SS1 - Topic: DATA REPRESENTATION

Meaning of Data Representation

Data representation is the method of converting information such as text, numbers, pictures and sound into a form that a computer can understand, process and store.
Computers represent and process data using the binary number system, which consists of two digits: **0 and 1**.

Bits and Bytes

Bit (Binary Digit): A bit is the smallest unit of data in a computer. It can have only one of two values: 0 or 1.
Byte: A byte is a group of 8 bits. It is commonly used to represent a character or a small amount of data.
Example:
`01000001` = 8 bits = 1 byte.

Common Units of Data Measurement

Unit	Equivalent
1 Bit	One binary digit (0 or 1)
1 Nibble	4 bits
1 Byte	8 bits
1 Kilobyte (KB)	1,024 bytes
1 Megabyte (MB)	1,024 KB
1 Gigabyte (GB)	1,024 MB
1 Terabyte (TB)	1,024 GB

ASCII (American Standard Code for Information Interchange)

ASCII is a character encoding system used to represent letters, numbers and symbols in computers.
It uses 7 bits to represent 128 different characters in its standard form.
Examples:

Character	ASCII Decimal Code
A	65
B	66
C	67
a	97
b	98
0	48
1	49

ASCII allows computers to store and exchange text information.

Unicode

Unicode is a universal character encoding standard designed to represent characters from different languages, symbols and writing systems around the world.
It supports English, Yoruba, Arabic, Chinese, French and many other languages.
Examples include:

- English letters: A, B, C.
- Yoruba characters: ẹ, ọ, ṣ.
- Arabic characters.
- Chinese characters.
- Mathematical and currency symbols.

Importance of Unicode:

1. Supports many languages.
2. Allows different characters to be represented digitally.
3. Makes international communication easier.
4. Supports special symbols and emojis.

5. Helps prevent problems when exchanging multilingual text.

Difference between ASCII and Unicode

ASCII	Unicode
Supports a limited set of characters.	Supports characters from many languages.
Standard ASCII uses 7 bits.	Uses encoding formats such as UTF-8 and UTF-16.
Mainly represents English letters, numbers and symbols.	Represents international characters and symbols.

5. Representation of Images

Computers represent digital images using tiny dots called **pixels**.
A pixel is the smallest individual element of a digital image.
An image is made up of many pixels arranged in rows and columns.

Important terms:

- **Pixel:** The smallest unit of a digital image.
- **Resolution:** The number of pixels that make up an image, usually expressed as width × height.
- **Colour Depth:** The number of bits used to represent the colour of each pixel.
- **Bitmap:** An image format that stores pictures as pixels.

Examples of image file formats:

- JPEG/JPG
- PNG
- GIF
- BMP

Example: An image with a resolution of 800 × 600 contains 480,000 pixels.

Representation of Sound

Computers represent sound by converting analogue sound waves into digital information.
This process is called **analogue-to-digital conversion**.
Digital sound is commonly represented through sampling and quantisation.

Important terms:

1. **Sampling:** Measuring the sound wave at regular intervals.
2. **Sample Rate:** The number of sound samples taken per second, measured in Hertz (Hz).
3. **Bit Depth:** The number of bits used to represent each sound sample.
4. **Quantisation:** Converting measured sound values into digital numerical values.

Examples of sound file formats:

- MP3
- WAV
- AAC
- WMA

Uses of digital sound:

1. Music playback.
2. Voice recordings.
3. Telephone communication.
4. Video production.
5. Online learning.

Summary

Data Type	Representation Method
Text	ASCII and Unicode
Numbers	Binary digits
Images	Pixels
Sound	Digital samples
Video	Combination of images and sound