Data Security and Privacy

Data Integrity

Data integrity is the overall **accuracy**, **completeness**, and **consistency** of data. We want data to be accurate of course. We also want data to be complete for example the ID cards department cannot have information on a person without having his or her address. Such data would be incomplete. Data must be consistent for example if two databases hold information on a person both databases must have the same address for the same person.

There is an assortment of factors that can affect the integrity of the data stored in a database. A few examples include:

- Human error: this means that an operator that is typing in data makes an error.
- Transfer errors: this means that data sent from a source arrives at the destination address corrupted.
- Bugs: errors in programs. Programs today contain a huge number of lines of code and it is not rare that programming errors (called bugs) are present in the program.
- Malware: e.g. virus or logic bomb that is written with the intention of altering or deleting data. Other malware like spyware steal your data but do not change or delete it.
- Compromised hardware: e.g. a hard disk that is malfunctioning can lose data forever.

Measures for Data Integrity

Risks to data integrity can easily be minimized or eliminated by doing the following:

- **Verification**: data is entered twice by two different operators and the computer checks that both entries match.
- Limiting access to data and changing permissions to restrict changes to information by unauthorized parties.
- **Validating data** to make sure it’s correct both when it’s gathered and used i.e. the computer will have a means to check that the input makes sense e.g. if you enter the day of the month, then this must be a number from 1 to 31.
- **Backing** up data.

- Using **logs** to keep track of when data is added, modified, or deleted and by whom.
- Conducting regular **data audits**: a data audit refers to the auditing of data to assess its quality or utility for a specific purpose. Apart from other tools data audits also examine the following questions: What is the integrity of this data? Does it have bias? Is some of the data more or less accurate?
- Using **error detection software**.

Data Security

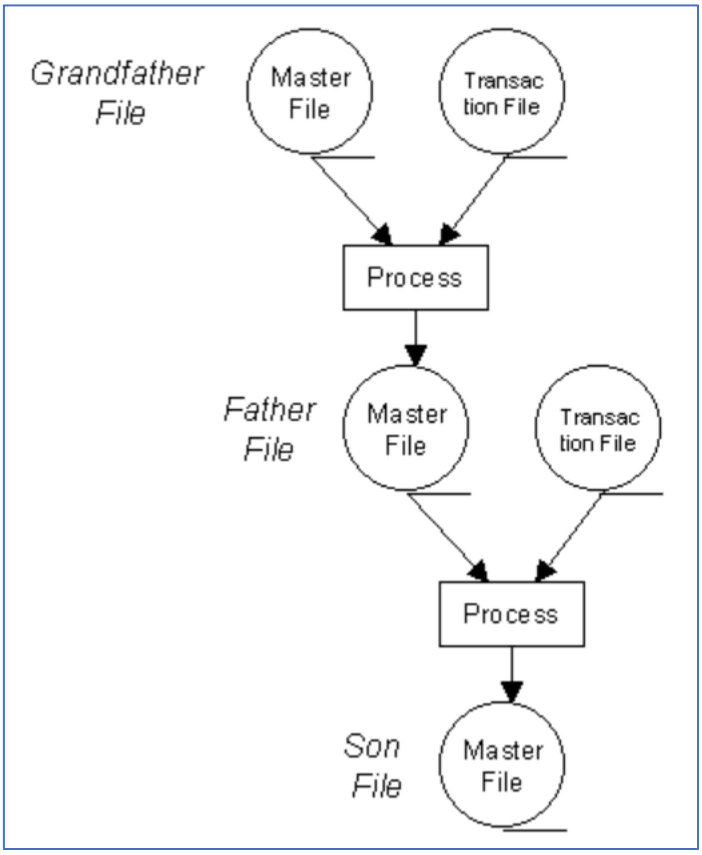
Data security refers to the process of protecting data from **unauthorized access** and **data corruption** throughout its lifecycle.

Measures for Data Security

There are many data security technologies and processes. Some are explained hereunder:

- **Authentication**: it is a process where the user, on log-on, will be checked and recognized (or not recognized) by the system. This process can consist of the user being asked for a username and password. Other means are **biometric** devices (e.g. a scanner that can recognize a fingerprint), **security tokens** or **swipe cards**.
- **Authorization**: it is a process by which a server determines if the client has permission to use a resource or access a file e.g. a user can have permission to read a document but not to change it. If a user asks to view a file, the system will check whether the user has or has not permission to do so. If he or she has the permission, then authorization is granted.
- **Backups** and **recovery**: Backups are copies of your data on a secondary storage medium (disk, tape) or on the cloud. Besides backup one should have a plan how to access backups in case of system failure, disaster, data corruption, or breach. It is very important to perform regular backups.

One way to apply a recovery is by keeping backup copies of a few file generations. If a master file M is updated by transaction file TF, then a new master file (call it MA) is generated. Now MA is called the child (or son) of M and M is called the parent (or father) of MA. For security purposes, besides MA even the files M and TF are stored. When another transaction file – call it TF2 – is used to update MA, then another new master file MB is created. MB will be the grandchild of M etc. Depending on the security requirements of the system, even the grandfather file together can be kept as backup together with the two transaction files.



- **Encryption** is a technique used to save or transmit data in a format that is unreadable to someone who succeeds in accessing it.

Note that security is kept both through physical measures and by means of software safeguards:

- Physical: biometrics, CCTV cameras, etc
- Software: passwords, encryption, etc

Privacy

Privacy of information is guarded by security measures and even protected by law.

Grades of Security

The size and usage of computer systems make different backup demands. For example, a mainframe computer system of a bank requires different and more sophisticated backup provisions than a small microcomputer system used by a shop owner for stock control purposes.

Malta Data Protection Act

The mission statement found at the Data Protection Commission website (http://idpc.gov.mt/) is this: “Our mission is to protect the individual's right to privacy by ensuring the correct processing of personal data.”

The Office of the Information and Data Protection Commissioner is committed to protect the individual's right to privacy by ensuring the correct processing of personal data. The Office aims to safeguard this fundamental human right which is enshrined in the Constitution of Malta.

The Data Protection Act of the European Parliament and of the Council was completely brought into effect on 15 July 2003. The main objectives of the law are twofold:

- The regulation of **data controllers**. A data controller is a person, company, or other body that determines the purpose and means of personal data processing.
- the protection of privacy rights of an individual, including the right to information, the right of access and the right to rectify (correct), block or erase personal data not processed in accordance with the Act.

In the implementation of the local legislation, the **data protection supervisory authority** has been vested in the **Information and Data Protection Commissioner** who enjoys the same independence attributed to a Judge in the local courts and acts independently without being subject to any direction or control of any other person or authority.

The Act empowers the Commissioner, among other things, with the following functions:

- To create and maintain a **public register** of all processing operations being notified by Data Controllers.
- To institute civil legal proceedings in cases where the provisions of the Act have been or are about to be violated.
- To encourage the drawing up of suitable codes of conduct by the various sectors.
- To order the blocking, erasure or destruction of data, to impose a temporary or definitive ban on processing, or to warn the controller.

- to collaborate with supervisory authorities of other countries to the extent necessary for the performance of his duties.

The following are the **Principles of Data Protection**. They are nine principles of ‘good information handling’. The controller shall ensure that:

- Personal data is processed fairly and lawfully.
- Personal data is always processed in accordance with good practice.
- Personal data is only collected for specific, explicitly stated and legitimate purposes.
- Personal data is not processed for any purpose that is incompatible with that for which the information is collected.
- Personal data that is processed is adequate and relevant in relation to the purposes of the processing.
- No more personal data is processed than is necessary having regard to the purposes of the processing.
- Personal data that is processed is correct and, if necessary, up to date.
- All reasonable measures are taken to complete, correct, block or erase data to the extent that such data is incomplete or incorrect, having regard to the purposes for which they are processed.
- Personal data is not kept for a period longer than is necessary, having regard to the purposes for which they are processed.

Plagiarism

What is plagiarism? Many people think of plagiarism as copying another's work or borrowing someone else's original ideas. But terms like "copying" and "borrowing" can disguise the seriousness of the offense. All of the following are considered as **plagiarism**:

- turning in someone else's work as your own
- copying words or ideas from someone else without giving credit
- failing to put a quotation in quotation marks
- giving incorrect information about the source of a quotation
- changing words but copying the sentence structure of a source without giving credit

- copying so many words or ideas from a source that it makes up most of your work, whether you give credit or not (see our section on "fair use" rules)

Most cases of plagiarism can be avoided, however, by citing **sources**. Simply acknowledging that certain material has been borrowed and providing your audience with the information necessary to find that source is usually enough to prevent plagiarism.

By the way, this information about plagiarism was taken from the site http://plagiarism.org/. Had I not told you so, then I would have plagiarised.

Plagiarism software can indicate the approximate percentage of plagiarism is a text.

Software Piracy and Copyright

Ethical Issues

Software piracy is the unauthorized use, copying or distribution of copyrighted software. It may take many forms.

Because a software pirate does not have proper permission from the software owner to take or use the software in question, piracy is the equivalent of theft and is, therefore, a crime.

There are contrasting ethical views on the issue of piracy.

Some think that there is nothing wrong with software piracy. They believe in the freedom of information and expression (ie. "information wants to be free"). According to them, it is acceptable and ethical to copy the software because they have a right to the information contained in the software. They also believe that, with the rising prices of software, software manufacturers are really not hurt by pirates making illegal copies of their programs.

Those who argue against piracy say that people who write the software have rights to profit from it, just as people who write books have the sole right to sell them.

They say that claim that pirates have a right to make illegal copies of software because the software is buggy, or too expensive, or not frequently used by the pirate, is also flawed. Someone might think a Rolls-Royce is too expensive and not worth the money, but this doesn't give him the right to steal it.

Types of Piracy

Here are some examples:

- Serial (product key). A **serial**, also known as a product key, is a series of alphanumeric characters (numbers and letters) that indicate the purchase of that specific software. Serials for very popular software can sometimes be found online and most software pirates use these serials to distribute pirated copies that can be unlocked.
  Some software makers are beginning to require online registration to mitigate this problem.
- A **crack** is more sophisticated than a serial. It is an application (program) that totally removes the authentication mechanism that has been embedded into the software.
- The act of obtaining copyrighted software legally and installing it on more computers than allowed is called **softlifting**.
- **Internet piracy** is a form of piracy in which unauthorized copies of copyrighted software are distributed in an electronic form.
- **Mischannelling** refers to the misuse of software that was acquired legally. An example would be a situation where, for example, an academic license is used for commercial gain.
- **Reverse engineering** is the study of a program to unravel the way it is working. This facilitates remaking a similar program or copying techniques used by the original program. Commonly used tools in reverse engineering are debuggers, disassemblers, and decompilers.
  - **Debuggers** enable the dynamic analysis of the program during its execution.
  - **Disassemblers** generate assembler code from executables.
  - The function of **decompilers** is to recreate high-level source code which then corresponds to the original code of the target software.

Hardware Techniques for Breaking Copy Protection

In these techniques, chips can be used to analyse what the software is doing. The techniques sometimes can work even if the code is encrypted (e.g. in the case of conditional branching). Memory locations and bus transactions can be traced. Encryption is one way of combating this type of piracy.

Ethical measures

The aim of the ethical measures is to make software piracy appear morally unappealing. In this case deterrents like shame, formal sanctions or moral beliefs have been utilized in the fight against software piracy.

Studies have demonstrated that ethics has an impact on individuals' tendency to engage in acts of software piracy and thus they should not be disregarded.

Legal measures

The legal measures attempt to prevent piracy by creating fear of consequences of being caught in the act of piracy. Legal measures act as a deterrent by threatening possible offenders with legal sanctions.

Legal deterrents are beyond the control of software publishers, and instead they rely on consumers' awareness of copyright laws and the enforcement of these laws by the government.

However, it is questionable whether the legal measures are effective in the fight against piracy. In fact, there is no evidence that increased awareness of piracy as a crime and the fear of legal sanctions have slowed down the spreading of piracy.

Technical measures

There are software-based and hardware-based technical measures against software piracy.

Software-based measures

Software-based measures include measures like:

- Software tokens
- **Watermarking**: In watermarking a copyright notice is embedded in the software code, enabling the owners of the software assert their intellectual property rights. The copyright notice can be extracted from the program to identify the owner of the copyright or an authentic user of the program. There is also a technique related to watermarking called **fingerprinting**. In this case a watermark that varies from one copy of software to another is created.
- **Obfuscation**: this is a method which attempts to transform the target program into a more complex version which is harder to reverse engineer. Code obfuscations aim to obscure the purpose of the code without changing its operation.

- **Encryption**.
- **Tamper-proofing**: this technique causes the program to malfunction when it detects that the program has been modified. One tamper-proofing technique uses **checksums**. For example, one method is to examine the program before running it and compare it with the original program using checksums. The problem with checksums, however, is that their nature may be difficult to conceal, and in case of detection the attacker can easily remove them.

Hardware-based measures

The measures based on hardware, use specific hardware tokens which are required for successfully installing or running the software. Hardware security tokens are typically portable and small.

The applications check for the presence of the token and often they will refuse to install or run if the token is not present.

The hardware items can be CDs, dongles (hardware keys), expansion cards, and smart cards.

Later Commodore PET cassette port dongle built by Brian Edmundson
Later Commodore PET cassette port dongle built by Brian Edmundson
Earliest Commodore PET cassette port dongle built by Graham Heggie at the request of Pete Dowson and Mike Lake
IBM PC 25 pin serial port dongle built by Brian Edmundson
IBM PC 9 pin serial port dongle built by Brian Edmundson
IBM PC 25 pin serial port dongle built by Brian Edmundson
Published by DATAVIEW Colchester, England
WORD CRAFT
© P.L. Dowson
CBM
K
DEMO
WORD CRAFT
Dataview Wordcraft Colchester & Derby England

Early protection dongles.

Other Measures

- Restricted access to computer areas.
- IDs and passwords.
- Log and audit trail. An audit trail is a record of the changes that have been made to a database or file. This will help find the culprit of purposeful corruption and give indications for system recovery.

Access Rights

Most networks will have been set up with 'access rights'. This means the administrator has set up each person who can log on, with the right to access certain files and folders. For instance, you may have a personal folder in which you have the right to open, read, write, create, and delete files.

Other parts of the network may have files you can only read but not write. And finally, there will be areas that you cannot enter at all. In which case a message will often pop up to inform you "You do not have sufficient access rights" or something similar.

You may also inherit access rights by being assigned to a 'group'. For example, the admin may have set up a student group and set access rights to the group. Then, once you become a member, you automatically get the same access rights.