billing systems and pay roll system.

In **real time processing** of data, input, processing and output as well as data storing occur simultaneously. Any type of monitoring system, any type of booking system, computer controlled systems such as automatic production line, an aircraft auto-pilot, the system controlling a nuclear power stations uses real time processing.

Output methods - Direct presentation to the user can be done by using a monitor, multimedia projector or a printer. Soft copies are presented with monitors and multimedia projectors whereas printer output printed copies which is said to be hard copies. Storing methods are used to store data securely for further processing needs.

Storage methods - Local storages such as Hard Disks Drives, Compact Disk, Digital Versatile Disks and Tape Drives can be used to store data inside the local computer. Remote storage such as cloud space can be used to store data securely. To get remote data when needed it is required the access to the internet

Competency Level 1.6: Investigates the application of ICT in different domains.

Education

The computer helps in providing a lot of facilities in the education system. The computer provides a tool in the education system known as CBE (Computer Based Education). CBE involves control, delivery, and evaluation of learning as follows.

1) Simulations: Computer simulators and prototypes can be used to perform the rather impossible experiments on the subjects which seemed impossible to be done in the labs especially in the field of medicine and engineering.

2) Distance Learning and E- Learning: This is the trend in 21st century education where the obstacles arise due to distance are being eliminated. One can attend classes and access learning materials remotely, without having to travel to the learning center.

3) Better Content Delivery: The motivation effect that comes about due to embracing ICT in teaching and learning is a common factor. For instance, multimedia and interactive content on interactive boards remain quite engaging and motivating making students pay more attention in class.

4) Individualized Learning: ICT help students with individual needs to have better suited assignments and making it easy to organize their learning.

Healthcare Computers have become an important part in hospitals, labs, and dispensaries. They are being used in hospitals to keep the record of patients and medicines. It is also used in scanning and diagnosing different diseases. ECG, EEG, ultrasounds and CT scans are also done by computerized machines.

Following are some major fields of health care in which computers are used.

- ✓ Diagnostic System - Computers are used to collect data and identify the cause of illness.
- ✓ Lab-diagnostic System - All tests can be done and the reports are prepared by computer.
- ✓ Patient Monitoring System - These are used to check the patient's signs for abnormality (ECG)

- ✓ Pharmacy Information System - Computer is used to check drug labels, expiry dates, harmful side effects, etc.
- ✓ Surgery - Nowadays, computers are also used as a supporting tool for surgery

Agriculture

ICT is applied in the field of agriculture, animal farm houses and fisheries industries. ICT is used to control light, moisture and air inside the greenhouses. Robotic machines are used to monitor plant growth levels, record them and harvesting in large scale farm lands. Crop harvesting and seedling planter robotics are also useful to keep efficiency of agriculture industry. Automatic weed removers and insect control devices protect the crop from weed plants and insects respectively. Monitoring devices are used to monitor factors affect crop production such as weather, climatic changes, rainfall, wind direction, soil fertility and water content.

Radio Frequency Identification Device (RFID) helps in identifying and counting the number of animals in farmhouses. Automated milking and examination of cows can be done with computer controlled systems which is helpful in monitoring the health status of the cows and the quality of milk.

Business and finance

A computer has high speed of calculation, diligence, accuracy, reliability, or versatility which has made it an integrated part in all business organizations.

Computer is used in business organizations for payroll calculations, budgeting, sales analysis, financial forecasting, managing employee database and maintenance of stocks etc.

Today, banking is almost totally dependent on computers. Banks provide online accounting facility, which includes checking current balance, making deposits and overdrafts, checking interest charges, shares, and trustee records. ATM machines which are completely automated are making it even easier for customers to deal with banks.

In marketing, advertising and home shopping, computers are widely used today. Advertising professionals use graphics, and animations fascinatedly to get customer attractions. Home shopping

has been made possible through the use of computerized catalogues that provide access to product information and permit direct entry of orders to be filled by the customers.

Engineering - CAD (Computer Aided Design) that provides creation and modification of drawing. 3D virtualization also can be done in engineering industry before development of buildings, machinery, and vehicles also. Some machinery can be developed virtually and can be monitored functioning animations on screens. Computer Aided Software Engineering tools (CASE) supports software engineers for designing, coding, testing and debugging of software.

Competency Level 1.7: Evaluates the impact of ICT in the society

Social and economic benefits of ICT

Benefits provided by ICT such as social benefits and economic benefits would tend people to embrace the information communication technology.

ICT impact in entertainment and leisure industry has improved the way in which we can entertain ourselves by enhancing and providing people with different ways of entertaining ourselves such as online games and chat with friends etc.

A growing ICT sector can contribute to aggregate increases in productivity, GDP and trade. As people continue to embrace ICT with an aim of making their life better, it may also affect how productive they are. Increase in the size and productivity of the ICT sector, it associate growth in labor productivity, manufacturing productivity and ultimately the economy.

ICT has bridged access to information from anywhere anytime which improves information dissemination around the world. Cloud computing activates better application, services, security, reliability and availability of information.

Social, economic, environmental, ethical, legal and privacy issues caused by ICT

ICT adoption in industries causes thousands of job losses. This is because computers reduce the number of people needed to perform a particular task.

Negative impacts of entertaining using ICT are time wastage, health issues. Continuous use of ICT for entertainment such as playing games brings health issues. For instance, constant use of thumbs, wrists and eyes can lead to problems such as eyes strain, RSI and other physical problems.

"Digital divide" is an economic and social inequality with regard to access to, use of, or impact of information and communication technologies. Access to appropriate technological resources similar and improvements in ICT education will helpful to overcome digital divide.

Electronic waste or e-waste is becoming a severe problem around the world. The major problem with e-waste is that the toxic chemicals such as lead, arsenic and cadmium can leach into the soil over time and contaminate the drinking water.

Green computing, the study and practice of efficient and eco-friendly computing resources, is now under the attention of not only environmental organizations, but also businesses from other industries. As a result, most countries in the developed world have introduced regulations to prevent e-waste being dumped into landfill. The majority of components in electronic goods now have to be recycled.

Intellectual property rights, patents on academic material, music and software difficult to be protected as how fast they are distributed over networks. Net etiquette/internet ethics are not kept by all of the Internet users.

One of the biggest concerns voiced by parents of young Internet users is the easy access to **pornography** that the Web provides. There are millions of porn sites online, making hardcore sexual images that were once very difficult to obtain now just a click away.

Trying to control porn in a global medium like the Internet is difficult at best. Most Internet pornography, while offensive and distasteful to many users, is not illegal. As well, countries have

different cultural standards and legislation regarding sexual material and content that is banned in one jurisdiction may be easily accessible on servers in another.

Confidentiality of data or information

Confidentiality of information can be kept with the use of password protection, encryption techniques and digital signatures.

Stealing / Phishing

Phishing is the attempt to acquire sensitive information such as usernames, passwords, and credit card details for malicious reasons, by masquerading as a trustworthy entity in an electronic communication. To protect data or information against stealing or phishing it is necessary to setup firewall, use of password, apply updates and not opening suspicious emails and clicking on links of untrusted senders.

Software piracy is the illegal copying, distribution, or use of software.

Copyright / Intellectual Property is the area of law that deals with protecting the rights of those who create original works. It covers everything from original plays and novels to inventions.

Plagiarism is the stealing of someone else's thoughts, ideas, expressions or work and the representation of them as one's own original work.

Licensed / unlicensed software - A software license is a legal instrument governing the use or redistribution of software. Two common categories for software under copyright law are proprietary software and free and open source software (FOSS). FOSS software licenses both rights to the customer and therefore bundles the modifiable source code with the software (opensource), while proprietary software typically does not license these rights and therefore keeps the source code hidden (closed source). License grants the licensee specific rights to use, edit or redistribution of software. Unlicensed software outside the copyright protection is either public software or software which is non-distributed, non-licensed and handled as internal resource of an organization.