# Unit 9 Security, Privacy and Data Integrity

Data privacy, also called information privacy, is the aspect of information technology (IT) that deals with the ability an organization or individual has to determine what data in a computer system can be shared with third parties.

**Data Privacy:** a requirement for data to be available only to authorized users.

The legal framework which penalizes offenders who breach this privacy is data protection law.

1. The major focus relates to personal, therefore private, data that an individual supply to an organization.
2. The data is supplied to allow the organization to use it but only for purposes understood and agreed by the individual.
3. Data protection laws oblige organizations to ensure the privacy and the integrity of this data.
4. Unfortunately, having laws does not guarantee adherence to them but they do act as a deterrent if wrong-doers can be subject to legal proceedings.

Data security means protecting data, such as a database, from destructive forces and from the unwanted actions of unauthorized users.

**Data Security:** a requirement for data to be available for use when needed, ensures that only authorized users have access to the system and data can be recovered if lost or corrupted.

**Data Integrity:** a requirement for data to be accurate and up to date.

## Threats to the security of a computer system and of the data stored in it.

The threats to the security of a system include the following types:

1. Individual user not taking appropriate care.
2. Internal mismanagement.
3. Natural disasters.
4. Unauthorized intrusion into the system by an individual.
5. Malicious software entering the system.

## Threats to computer and data security posed by networks and the Internet

As internet is not a stand-alone system; one cause of concern is the hacker who is someone intent on gaining unauthorized access to a computer system. A hacker who achieves this aim

might gain access to private data. Alternatively, a hacker might cause problems by deleting files or causing problems with the running of the system. The other major cause of concern is malicious software entering the system.

## Types of malware

Malware is the everyday name for malicious software. It is software that is introduced into a system for a harmful purpose. One category of malware is where program code is introduced to a system. The various types of malware-containing program code are:

1. **Virus:** tries to replicate itself inside other executable code.
2. **Worm:** runs independently and transfers itself to other network hosts.
3. **Logic bomb:** stays inactive until some condition is met.
4. **Trojan horse:** replaces all or part of a previously useful program.
5. **Spyware:** collects information and transmits it to another system.
6. **Bot:** takes control of another computer and uses it to launch attacks.

The differences between the different types are not large and some examples come into more than one of these categories. The virus category is often subdivided according to the software that the virus attaches itself to. Examples are boot sector viruses and macro viruses. Malware can also be classified in terms of the activity involved:

1. **Phishing:** sending an email or electronic message from an apparently legitimate source requesting confidential information.
2. **Pharming:** setting up a bogus website which appears to be a legitimate site.
3. **Keylogger:** recording keyboard usage by the legitimate user of the system.

## System vulnerability arising from user activity

Many system vulnerabilities are associated directly with the activities of legitimate users of a system. Two examples which do not involve malware are as follows.

1. The use of weak passwords and particularly those which have a direct connection to the user. It gives the would-be hacker a strong chance of guessing the password and thus being able to gain unauthorized access.
2. A legitimate user not recognizing a phishing or pharming attack and, as a result, giving away sensitive information.

A legitimate user with a grievance might introduce malware deliberately. More often, malware is introduced accidentally by the user. Typical examples of actions that might introduce malware are:

1. Attaching a portable storage device.
2. Opening an email attachment.

3. Accessing a website.
4. Downloading a file from the Internet.

## Vulnerability arising from within the system itself

1. Operating system often lack good security. Operating systems have regular updates, often because of a newly discovered security vulnerability.
2. In the past, commonly used application packages allowed macro viruses to spread, but this particular problem is now largely under control.
3. A very specific vulnerability is buffer overflow. Programs written in the C programming language, of which there are very many, do not automatically carry out array bound checks. A program can be written to deliberately write code to the part of memory that is outside the address range defined for the array, set up as a buffer. The program overwrites what is stored there so when a later program reads this overwritten section it will not execute as it should. Sometimes this only causes minor disruption, but a cleverly designed program can permit an attacker to gain unauthorized access to the system and cause serious problems.

## Security measures for protecting computer systems

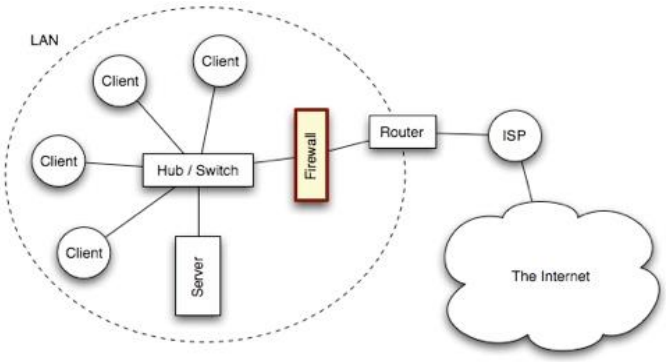
**Disaster Recovery:** Continuity of operation is vital for large computer installations. If there has to be a system shut-down, at the very least to guarantee that the service will start again within a very short time. If an organization has a full system always ready to replace the normally operational one (host site), such a system has to be remote from the original system to allow recovery from natural disasters such as earthquake or flood.

**Safe System Update:** A company is never closed for business. As a result, organizations may need to have the original system and its replacement running in parallel for a period to ensure continuity of service.

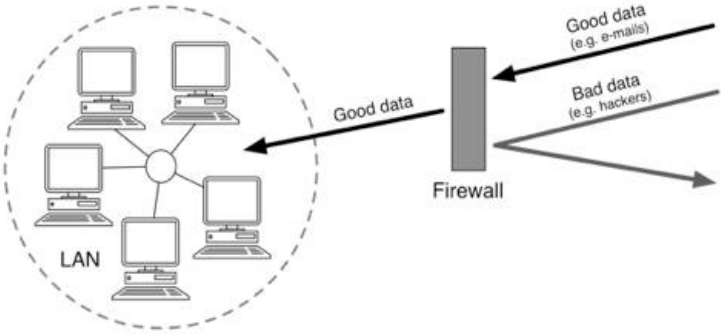
**Preventing Unauthorized Access:** There are a number of security measures that you can take to prevent hackers accessing your computer and all of the data stored on it. Even if a PC is used by only one person there should be a user account set up. User accounts are, of course, essential for a multi-user (timesharing) system. The normal method is to associate a password with each account. Alternative methods of authentication include biometric methods and security tokens.

**Good Practice:** The first thing to make sure of is that no unauthorized people can physically access (sit down in front of) any of the computers on your network. For example, by keeping office doors locked. Portable devices increase the risk of transferring malware into the system. This risk is reduced by an organization having a policy banning the use of such devices or at least limiting their use.

**Always Install and Use a Firewall:** A firewall is a device, or a piece of software that is placed between your computer / LAN and the rest of the network / WAN (where the hackers are!). If you wish to protect your whole LAN from hackers out on the Internet, you would place a firewall between the LAN and the Internet connection.




A firewall blocks unauthorized connections being made to your computer or LAN. Normal data is allowed through the firewall (e.g. e-mails or web pages) but all other data is blocked.




**Digital Signature:** A digital signature is a mathematical scheme for demonstrating the authenticity of a digital message or documents. A valid digital signature gives a recipient reason to believe that the message was created by a known sender, that the sender cannot deny having sent the message (authentication and non-repudiation), and that the message was not altered in transit (integrity). Digital signatures are commonly used for software distribution, financial transactions, and in other cases where it is important to detect forgery or tampering.

## Security measures designed to protect the security of data,

In addition to problems arising from malicious activity there are a variety of reasons for accidental loss of data:

1. A disk or tape gets corrupted.
2. A disk or tape is destroyed.
3. The system crashes.

4. The file is erased or overwritten by mistake.
5. The location of the file is forgotten.

**Data Backup - backup**, or the process of backing up, refers to the copying and archiving of computer data so it may be used to *restore* the original after a data loss event. Data is stored on some data storage medium such as, +, solid state storage and remote backup system.

**Disk Mirroring** - Disk mirroring is a form of disk backup in which anything that is written to a disk is simultaneously written to a second disk. This creates fault tolerance in the critical storage systems. If a physical hardware failure occurs in a disk system, the data is not lost, as the other hard disk contains an exact copy of that data. Mirroring can be either hardware or software based.

**Restricting Access to Data:** If a user has logged in, they have been authorized to use the computer system. The solution is to have an authorization policy which gives different access rights to different files for different individuals.

**Encryption:** It means to scramble data in such a way that only someone with the secret code or key can read it. Encryption works by scrambling the original message with a very large digital number (key). This is done using advanced mathematics. Commercial-level encryption uses 128-bit key that is very, very hard to crack. The computer receiving the message knows the digital key and so is able to work out the original message.

**Symmetric Encryption:** It is the oldest and best-known technique. A secret key, which can be a number, a word, or just a string of random letters, is applied to the text of a message to change the content in a particular way. As long as both sender and recipient know the secret key, they can encrypt and decrypt all messages that use this key.

**Asymmetric Encryption:** The problem with secret keys is exchanging them over the Internet or a large network while preventing them from falling into the wrong hands. Anyone who knows the secret key can decrypt the message. One answer is asymmetric encryption, in which there are two related keys--a key pair. A public key is made freely available to anyone who might want to send you a message. A second, private key is kept secret, so that only you know it.

Any message (text, binary files, or documents) that are encrypted by using the public key can only be decrypted by applying the same algorithm, but by using the matching private key. Any message that is encrypted by using the private key can only be decrypted by using the matching public key.

**Validation:** It is an automatic computer check to ensure that the data entered is sensible and reasonable. It does not check the accuracy of data.

For example, a secondary school student is likely to be aged between 11 and 16. The computer can be programmed only to accept numbers between 11 and 16. This is range check.

However, this does not guarantee that the number typed in is correct. For example, a student's age might be 14, but if 11 is entered it will be valid but incorrect.

**Types of validation**

There are a number of validation types that can be used to check the data that is being entered.

| Validation type | How it works | Example usage |
|---|---|---|
| Check digit | the last one or two digits in a code are used to check the other digits are correct | bar code readers in supermarkets use check digits |
| Format check | checks the data is in the right format | a National Insurance number is in the form LL 99 99 99 L where L is any letter and 9 is any number |
| Length check | checks the data isn't too short or too long | a password which needs to be six letters long |
| Lookup table | looks up acceptable values in a table | there are only seven possible days of the week |
| Presence check | checks that data has been entered into a *field* | in most *databases* a *key field* cannot be left blank |
| Range check | checks that a value falls within the specified range | number of hours worked must be less than 50 and more than 0 |
| Spell check | looks up words in a dictionary | when word processing |

**Verification** - Verification is performed to ensure that the data entered exactly matches the original source.

There are two main methods of verification:

1. **Double entry** - entering the data twice and comparing the two copies. This effectively doubles the workload, and as most people are paid by the hour, it costs more too.

2. **Proofreading data** - this method involves someone checking the data entered against the original document. This is also time consuming and costly.

**Parity Check- Use your O level Notes.**

**CheckSum – Use your O level Notes.**