UNIVERSITY OF WESTMINSTER

7BUIS030W
Data System Concepts and Fundamentals

Lecture-1

Dr. Saumya Kareem Reni

Lecture-1 Outline

Information Systems, Information systems life cycle, Data and data cycle, Data actors, Data security , compliance and frame work, Data representation

Information Systems

- **Information systems** are resources that enable the **collection**, **management**, **control** and **dissemination** of information throughout an organisation.
- Information systems support decision making, coordination , analysis and visualisation of the organisational resources.
- Information systems take data and turn it into information.
- Business firms and other organizations rely on information systems to carry out and manage their operations, interact with their customers and suppliers, and compete in the marketplace.

Applications of Information System

- Information systems are used to run inter-organizational supply chains and electronic markets.
- Corporations use information systems to process financial accounts, to manage their human resources, and to reach their potential customers with online promotions.
- Companies like ebay and amazon are built entirely around information systems.
- Most of the revenue from Google is through keyword advertising on internet searches.
- Governments deploy information systems to provide services cost-effectively to citizens.
- Individuals rely on internet and media based information systems, for socializing, study, shopping, banking, and entertainment.

Data and Information

Data: A single quality or quantity of some object or phenomenon.

It is raw, unorganized facts that need to be processed.

Eg: Temperature reading around the world, Collection of Medical images, number of people in a county

Information: Data when processed and transformed into something useful. When data is processed, organized, structured or presented in a given context so as to make it useful, it is called information.

Eg: The temperature readings are analysed to establish that there is a rise in global warming, Medical images processed and analysed to diagnose a medical condition, Number of people in a county to establish the population in the county

Data and Information

Peter earns £10 an hour

Peter 's company does not comply with the national hourly wages payment plan

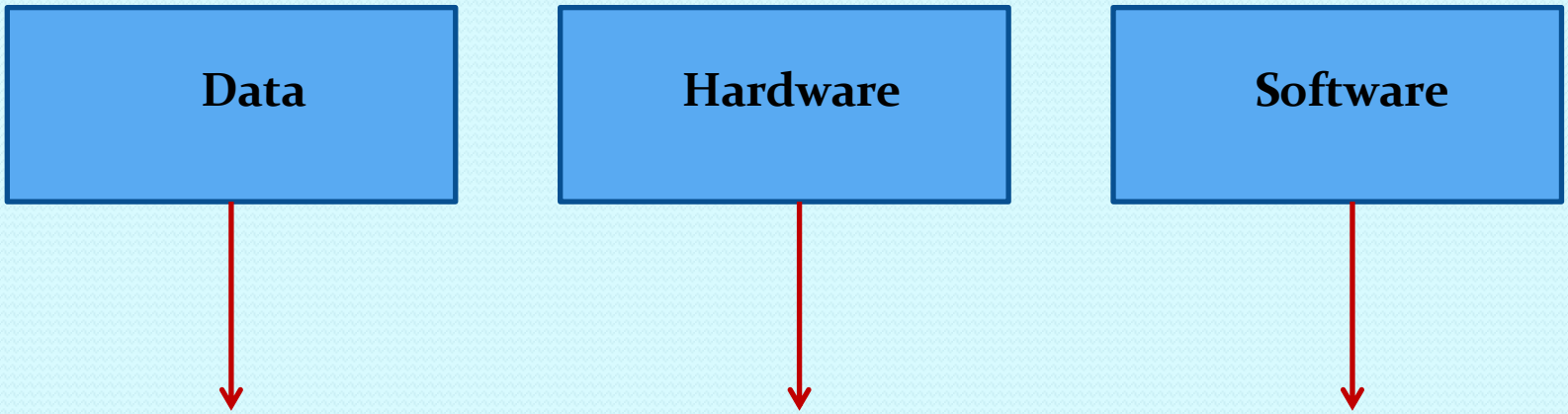
John spend £130 in Tesco on Dec 28th 2019

The total sale of Tesco is £710,115 in Dec 2019

George scored 72% in his In-Class test

The average of the test scores is 58%

Components of information systems

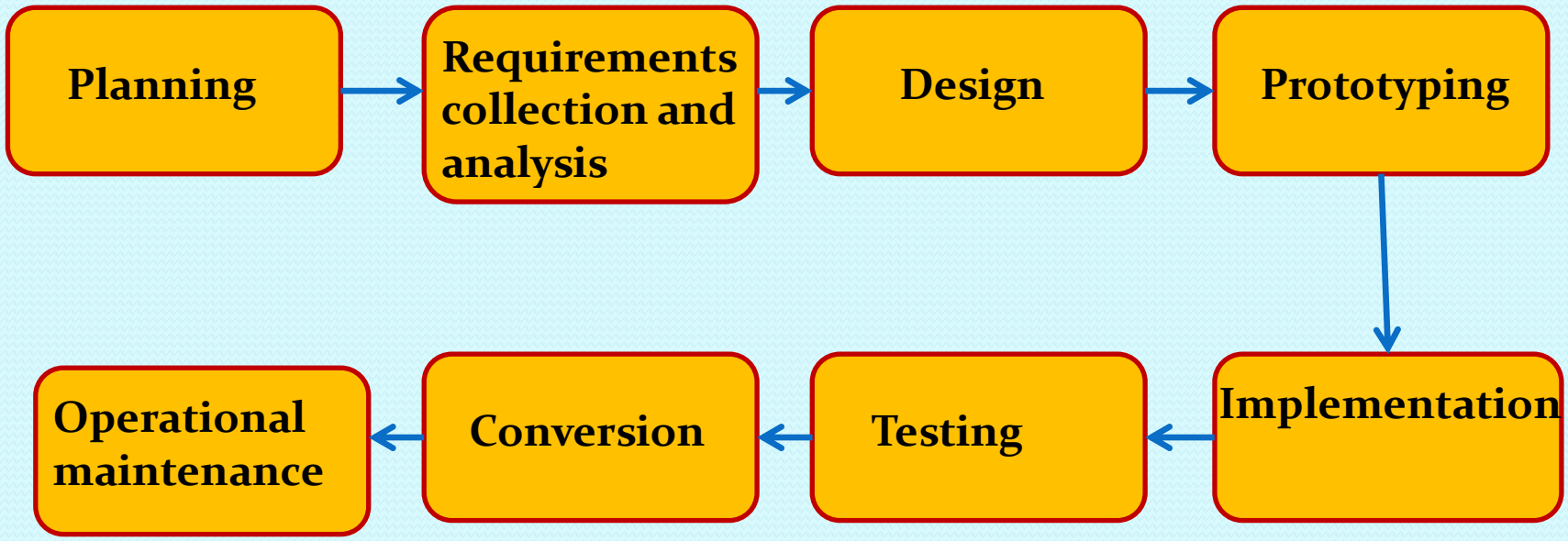


Data represented in the form of variables , numbers or any other file format

Physical part of IS
Computers,
keyboards,
disk drives,
iPads,
flash drives

A set of instructions that tells the hardware what to do.
Data base,
Database software,
Operating system software,
Application software

Information system life cycle



Information system life cycle

Planning

How different stages of the life cycle are implemented effectively

Requirement Collection and analysis

Process of collecting and analysing data about the required organisation or fact to be converted to information and identify new requirements

Design

Conceptual and physical design of the information system

Information system life cycle

Prototyping

Building a working model of the systems for the designers or users to visualise and evaluate how the final system will look and function.

Implementation

Creating physical systems and applications

Testing

Test the system of errors and validate against the requirements

Information system life cycle

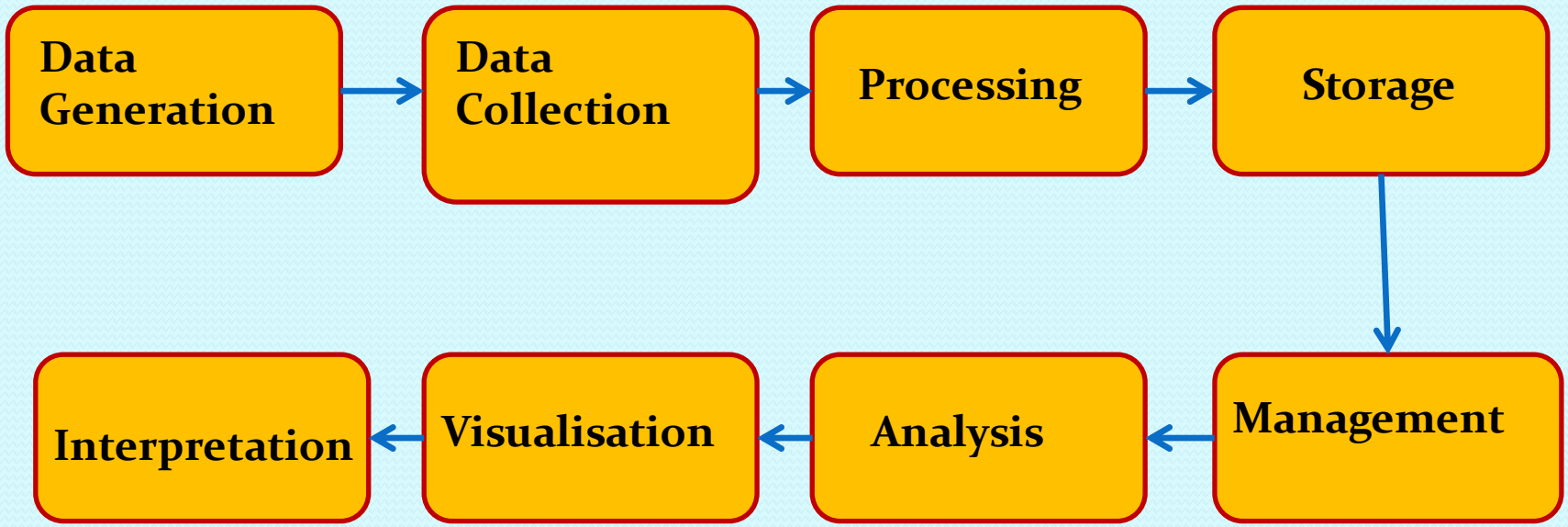
Conversion

Optional, loading the data or facts from old systems to new system

Operational maintenance

Continuous monitoring of the implemented system, incorporate new requirements.

Data Life Cycle



Data Life cycle

Data Generation:

Data is generated on a variety of areas.

- Digital footprint,
- Market analysis,
- Medical sensors,
- Images,
- Surveys, etc

Data Collection:

- Not all data generated is collected.
- Generated data is filtered
- Filtering will aid in deciding whether to collect it based on its usefulness and processing compatibility.

Data Life Cycle

Data Processing:

Processing involves:

- data cleaning,
- data wrangling,
- data formatting,
- data compression for efficient storage (image and audio files),
- data encryption for secure storage

Data Life Cycle

Data compression:

Loss less compression-

- The file is compressed without losing information.
- Compressed file can be decompressed to its original format

Eg: ***reducing the size of repeating string of items through encoding-Run Length Encoding[cccmmmmm- 3c5m]***

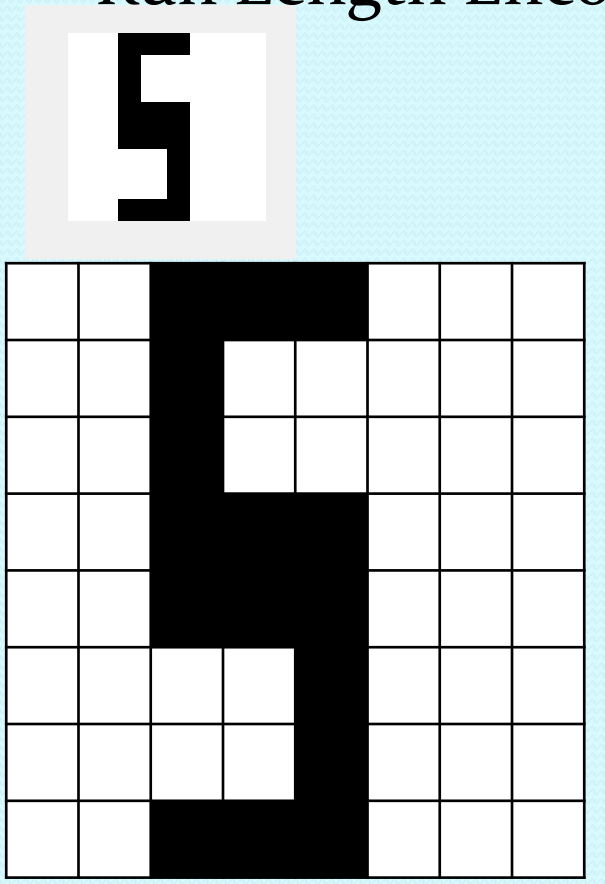
Lossy compression:

- Data is compressed by removing some of it.
- Reconstruction of the original file is not possible

Eg: ***jpeg, bmp images, wav, mp3 audio files***

Data Life Cycle

Run Length Encoding Demonstration- bitmap image



Code	RLE version	Size of RLE code
wwbbbwww	2w3b3w	6
wwbwwwww	2w1b5w	6
wwbwwwww	2w1b5w	6
wwbbbwww	2w3b3w	6
wwbbbwww	2w3b3w	6
wwwwbwww	4w1b3w	6
wwwwbwww	4w1b3w	6
wwbbbwww	2w3b3w	6
64bytes		48bytes

Can we reduce the size even further?

Data Life Cycle

Data Storage:

- Processed data is stored in binary format in the computer systems.
- Data stored in different operating system will have different size.

Data Management:

- Stored data is organised in an ordered way –Database
- Information is quickly retrieved through queries

Data Life Cycle

Data Analysis:

Deploy computational and statistical techniques for analyzing data .

Analysed data is used to gain knowledge or insights, build classifiers and predictors, or infer causality.

Eg:

- ***algorithms and methods that underlie artificial intelligence (AI),***
- ***data mining,***
- ***machine learning,***
- ***statistical inference***

Data Visualisation:

- Presenting results in a clear and simple way that a human can readily understand and visualize.
- Along with functionality, aesthetics, and human visual perception to convey the results of data analysis is considered.

Data Interpretation

Provide the user an explanation of what the visualisation means.

Data Security

Data security is protecting digital data that are stored and organised in the database from unwanted actions, unauthorised users or destruction.

Cyberattack is any type of offensive activity that targets computer information systems, infrastructures, computer networks, or personal computer devices.

Data Breach is the intentional or unintentional release of secure or private/confidential information to an untrusted environment. This includes unintentional information disclosure, data leak, information leakage and also data spill.

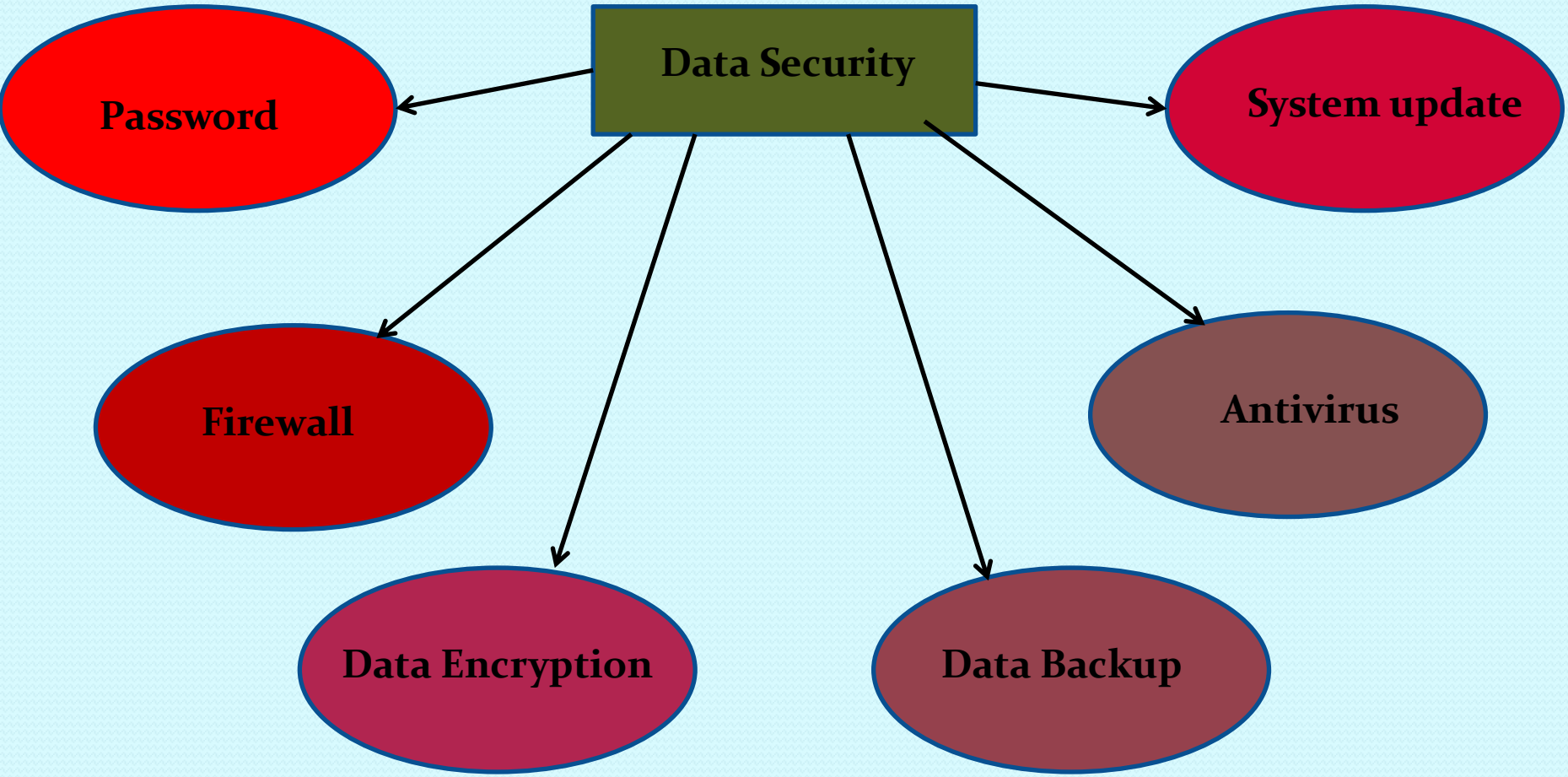
Data Security

Data is compromised widely!

- Websites and mobile applications integrate third party services for behavioral targeting, user analytics, navigation, and many other functionalities.
- Governments build central infrastructures to share data efficiently between different branches of government and with other organisations.
- Social media usage leads to more and more information gets placed online.
- There is an increased danger of hackers, companies, and malicious interlopers mining your data in ways that undermine personal privacy.

Certain cases leads to data stealing!

Data Security Measurements



Data Encryption

- Data can be encrypted on a hard drive through software or hardware.

Disk encryption hardware:

- Built in encryption hardware with an encryption key which is maintained independently from the CPU thus removing computer memory as a potential attack vector.

Disk encryption software

- Bit locker for windows
- File Vault for Apple OS/X
- LUKS , free software for Linux

Data Protection

Data Protection deals with personal data and relates to the following fundamental principles:

1. Personal data shall be processed fairly and lawfully,
2. Personal data shall be obtained only for specified, lawful purposes, and shall not be further processed in any manner incompatible with that purpose or those purposes,
3. Personal data shall be adequate, relevant and not excessive in relation to the purpose or purposes for which they are processed.
4. Personal data shall be accurate and, where necessary, kept up to date.
5. Personal data processed for any purpose or purposes shall not be kept for longer than is necessary for that purpose or those purposes.

Data Protection

- Data Security is **NOT** Data protection!!
- Data security is NOT a substitute for data protection

Data Protection	Vs	**Data Security**
•Deals with personal data •Based on fundamental principles •Data protection compliance audits are mainly legal work		•Deals with the protection of databases and systems •Mainly a technical and a procedural issue •Data security audits do not have a legal component

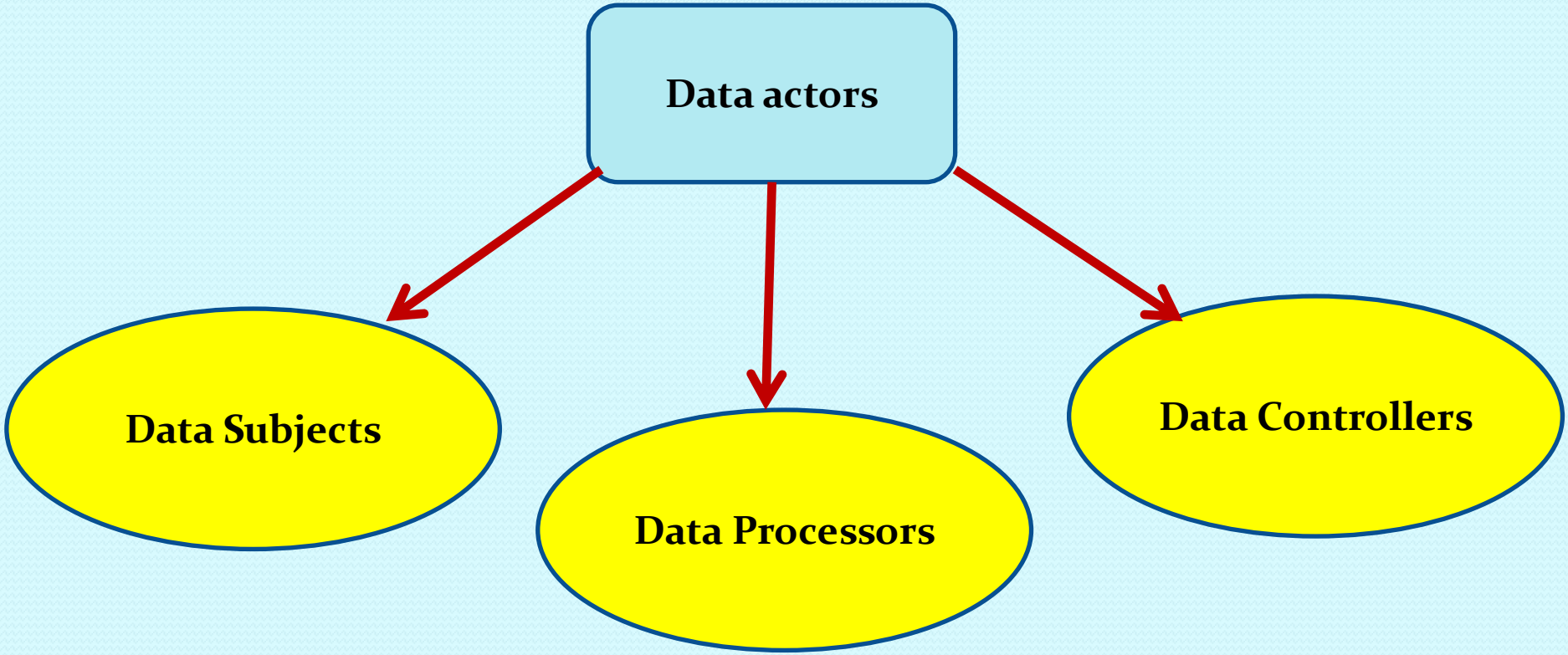
Regulations & Compliance Framework

In order to maintain data protection , the European union has issued a pan-European data protection law called **General Data Protection Regulation (GDPR)** in 2018.

- GDPR addresses the transfer of personal data outside the EU and EEA areas.
- The primary aim is to let individuals control their personal data
- For organisations, a unified regulation was put in place to simplify the regulatory environment for international businesses.

Data actors

- GDPR clarifies the role of various parties involved in data protection. They are collectively called **data actors**.



Data actors

Data Subjects

- A data subject is any person whose personal data is being collected, held or processed.
  - Personal data can refer to anything from your name, home address or your posts on social media.
  - Anyone becomes at some point a data subject when they are applying for a job, booking a flight, using their credit card or just browsing the internet, they disclose some personal data.
  - GDPR rules that in order for their personal data to be processed, data subjects must give their consent unless it is for legal purposes.
  - Data subject has the right to access data about them.
  - They also have the right to their processed data

Data actors

Data Controllers

- A data controller determine the purpose and means of the personal data processing.
  - Data controllers can be a natural person, a legal person, a public authority, agency or other.
  - Data controllers are in charge of the processing of personal data belonging to data subjects.
  - Data controllers can act alone or jointly with others

  For example, your university has a data controller who decides how and when your personal data should be processed

Data actors

Data Processors

- A data processor is the person processing the personal data of the data subjects.
  - Data processor can be an individual , legal person, a corporation, a public authority, agency or any other body which processes personal data .
  - Data processors acts in behalf of data controllers
  - Data processors act in accordance with the instructions of data controllers.
  - The data processor is usually a third party external to the organisation

Data Protection Act

- Data Protection Act (DPA) is the implementation of GDPR executed by individual countries.
- The DPA modifies the EU GDPR by filling in the sections of the Regulation that were left to individual member states to interpret and implement.
- DPA rules are applied to the respective countries.
- For Eg; The Data Protection Act 2018 is the UK's implementation of the General Data Protection Regulation (GDPR).

Compliance Framework

Provides a structure to manage the personal data an organisation can use.

The structure should comply with the GDPR .

- Organisations can have their own have privacy compliance frameworks .
  - They should obtain certification to national and international standards to demonstrate to regulators that due diligence and compliance efforts have been made.
- Organisations that have not developed their own privacy compliance frameworks can use a standardised framework to ease their path to GDPR compliance.

Standardised compliance frameworks

There are two categories of standard compliance framework

Personal information management system (PIMS)

- Provides a well-defined structure for managing data protection,
- Designed to follow the plan-do-check-act cycle (PDCA) to ensure continual improvement
- **BS10012:2017** is the British standard that specifies the requirements for a personal information management system (PIMS)
- **BS10012:2017** is aligned with the requirements of the GDPR.

Standardised compliance frameworks

Information security management system (ISMS)

- Demonstrates that your company follows information security best practice, delivers an independent, expert assessment of whether your data is adequately protected.
- **Certification to ISO 27001** is internationally recognised and does not favour any one technology or solution, and can be used by organisations of any size.
- ISO27001 risk based approach is an excellent in meeting the GDPR requirement that organisations implement with regards to the "**confidentiality, integrity and availability**" of processing systems and services.