

Paul Long, Sarah Lawrey and Victoria Ellis

Cambridge International  
AS and A Level

**SAMPLE IT**

Coursebook



Completely **Cambridge**  
Cambridge resources  
for  
Cambridge qualifications

This book will cover all the topics you need to learn for the international A Level in Information and Communication Technology (ICT). Throughout the book you will find examples of ICT in practice, tasks for you to complete, discussion points or extension exercises to provoke further thought and discussion, and questions that will test your knowledge and understanding.

You will discover that the practical aspects of this book do not include step-by-step tutorials for how to use software, but instead tell you about the capabilities of different types of software. This is deliberate because the book is designed to provide you with knowledge and understanding that you can use in an examination. We also recognise that not all schools and colleges use the same software and there are many of tutorials available on the World Wide Web.

## Examination

The examination will test your knowledge of ICT and your understanding of the impact of ICT on individuals, organisations and society as a whole.

To support your learning, you will find review questions at the end of each chapter. These are examination style questions that will help you to practise your approach to answers. You should also spend time practising answering questions from past examination papers.

Through studying ICT, you will have a holistic understanding of real-world computing that will open many doors into employment and further study. Most importantly, you will understand the importance of a user-centric approach to computing, together with project management and systems development skills that are all-important workplace skills.

SAMPLE

# Chapter 1

## Data, information, knowledge and processing

### Learning objectives

*By the end of this chapter, you will be able to:*

- define the differences between data, information and knowledge
- define static and dynamic data
- compare static information sources with dynamic information sources
- define direct and indirect data sources
- understand the advantages and disadvantages of gathering data from direct and indirect data sources
- understand how the accuracy, relevance, age, level of detail and completeness of information can affect its quality
- describe the coding of data and discuss its advantages and disadvantages
- evaluate the need to encode data and analyse different methods for encoding data
- define encryption and describe different methods of encryption
- evaluate the need for encryption and how it can be used to protect data
- discuss encryption protocols
- define validation and analyse a range of validation methods
- define verification and analyse verification methods
- explain the need for both validation and verification
- define proof reading

## 1.01 Data, information and knowledge



### KEY TERMS

**Data:** raw numbers, letters, symbols, sounds or images without meaning

**Information:** data with context and meaning

**Knowledge:** information to which human experience has been applied

### Data

**Data** are raw numbers, letters, symbols, sounds or images with no meaning.

#### EXAMPLE

P952BR  
@bbcclick  
359  
23557.99

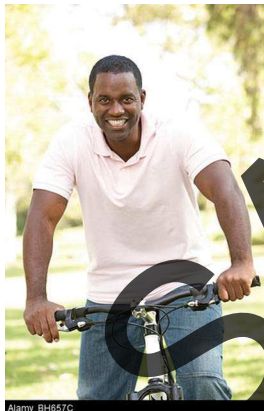


Figure 1.01

The data P952BR could have several meanings. It could possibly be:

- a product code
- a postal / ZIP code
- a car registration number.

As it is not known what the data means, it is meaningless.



### DISCUSSION POINT

When answering a question in an exam such as ‘Give one item of data’, do not try to explain what the data means because it then becomes **information**. Just give the raw numbers, letters, symbols or image.

### Information

When data items are given context and meaning, they become information. A person reading the information will then know what it means.

Data is given context by identifying what sort of data it is. This still does not make it information but it is a step on the way to becoming information as shown in the next example.

#### EXAMPLE

Data	Context	Comment
P952BR	A product code	This is a product code but it is still not known what it is a product code for so it is still data.
@bbcclick	A Twitter handle	This is an address used for Twitter but it is not information unless it is known to be a Twitter handle or used within Twitter software. It's also not known who the address is for.
359	Price in Pakistani Rupees	This is a currency value but it is not known what the price is for so it is still data.

For the data to become information, it needs to be given meaning. Information is useful because it means something.

#### EXAMPLE

Data	Context	Meaning
P952BR	A product code	A product code for a can of noodles.
@bbcclick	A Twitter handle	The Twitter address for the BBC's weekly technology show, <i>Click</i> , which is worth watching on BBC World News and BBC2 to keep up to date with technology.
359	Price in Pakistani rupees	The price of a pack of playing cards.



## Knowledge

**Knowledge** is basically what a person knows. This is known as their knowledge base. A knowledge base gets larger over time as a person gains experience or learning. Knowledge requires a person to understand what information is based on their experience and knowledge base. Crown Prince Salman was appointed Crown Prince of Saudi Arabia on 16 June 2012. This is information. Knowing that he had been Crown Prince for 2 years on 1 August 2014 is knowledge. Knowledge allows data to be interpreted. In computing terms, knowledge is also what a machine knows through the use of a knowledge base consisting of rules and facts.

### EXAMPLE

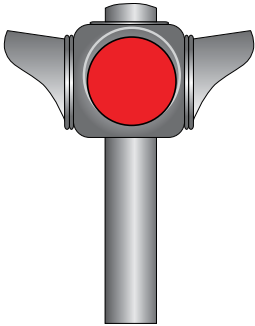
Information	Knowledge
70 mph is the speed limit on British motorways.	Travelling at 80mph on British motorways is illegal.
	A red traffic light means a car should stop.
359 Pakistani rupees is the price of a pack of playing cards.	The price of a pack of cards is more expensive than a cup of coffee.

Figure 1.02 - A red traffic light.

- For example, context is added to 000000 because we are told it is a colour code. Meaning is added by telling us it is the colour code for black.
- Knowledge is applying experience to information so that the information can be interpreted.

## 1.02 Sources of data

### KEY TERMS

**Static data:** data that does not normally change

**Dynamic data:** data that changes automatically without user intervention

**Direct data source:** data that is collected for the purpose for which it will be used

**Indirect data source:** data that was collected for a different purpose (secondary source)

### Static data



Figure 1.03

Static means 'still'. It is data that does not normally change. **Static data** is either fixed or has to be changed manually by editing a document.

### QUESTIONS

A company creates websites using style sheets.

- Identify one item of data that will be used by the company.
- Describe how this item of data can become information.
- Describe the term knowledge.

### Answers

- For example 000000 (there must be no explanation, just the data).

### EXAMPLE

- title of a web page
- magazines
- CD-ROMs
- column headings for a football league table (e.g. P, W, D, L, F, A, GD)
- instructions on a data entry screen

### Dynamic data

Dynamic means ‘moving’. It is data that updates as a result of the source data changing. **Dynamic data** is updated automatically without user intervention.

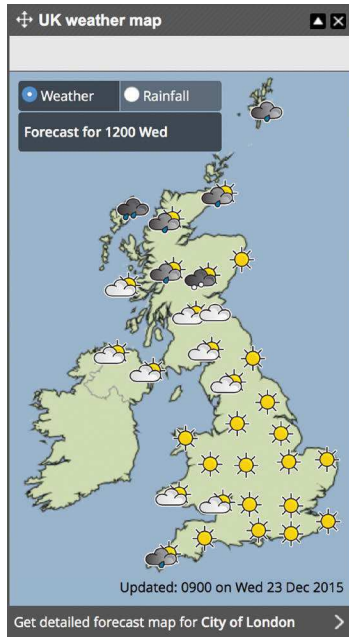


Figure 1.04

### Static information sources compared with dynamic information sources

Static information source	Dynamic information source
The information does not change on a regular basis.	Information is updated automatically when the original data changes.
The information can go out of date quickly because it is not designed to be changed on a regular basis.	It is most likely to be up to date as it changes automatically based on the source data.
The information can be viewed offline because live data is not required.	An internet or network connection to the source data is required, which can be costly and can also be slow in remote areas.
It is more likely to be accurate because time will have been taken to check the information being published, as it will be available for a long period of time.	The data may have been produced very quickly and so may contain errors.

Table 1.01 - Static and dynamic information sources.

#### EXAMPLE

- live sports results on a website (when a goal is scored, the scores will update on the website)
- news feeds on a mobile phone app (when the news is changed in the main database, the news feed will be updated on the phone)
- availability of tickets for a concert (when somebody books a ticket, the ticket is no longer available)
- product prices for a till/point of sale (if a price is reduced or increased in the database, this new price will be used the next time the barcode is scanned)
- train expected arrival times (these will update automatically based on the train’s position)
- profit for a product in a spreadsheet (profit = price – cost so when either the price or cost changes, then the profit changes, too)



#### DISCUSSION POINT

Some people get a little confused by dynamic data because they think it can be any data that changes at any time. For example, some people think that any website includes dynamic data. However, if you look at [www.bbc.co.uk/contact](http://www.bbc.co.uk/contact), the information on this page is mainly static, with links to other pages. Although it can be changed, it can only be done by changing the actual data on the page, whereas dynamic data changes because the original source has changed. If you look at [www.bbc.co.uk/](http://www.bbc.co.uk/) click, you will see that most of the information on this page is sourced from programme schedules and programme information, which will be stored in another database. This makes it dynamic data because it will update when a new programme has been shown.

#### TASK

Look at the BBC website [www.bbc.co.uk](http://www.bbc.co.uk) and identify which data is static and which data is dynamic (updates automatically).

**Direct data source**

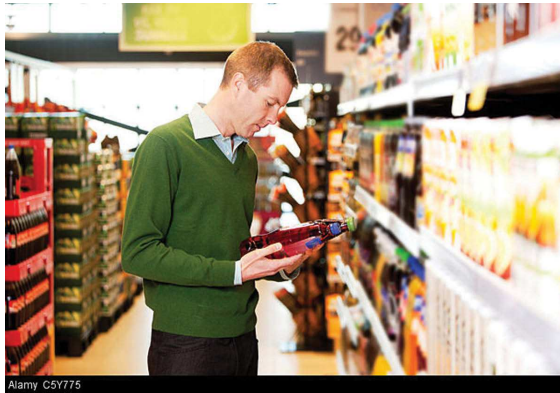


Figure 1.05

Data collected from a **direct data source** (primary source) must be used for the same purpose for which it was collected. It is often the case that the data will have been collected by the person who intends to use the data or has been requested by the person who intends to use the data. The data must not already exist for another purpose though. When collecting the data, the person collecting should know what purpose they intend to use the data for.

**EXAMPLE**

A confectionary shop wants to find out what other shops are charging for sweets. There are various direct sources that this data can be collected from. These could include:

- visiting the other shops and noting down the prices
- visiting the other shops' websites and noting down the prices
- giving out a survey to other shop owners to ask their prices (although they are unlikely to want to do this!).

**Indirect data source**



Figure 1.06

Data collected from an **indirect data source** (secondary source) already existed for another purpose. Although it can still be collected by the person who intends to use it, it was often collected by a different person or organisation.

**EXAMPLE**

The confectionary shop could use various indirect sources to find out what other shops are charging for sweets including:

- giving out a survey to customers who have purchased sweets from the other shops (the price was originally given to the customer for the purpose of selling the sweet which may have been given on a different date to when it is now being used or it may have been discounted at the time)
- looking at till receipts from the shop (the price is printed on the till receipt for the purpose of providing proof of purchase, not for identifying prices).

**TASK**

Which of the following are direct data sources and which are indirect data sources?

Data	Reason collected	Reason used
Names and email addresses of members of a political party	To record their membership and to be able to contact them.	To contact members by email to see if they will donate some money.
Employee attendance dates and times	To identify when employees attended work and to calculate their wages.	By a police officer to check the employee's alibi for when a crime was committed.
Flight times and prices from airline websites	To compare the prices and times for a trip to Florida.	To decide the best flight to use for a trip to Florida.
Names, ages and addresses of people	For a national census.	By a marketing company to find out which areas have the highest population of children.
Weather measurements from a weather station	To record the current weather.	To show the current temperature and rainfall on a website.



**DISCUSSION POINT**

Remember that direct data is usually used by the person that collected it and for the purpose they collected it for. However, it's also possible for a person to collect data from an indirect (secondary) source. For example, if a journalist is writing a news article and bases his story on existing news articles, then he has used indirect sources rather than interviewing the people involved in the original story.

**Advantages and disadvantages of gathering data from direct and indirect data sources**

The general rule is that data collected directly for the purpose for which it is intended is more likely to be accurate and relevant than data that is obtained from existing data (indirect source).

Direct data source	Indirect data source
The data will be relevant because what is needed has been collected.	Additional data that is not required will exist that may take time to sort through and some data that is required may not exist.
The original source is known and so can be trusted.	The original source may not be known and so it can't be assumed that it is reliable.
It can take a long time to gather original data rather than use data that already exists.	The data is immediately available
A large sample of statistical data can be difficult to collect for one-off purposes.	If statistical analysis is required, then there are more likely to be large samples available.
The data is likely to be up to date because it has been collected recently.	Data may be out of date because it was collected at a different time.
Bias can be eliminated by asking specific questions.	Original data may be biased due to its source.
The data can be collected and presented in the format required.	The data is unlikely to be in the format required, which may make extracting the data difficult.

Table 1.02 - Direct and indirect data sources.

**QUESTIONS**

The spreadsheet below is used to calculator the area of a driveway.

	A	B	C
1	<b>Area calculator</b>		
2	Length =	3	m
3	Width =	5	m
4	Area =	15	m <sup>2</sup>

Figure 1.07 - Part of a spreadsheet.

- 1 Identify one item of data that is static.
  - 2 Identify one item of data that is dynamic.
- The builder using the spreadsheet needs to know the length and width of a driveway for a customer.
- 3 Identify one direct source the builder could use to find the length and width.
  - 4 Identify one indirect source the builder could use to find the length and width.
  - 5 Give one advantage of using the direct source instead of the indirect source to find the length and width.

**Answers**

- 1 Area calculator, Length =, Width=, Area=, m, m<sup>2</sup> (not 3, 5 or 15).
- 2 15 (not 3, 5).
- 3 For example, measure it himself.
- 4 For example, accept measurements given by the customer.
- 5 For example he can rely on the measurements he has taken himself to be accurate.

**1.03 Quality of information**

The quality of information is determined by a number of attributes.

**Accuracy**

Information that is inaccurate is clearly not good enough. Data must be accurate in order to be considered of good quality. Imagine being told that you need to check in at the airport 45 minutes before the flight leaves, so you turn up at 18:10 for a 19:05 flight only to find that you were actually supposed to check in one hour early.



**EXAMPLE**

Examples of inaccurate information include:

- decimal point in the wrong place, for example £90.30 instead of £903.00 could suggest a product is cheaper than it really is
- misspelling such as 'stair' instead of 'stare', where words have completely different meanings
- misplaced characters, such as a licence plate of BW9EP3T instead of BW93PET.

**Relevance**

Information must be relevant for its purpose. Having additional information that is not required means that the user has to search through the data to find what is actually required.

**EXAMPLE**

Examples of irrelevant information include:

- being given a bus timetable when you want to catch a train
- being told the rental price of a car when you want to buy the car
- a user guide for a mobile phone that includes instructions on how to assemble a plug.

**Age**

Information must be up to date in order to be useful. Old information is likely to be out of date and therefore no longer useful. When using indirect data sources, always check when the information was produced.

**EXAMPLE**

Examples of out-of-date information include:

- the number of residents in a town based on a census from 2011, but 500 new homes have been built in the town since then
- a rugby score that has not been updated for 5 minutes in which time a try has been scored.

**Level of detail**

There needs to be the right amount of information for it to be good quality. It's possible to have either too little or too much information provided. If there is too much information, then it can be difficult to find the exact information required. If there is not enough information, then it is not possible to use it correctly.

**EXAMPLE**

A person orders a pizza. They ask for a large pepperoni to be delivered. They forgot to say what type of base they wanted and where it should be delivered to. The pizza company does not have enough information to fulfil the order.

A traveller needs to catch a train from Bhopal to Kacheguda. They phone up to find out the time of departure and arrival, but they have to listen to all the times of the stations in between before they get the arrival time at Kacheguda.

**Completeness**

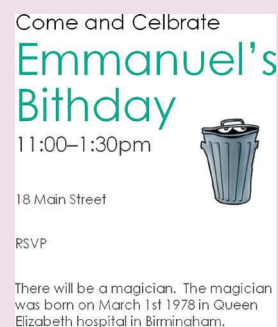
All information that is required must be provided in order for it to be of good quality. Not having all the information required means it cannot be used properly.

**EXAMPLE**

A person has booked their car in for a service over the phone. The garage owner tells them the name of the street but doesn't give the postcode or building number.

**TASK**

Look at the invitation below.



**Figure 1.08 - Wedding invitation.**

Describe how accuracy, relevance, level of detail and completeness affect the quality of information in the invitation.

**QUESTIONS**

- 1 Identify three factors that could affect the quality of information.
- 2 Describe how the age of information could affect the quality of information within a user guide for a mobile phone.

**Answers**

- 1 Accuracy, relevance, age, level of detail, completeness.
- 2 The user guide could be written for an old operating system that has since been updated on the phone.

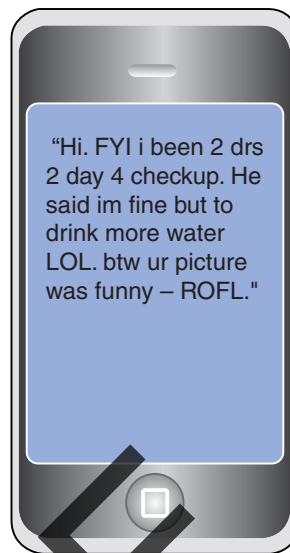


Figure 1.09 - Text message on a mobile phone.

## 1.04 Coding, encoding and encrypting data



**KEY TERMS**

**Coding:** representing data by assigning a code to it for classification or identification

**Encoding:** storing data in a specific format

**Encryption:** scrambling data so it cannot be understood without a decryption key to make it unreadable if intercepted

**HTTPS:** hypertext transfer protocol secure

**EXAMPLE**

Using numbers for international dialling codes:

- 44 = Great Britain
- 33 = France
- 49 = Germany
- 34 = Spain
- 93 = Afghanistan
- 971 = United Arab Emirates
- 81 = Japan

Using abbreviations for international vehicle registration plates:

- GB = Great Britain
- F = France
- D = Germany (Deutschland)
- E = Spain (España)
- AFG = Afghanistan
- UAE = United Arab Emirates
- J = Japan

### Coding data

You are probably very familiar with **coding** data already. When you send a text message or instant message you probably code your data. You might use codes like:

- LOL = laugh out loud
- ROFL = roll on floor laughing
- FYI = for your information
- BTW = by the way
- 2 = to
- 4 = for
- BRB = be right back

In a similar way, data stored on a computer can be coded. Coding is the process of representing data by assigning a code to it for classification or identification.

### Advantages and disadvantages of coding data

There are a number of reasons for coding data. In the examples used above, it would take a long time to use all the letters of a country to dial a telephone number and there would not be enough space on the rear of a car to display the full country name.

The advantages of coding data can be summarised as shown in Table 1.03.

<b>Presentation</b>	Data can be presented in small spaces, such as on labels or when listing large amounts of data.
<b>Storage</b>	Less storage space is required because less characters are used to store the data.
<b>Speed of input</b>	Data can be input more quickly because only a few numbers or letters need to be input instead of whole words or phrases.
<b>Processing</b>	Processors can process data more quickly because they only have to process the coded data instead of the whole word or phrase.
<b>Validation</b>	It's possible to validate data by checking it is a particular length or within a certain range or matching other rules. For example, an international vehicle registration code must be a maximum of three letters and not contain any numbers or symbols.
<b>Confidentiality</b>	Sometimes data can be coded so that it only makes sense to people who know what the codes mean.
<b>Consistency</b>	If the correct codes are used then data will be input consistently without spelling errors, making it easier to be searched or summarised.

Table 1.03 - Advantages of coding.

#### TASK

Create a coding system for days of the week in English and another language of your choice. Did you come across any problems? How did you overcome these problems?



#### DISCUSSION POINT

Can you see problems that might occur when these numbers and letters are used in codes?

- 0 or o or O
- l or L or l or i
- Z or 2
- 7 or 1

The disadvantages of coding can be summarised as shown in Table 1.04.

<b>Limited codes</b>	There may not be enough codes available to use, meaning that more letters or numbers need to be added. For example, with international vehicle registration codes, E is already used for Spain and so Egypt has to use ET and Ecuador has to use EC. Now that Egypt has taken ET, Ethiopia can't use it and so has to use ETH.
<b>Interpretation</b>	Codes may be difficult to interpret. For example, with international vehicle registration codes, somebody might look at ET and assume it is Ethiopia or look at S and assume it is Spain. It's even more difficult when considering international dialling codes that are represented by numbers.
<b>Similarity</b>	Some letters and numbers can be difficult to distinguish such as O and 0 or Z and 2, especially if handwritten. With codes, it's difficult to guess what the value might be and so mistakes can occur. At least with words it's possible to guess what the badly written letter might be.
<b>Efficiency</b>	If a user inputting codes does not know what code to use, then they will not be able to enter the data efficiently. Instead they will have to look up the code.
<b>Missing information</b>	It's possible that some information gets lost during the process of coding. For example, if devices on a network are coded as L for laptop, P for printer, D for desktop computer and M for mobile phone, then information about whether or not the mobile phone or laptop is touch screen is lost.

Table 1.04 - Disadvantages of coding.

**TASK**

Create a coding system for the names of people in your class. Show it to a friend. Can they interpret the code efficiently? Did you have any duplicate codes?

**Encoding data**

When data is **encoded**, it is stored in a specific format. Computers do not recognise text, sound and images in the same way we do. Computers use binary digits which are 1s and 0s. One means on and zero means off. A typical binary number would look like this: 11011010. Therefore, data needs to be encoded into a format which the computer understands. Codecs are programs that are used to encode data for images, audio and video. The codecs are also needed to read the data.

**Text**

Text is encoded as a number that is then represented by a binary number. A common encoding method is ASCII (American Standard Code for Information Interchange). ASCII consists of 256 codes from 0 to 255. Here is a subset of the ASCII code table:

Character	Decimal number	Binary number
@	64	01000000
A	65	01000010
B	66	01000011
C	67	01000100
a	97	01100001
b	98	01100010
c	99	01100011

Table 1.05 - ASCII code table.

There are other encoding methods, too, such as extended ASCII which allows for more characters to be used, particularly for international languages. Other encoding methods include UNICODE and EBCDIC but you don't need to learn about these.

**TASK**

Visit the website [www.ascii-code.com](http://www.ascii-code.com) and look at the ASCII code table. What sentence does the following ASCII code represent?

084 104 101 032 098 105 103 032 098  
097 100 032 119 111 108 102 033

Visit the website [www.unicode.org/charts](http://www.unicode.org/charts) or [www.unicode-table.com](http://www.unicode-table.com).

Choose a language script and then choose a short sentence (less than ten words). Encode your sentence into UNICODE. Now show the UNICODE to a friend and see if they can convert it back into the original sentence.

**Images**

Encoding is also used to store images. At the most basic level, images are encoded as bitmaps. A Microsoft Windows bitmap includes the following data when encoding an image (this is not the complete list):

- width of bitmap
- height of bitmap
- bit count which represents number of colours:
  - 1 = monochrome
  - 4 = 16 colours
  - 8 = 256 colours
  - 16 = 65536 colours
  - 24 = 16 million colours
- compression type (no compression, eight-bit run-length encoding or four-bit run-length encoding)
- horizontal resolution in pixels per metre
- vertical resolution in pixels per metre
- raster data (the actual shape of the image in pixels).

For the raster data, we will assume a monochrome image (black and white). Each bit will represent either a black or white pixel. A byte consists of eight bits and so will represent eight pixels. The encoding starts from the bottom left of the image, works to the right-hand side and then up to the next row and to the right-hand side.



Figure 1.10 - Empty raster grid.

**EXAMPLE**

The following image is 16 pixels wide and 6 pixels high. Its bit count (colour depth) is 1 so it is monochrome. There is no compression. The pixels in the image look like this:



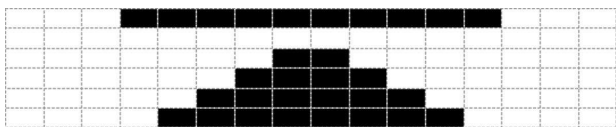


Figure 1.11 - Filled raster grid.

Imagine a 1 each time there is a black pixel and a zero each time there is a white pixel, but remember the encoding starts at the bottom left:

0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0
0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0
0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0

Figure 1.12 - Filled raster grid with binary digits.

Starting at the bottom left, the raster data to be stored would be:

```
00001111 11110000 00000111 11110000
00000011 11000000
00000001 10000000 00000000 00000000
00011111 11111000
```

This time we are going to use a bit count (colour depth) of 4 so the image can have up to 16 colours. Each byte stores two pixels. The byte consists of two bit patterns such as 1100 1010. The first bit pattern is the first pixel and the second bit pattern is the second pixel. Each bit pattern represents the colour to be used for the pixel.

The following image is 16 pixels wide and 6 pixels high. Its bit count (colour depth) is 4 so it has up to 16 colours. It uses four bits to store each pixel. There is no compression. The pixels in the image look like this:



Figure 1.13 - Filled colour raster grid.

Each byte (eight bits) represents two pixels. The colours in this example are represented by the following bit patterns:

- Black = 1111
- White = 0000
- Blue = 0001

Green = 0010

Yellow = 0011

Red = 0100

The bit pattern is added to represent the colour of each pixel:

0000	0000	0000	1111	1111	1111	1111	1111	1111	1111	1111	1111	1111	0000	0000	0000
0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000
0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000
0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000
0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000
0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000

Figure 1.14 - Filled colour raster grid with binary digits.

Starting at the bottom left, the raster encoding for the bottom two rows would be:

```
00000000 00000000 01000100 01000100 01000100
01000100 00000000 00000000
00000000 00000000 00000011 00110011 00110011
00110000 00000000 00000000
```

As the image gets bigger, it takes up a lot of storage space. Therefore a method called run-length encoding (RLE) can be used to reduce the amount of storage space that is used. This is known as compression. RLE takes consecutive values and combines them together. At a basic level, the image above has ten black pixels in a row. Instead of storing ten separate pixels, RLE would store the quantity and colour, for example ten x black or ten x 1111.

Sometimes when files are compressed, they use lossy compression, which means some of the original data is removed and the quality is reduced. For example, Figure 1.15 shows how the quality of an image can be reduced if lossy compression is used.

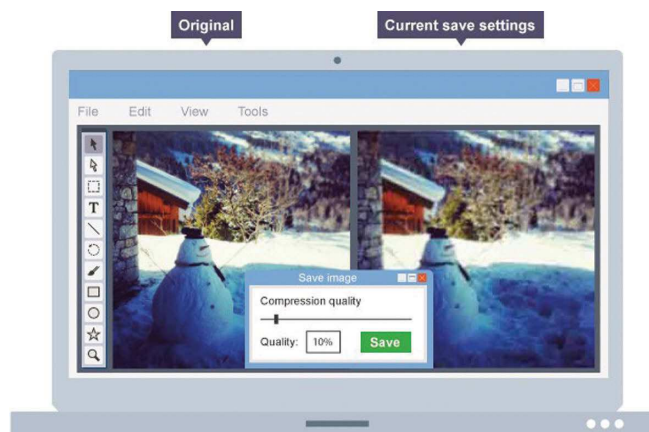


Figure 1.15 - Lossy compression.

**TASK**

Images can be encoded into a variety of different file types. Find out what the purpose of each of the following file types is:

- JPEG/JPG (Joint Photographic Experts Group)
- GIF (Graphics Interchange Format)
- PNG (Portable Network Graphics)
- SVG (Scalable Vector Graphics)

**Sound**

When sound is encoded, the sample rate, bit depth and bit rate are stored. When sound is recorded, it is converted from its original analogue format into a digital format, which is broken down into thousands of samples per second.

The sample rate, also known as the frequency, is the number of audio samples per second. The higher the sample rate, the higher the quality of the music, but also the more storage that is required. Each sample is stored as binary data. The sample rate is measured in hertz (Hz). Typically, music on a CD will use a sample rate of 44.1 kHz (kilohertz) whereas a simple telephone call would find 8 kHz sufficient.

The bit depth is the number of bits (1s and 0s) used for each sound clip. A higher bit depth will give a higher quality sound. CDs use a bit depth of 16 which means 16 bits (0s and 1s) are used. 16 bits gives 65 536 combinations of binary storage.

The bit rate is the number of bits that are processed every second. It is measured in kilobits per second (kbps). The bit rate is calculated using this calculation:

$$\text{bit rate} = \text{sample rate} \times \text{bit depth} \times \text{number of channels}$$

**EXAMPLE**

A CD sound file has a sample rate of 44.1 kHz (44 100 Hz), a bit depth of 16 bits and two channels (left and right for stereo).

$$\text{bit rate} = 44\,100 \times 16 \times 2 = 1\,411\,200 \text{ bps} = 1.4 \text{ mbps}$$

(megabits per second)

That means that 1.4 megabits are required to store every second of audio. Therefore, we can multiply the bit rate by the number of seconds to find the file size. So for a 3 minute 30 second audio file (210 seconds):

$$\text{file size (in bits)} = 1\,411\,200 \times 210 = 296\,352\,000$$

(296 megabits)

There are eight bits in a byte and we use bytes to measure storage, so the file size in bits is divided by eight:

$$\begin{aligned} \text{file size (in bytes)} &= 296\,352\,000 \div 8 = 37\,044\,000 \\ \text{megabytes} &= 37 \text{ MB (megabytes)} \end{aligned}$$

**TASK**

Calculate the file size of a sound file with the following properties:

- sample rate = 8 000 bps
- bit depth = 16
- channels = 1

The encoding method above does not compress the file and a typical storage format is WAV (Waveform Audio File Format), which is used on PCs. However, music files are often compressed to allow for efficient streaming across the internet and to enable lots of music files to be stored on a single storage medium. Compression reduces the file size.

There are two types of compression: lossy compression and lossless compression. Lossless compression reduces the file size without losing any quality, but it can only reduce the file size by about 50%.

Lossy compression reduces the file size by reducing the bit rate, which means that some of the quality is lost. Most human ears won't be able to detect the loss in quality, but an experienced musician would notice. When compressing an audio file, it is possible to choose the bit rate. A bit rate of 128 kbps is equivalent to what would be heard on the radio.

**EXAMPLE**

A CD file originally has a bit rate of 1.4 Mbps (1 411 200 bps). This is reduced to 128 kbps (128 000 bps). The original file size was 37 MB. To calculate the change in file size, divide by the original bit rate and multiply by the new bit rate.

$$\begin{aligned} \text{Compressed file size} &= 37 \div 1\,411\,200 \times 128\,000 \\ &= 3.36 \text{ MB.} \end{aligned}$$

In this example, the file size has been reduced to approximately 10% of the original size.

## Video

When video is encoded it needs to store images as well as sound. Images are stored as frames. A standard quality video would normally have 24 frames per second (fps). High definition (HD) uses 50 fps and 60 fps. The higher the number of frames per second, the more storage that is required, but the higher quality the video will be.

The size of the image is also important. A HD video will have an image size of 1920 pixels wide and 1080 pixels high. The higher the image size, the more storage that is required. Other common resolutions include:

- 7680 x 4320 (8K / ultra high definition)
- 3840 x 2160 (4K / ultra high definition)
- 1024 x 768 (Extended Graphics Array (XGA) – often used with digital projectors)
- 1280 x 720 (Wide Extended Graphics Array (WXGA) – wide screen)
- 800 x 600 (Super Video Graphics Array (SVGA))
- 480 x 320 (useful for mobile phones where the screen is smaller and internet bandwidth could be low).

The bit rate for videos combines both the audio and frames that need to be processed every second. The bit rate is the number of bits that need to be processed every second. A higher frame rate requires a higher bit rate.

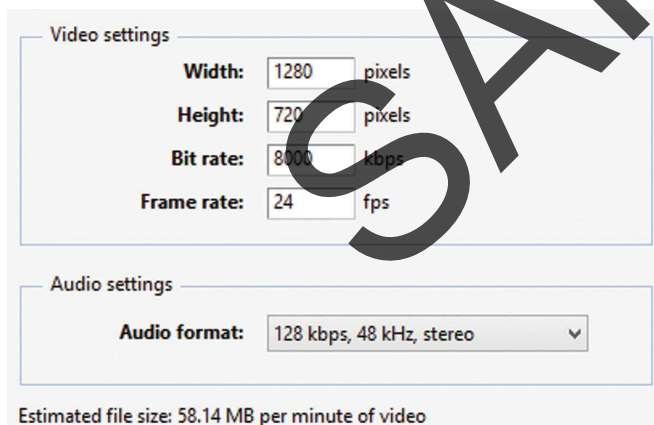


Figure 1.16 - Video and audio settings.

A one hour, eight-bit HD video with 24 fps would require 334 GB (gigabytes) of storage. This would be far too much data to download or even broadcast. Therefore, compression is required. Compression usually involves reducing the:

- resolution
- image size or
- bit rate.

These all result in lossy compression. A common lossy compression format is MP4, which is a codec created by MPEG (Moving Pictures Expert Group). There are also lossless compression methods such as digital video (DV).

### TASK

Edit a short video using video editing software. Try saving it with different resolutions, different frame rates and different audio formats. Compare the file sizes of each video file that is created.

## Advantages and disadvantages of encoding data

Data has to be encoded in order to be able to be stored and processed by a computer system. It is essential that data is encoded into binary digits (1s and 0s).

One purpose of encoding data is often to reduce the file size. The advantages of reducing the file size include:

- enabling real-time streaming of music and video over restricted bandwidth (e.g. home broadband, mobile broadband)
- reducing the time taken to download files
- websites with images, music or video downloading more quickly
- enabling more music and video to be stored on CDs, DVDs, flash memory and other storage media.

Another purpose of encoding is to enable different formats to be used. The advantages of different formats for images include:

- formats such as PNG and GIF enabling transparent backgrounds
- formats such as GIF allowing animated images
- formats such as JPG allowing a very large colour depth, meaning that the image will be of high quality
- formats such as bitmap image file (BMP) including all the original uncompressed data so that the image is of the highest possible quality and can be manipulated
- formats that are designed for specific software such as Paint Shop Pro and Photoshop containing layers which allow for various adjustments to be made to the image and being able to enable or remove those layers
- vector formats such as SVG storing the shapes rather than the pixels so that graphics can be enlarged and reduced without loss of quality due to pixilation.

The main disadvantages of encoding data is the variety of encoding methods, resulting in a large variety of file types. Problems include:

- the required codecs may not be installed and so a file cannot be saved in the desired format
- the necessary codecs must be installed, in order to open an encoded file
- not all software is capable of opening different file types
- some hardware such as music and video players only playing files encoded in certain formats (for example, a CD player may only be able to play MP3 files but not Audio Interchange File Format (AIFF) or Advanced Audio Coding (AAC) audio files)
- quality of images, sound and videos is lost when files are compressed using lossy compression
- text that has been encoded using ASCII or UNICODE needs to be decoded using the correct format when it is opened. If some international characters have been included using UNICODE and the file is opened as an ASCII file, then the international characters will not be recognised.

**EXAMPLE**

In this example the alphabet is to be shifted by three (+3) letters so that A = D, B = E and so on:

Original	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Encrypted	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

So to encrypt the word 'Hello', we would use:

$$H = K, E = H, L = O, O = R$$

which gives KHOOR.

**TASK**

Compare the file types that can be opened and saved using different graphics software packages.

**Encryption**

One specific type of encoding is **encryption**. This is when data is scrambled so that it cannot be understood. The purpose of encryption is to make the data difficult or impossible to read if it is accessed by an unauthorised user. Data can be encrypted when it is stored on disks or other storage media, or it can be encrypted when it is sent across a network such as a local area network or the internet. Accessing encrypted data legitimately is known as decryption.

**Caesar cipher**

A cipher is a secret way of writing. In other words it is a code. Ciphers are used to convert a message into an encrypted message. It is a special type of algorithm which defines the set of rules to follow to encrypt a message. Roman Emperor Julius Caesar created the Caesar cipher so that he could communicate in secret with his generals.

The Caesar cipher is sometimes known as a shift cipher because it selects replacement letters by shifting along the alphabet.



**DISCUSSION POINT**



Alamy @WADWJ

Figure 1.17 - The Enigma machine.

The Germans used encrypted messages during World War II using the Enigma machine. An operator would type in a message and then scramble it by using three to five notched wheels. The recipient would need to know the exact settings of the wheels in order to decipher the message. There were  $159 \times 10^{18}$  possible combinations.

Encrypted messages from the Lorenz cipher machine led to the development of the world's first electric programmable computer, Colossus, which helped British code breakers read encrypted German messages.

The first major encryption algorithm for computers was the Data Encryption Standard (DES) used in the 1970s. It used a 56-bit (1s and 0s) key. This offered  $70 \times 10^{15}$  possible combinations, but with the development of computers meant that it eventually became possible to 'crack' the code. This was replaced by the Advanced Encryption Standard (AES) which uses up to 256-bit keys. This can offer  $300 \times 10^{11}$  possible combinations.



**TASK**

- 1 Using the Caesar cipher +3 example above, write an encrypted message to a friend. Ask your friend to decipher it.
- 2 Choose how many letters you are going to shift by and write another encrypted message to a friend. Don't tell your friend how many letters you shifted by. Your friend should try to decipher the code by working out which letters appear most commonly.
- 3 See [www.math.uic.edu/CryptoClubProject/CCpacket.pdf](http://www.math.uic.edu/CryptoClubProject/CCpacket.pdf) to create a cipher wheel and use it to encrypt and decrypt messages.

**Symmetric encryption**

This is the oldest method of encryption. It requires both the sender and recipient to possess the secret encryption and decryption key. With symmetric encryption, the secret key needs to be sent to the recipient. This could be done at a separate time, but it still has to be transmitted whether by post or over the internet and it could be intercepted.



Figure 1.18 - Symmetric encryption.

**Asymmetric encryption**

Asymmetric encryption is also known as public-key cryptography. Asymmetric encryption overcomes the problem of symmetric encryption keys being intercepted by using a pair of keys. This will include a public key which is available to anybody wanting to send data, and a private key that is known only to the recipient. They key is the algorithm required to encrypt and decrypt the data.

The process works like this:



Figure 1.19 - Asymmetric encryption.

In this example, Bob sends a message to Alice. Bob encrypts the message using Alice's public key. Alice receives the encrypted message and decrypts it using her private key.

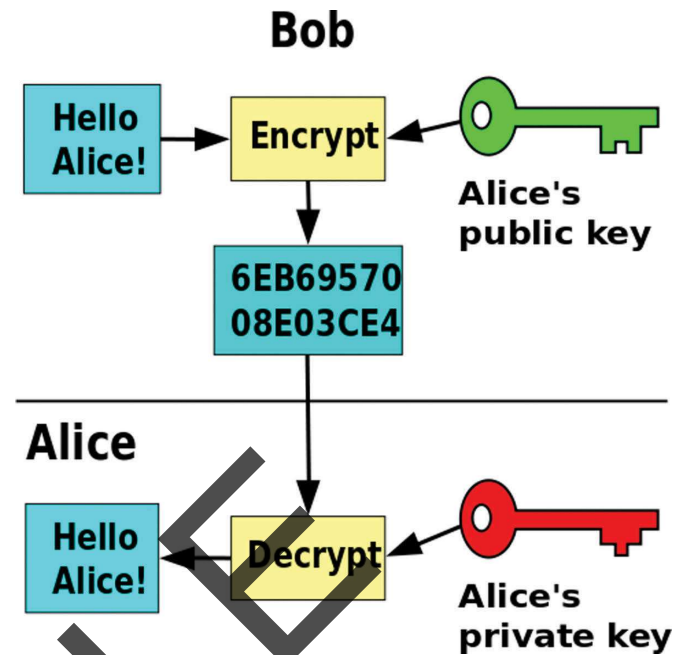


Figure 1.20 - Asymmetric encryption example.

This method requires a lot more processing than symmetric encryption and so it takes longer to decrypt the data.

In order to find a public key, digital certificates are required which identify the user or server and provide the public key. A digital certificate is unique to each user or server. A digital certificate usually includes:

- organisation name
- organisation that issued the certificate
- user's email address
- user's country
- user's public key.

When encrypted data is required by a recipient, the computer will request the digital certificate from the sender. The public key can be found within the digital certificate.

**Hard disk**

Disk encryption will encrypt every single bit of data stored on a disk. This is different to encrypting single files. In order to access any file on the disk, the encryption key will be required. This type of encryption is not limited to disks and can be used on other storage media such as backup tapes and Universal Serial Bus (USB) flash memory. It is particularly important that USB flash memory and backup tapes are encrypted because these are portable storage media and so are susceptible to being lost or stolen. If the whole medium is encrypted, then anybody trying

to access the data will not be able to understand it. The data is usually accessed by entering a password or using a fingerprint to unlock the encryption.

**HTTPS**

Normal web pages that are not encrypted are fetched and transmitted using Hypertext Transfer Protocol (HTTP). Anybody who intercepts web pages or data being sent over HTTP would be able to read the contents of the web page or the data. This is particularly a problem when sending sensitive data, such as credit card information or usernames and passwords.

Hypertext Transfer Protocol Secure (**HTTPS**) is the encryption standard used for secure web pages. It uses Secure Sockets Layer (SSL) or Transport Layer Security (TLS) to encrypt and decrypt pages and information sent and received by web users. This is the encryption method that is used by banks when a user logs onto online banking. A secure web page can be spotted by its address beginning with https:// and some browsers display a small padlock.

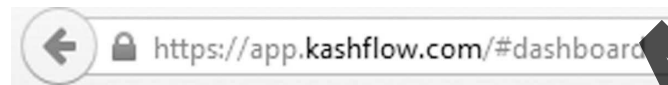


Figure 1.21 - The 's' after 'http' and the padlock indicate that this is a secure website.

When a browser requests a secure page, it will check the digital certificate to ensure that it is trusted, valid and that the certificate is related to the site from which it is coming. The browser then uses the public key to encrypt a new symmetric key that is sent to the web server. The browser and web server can then communicate using a symmetric encryption key, which is much faster than asymmetric encryption.

**EXAMPLE**

The web browser requests the certificate from the web server.

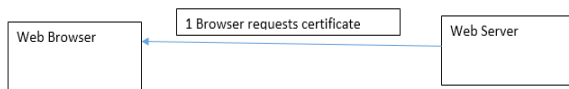


Figure 1.22 - Asymmetric cryptography.

The web browser then uses the web server's public key to encrypt a new symmetric key and sends that encrypted symmetric key to the web server. The web server uses its own private key to decrypt the new symmetric key.

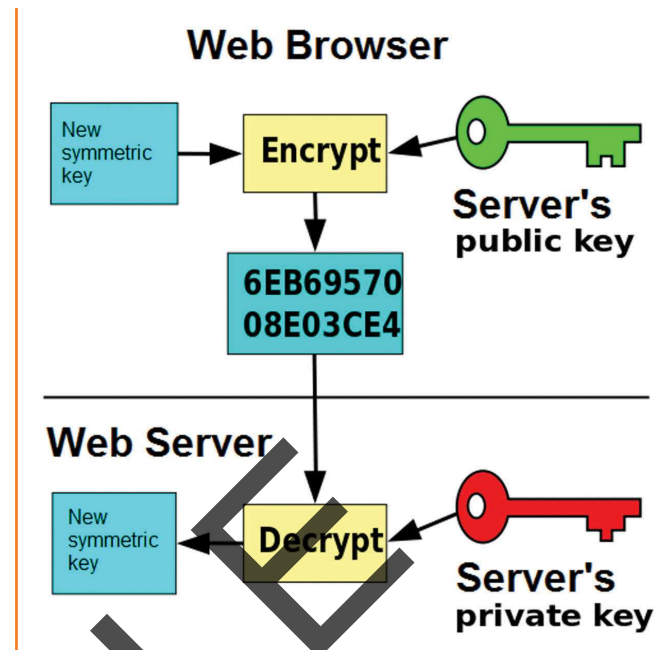


Figure 1.23 - Secure website identification.

The browser and web server now communicate using the same symmetric key.

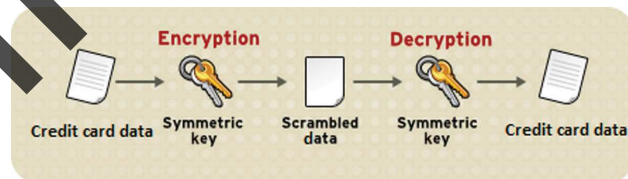


Figure 1.24 - Symmetric encryption.

**Email**

Email encryption uses asymmetric encryption. This means that recipients of emails must have the private key that matches the public key used to encrypt the original email. In order for this to work, both the sender and recipient need to send each other a digitally signed message that will add the person's digital certificate to the contact for that person. Encrypting an email will also encrypt any attachments.

**How encryption protects data**

Encryption only scrambles the data so that if it is found, it cannot be understood. It does not stop the data from being intercepted, stolen or lost. However, with strong 256-bit AES encryption it is virtually impossible for somebody to decrypt the data and so it is effectively protected from prying eyes.



**DISCUSSION POINT**

Most Wi-Fi access points and Wi-Fi routers use encryption. This serves two purposes. The first is to only allow people who know the 'key' (usually a password) to access the network, so that any unauthorised users cannot gain access. The second is to encrypt the data, so that it cannot be understood by somebody 'snooping' on the Wi-Fi network.

Wi-Fi networks are particularly susceptible to 'snooping' because no wires are required to connect to the network. It is possible to sit in a car outside somebody's house and see the Wi-Fi network. The 'key' stops that person from accessing the network and also stops that person from understanding the data that is moving around the network.

Did you know that if you access a public Wi-Fi hotspot that is 'open' and therefore not encrypted that anybody with the right software can see what you are sending over the network including your emails? This applies to laptops, tablets and mobile phones or any other device using public Wi-Fi.

**Validation**

**Validation** takes place when data is input into a computer system. The purpose is to ensure the data is sensible and conforms to defined rules. A railway season ticket will have an expiry date. The season ticket is valid until it expires. Once it expires it is invalid. The rule here is that the date the season ticket is used must be before its expiry date.

When data is validated, if it conforms to the rules then it will be accepted. If it does not conform to the rules, then it will be rejected and an error message will be presented. Validation does not ensure that data is correct.

**TASK**

Create a flow chart to describe the process of validation. You should include the following:

- Start
- End
- Input of data
- Error message
- Data accepted
- Data rejected
- Validation decision

There are a variety of different validation checks that can be used to check whether data is acceptable. These different checks are the different types of rules that are used.

**QUESTIONS**

- 1 Give two reasons for encoding data.
- 2 Explain why encoding is necessary for text.
- 3 Identify two factors that affect the size of a sound file.
- 4 Decipher the following text that has been encrypted using the Caesar cipher with a shift of +4.

AIPP HSRI

- 1 Saving storage space, enables validation, can help with presentation of lists, speed of input, speed of processing, confidentiality.
- 2 Computers can only understand on and off (1 and 0), so text needs to be encoded into a standard character set that uses numbers represented by binary.
- 3 Sample rate, bit depth, number of channels.
- 4 WELL DONE.

**Presence check**

A presence check is used to ensure that data is entered. If data is entered, then it is accepted. If data is not entered, then the user will be presented with an error message asking them to enter data.

**1.05 Checking the accuracy of data**



**KEY TERMS**

**Validation:** the process of checking data matches acceptable rules

**Verification:** ensuring data entered into the system matches the original source

**Proof reading:** checking information manually

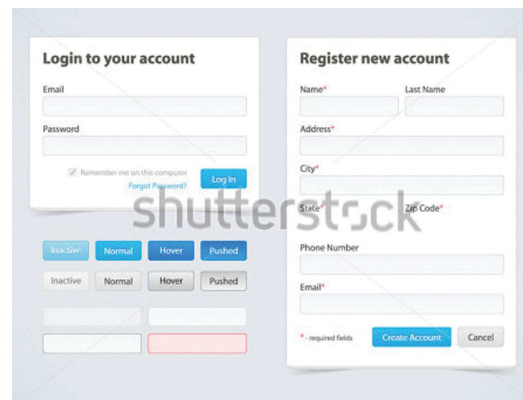


Figure 1.25 - Required data on a website.

**EXAMPLE**

When filling in a 'contact us' form on a website, it is essential that an email address is entered. The following would be valid if only a presence check is carried out:

- a
- a@b
- a@bc.d
- @
- 372823

Notice that none of these are correct but they pass the rule that data must be present.

**Range check**

A range check ensures that data is within a defined range. A limit check has a single boundary. This could be the highest possible value or the lowest possible value. A range check includes two boundaries, which would be the lower boundary and the upper boundary. The following symbols are used when comparing with a boundary:

- > greater than
- < less than
- >= greater than or equal to
- <= less than or equal to

**EXAMPLE**

An opinion poll is taken and asks for the respondent's age. The respondents have to be at least 18 years old. The lower boundary is 18. There is no upper boundary, so this is a limit check. This could be written as:

Age >= 18

Letters representing grades for an exam are entered. Only the letters A–E are valid grades. The grade must be less than F. The upper boundary is E. There is no lower boundary, so this is a limit check. This could be written as:

Grade < F

The number of students in a class must be between 5 and 28. The lower boundary is 5 and the upper boundary is 28, so this is a range check. This could be written as:

Number of students >= 5 and <= 28

Data that is within the boundaries is valid. Data that is outside the boundaries is invalid. Data that is valid and within the boundaries is not necessarily correct. A grade of C could be entered when a grade A should have been entered. C is valid but incorrect.

**Type check**

A type check ensures that data must be of a defined data type.

**EXAMPLE**

- If an age is entered, it must be an integer.
- If a grade is entered, it must be text with no numbers.
- If a price is entered, it must be numerical.
- If a date of birth is entered, it must be a date.

Data that is of the correct data type is valid. Data that is valid and of the correct data type is not necessarily correct. A date of birth of 28/12/2087 could be entered. The date is valid because it is a date data type, but it is clearly incorrect.

**Length check**

A length check ensures data is of a defined length or within a range of lengths.

**EXAMPLE**

- A password must be at least six characters long.
- A grade must be exactly one character long.
- A product code must be at least four characters and no more than six characters.

Data that is of the allowed length is valid. Data that is valid and of the allowed length is not necessarily correct. A grade of H is valid because it has one character, but it is not valid.

**Format check**

A format check ensures data matches a defined format. It is sometimes known as a pattern check because data has to follow a pattern.



**EXAMPLE**

An email address must include an @ symbol preceded by at least one character and followed by other characters. The following data would be valid:

- john@bldef.co
- a@b.dek
- fdc@jb

A student ID must be four numbers followed by two letters. The following data would be valid:

- 3827BD
- 1111AA

Data that matches the pattern is valid. Data that is valid and of the defined format is not necessarily correct. An email address of fdc@jb meets the rules above but is clearly incorrect.

**Lookup check**

A lookup check tests to see if data exists in a list. It is similar to referential integrity in Chapter 9, but uses a list defined within the validation rule.

**EXAMPLE**

When asking a user for their gender, they can respond with 'Male' or 'Female'. A lookup validation rule would check to see that the values are within this list. Students taking a qualification could be issued grades of pass, merit and distinction. When inputting the data, a validation rule could check that only 'X', 'P', 'M' or 'D' are entered ('X' would be for fail).

**Consistency check**

A consistency check compares data in one field with data in another field that already exists within a record, to see whether both are consistent with each other.

**EXAMPLE**

When entering the gender of 'M' or 'F', a consistency check will prevent 'F' from being entered if the title is 'Mr' and will prevent 'M' from being entered if the title is 'Mrs' or 'Miss'.

When entering data about dispatching products, it would not be possible to mark an item as being dispatched until after it has been packaged.

**Check digit**

A check digit is a number (or letter) that is added to the end of an identification number being input. It is a form of redundancy check because the check digit is redundant (not needed for the identification number, but just used for validation). When the identification number is first created, an algorithm (a series of calculations) is performed on it to generate a check digit. When the identification number is input, the same algorithm is performed on it. The result of the algorithm should match the check digit. If it matches, then the data is valid. If it does not match then the data is invalid.

Original identification number = 20392  
Algorithm is performed on 20392  
Check digit = 4

Data including check digit = 203924

**Valid Example**

Identification number including check digit is entered into the computer: 203924  
Algorithm is performed on 20392  
Result of algorithm = 4

Result of algorithm (4) is compared with check digit that was entered (4).  
They match.  
Data is valid.

**Invalid Example**

Identification number including check digit is entered into the computer: 205924  
Algorithm is performed on 20592  
Result of algorithm = 7

Result of algorithm(7) is compared with check digit that was entered (4).  
They do not match.  
Data is invalid.

Figure 1.26 - Valid and invalid check digit calculations.

**EXAMPLE**

Figure 1.27 - Unique Product Code check digit.

The Unique Product Code (UPC) check digit is used with 13 digit barcodes. It is the last digit shown on a barcode. The algorithm for calculating the check digit is:

- 1 Add all the digits in even numbered positions together.
- 2 Multiply the result (1) above by 3.
- 3 Add all the digits in odd numbered positions together.
- 4 Add results (2) and (3) together.
- 5 Divide the result (4) above by 10.
- 6 Calculate the remainder (modulo 10) of result (5) above.
- 7 Subtract (6) above from 10.

**Valid example**

In this example, the International Standard Book Number (ISBN) is 978095734041-1 where the last 1 is the check digit. To calculate the check digit, the following algorithm is performed on the ISBN (excluding check digit):

- 1 Add all the digits in even numbered positions together (978095734041) :  $7 + 0 + 5 + 3 + 0 + 1 = 16$ .
- 2 Multiply result (1) above by 3:  $16 \times 3 = 48$ .
- 3 Add all the digits in odd numbered positions together (978095734041) :  $9 + 8 + 9 + 7 + 4 + 4 = 41$ .
- 4 Add results (2) and (3) together:  $48 + 41 = 89$ .
- 5 Divide the result (4) above by 10:  $89 \div 10 = 8.9$ .
- 6 Calculate the remainder (modulo 10) of result (5) above:  $89 \text{ MOD } 10 = 9$ .

- 7 Subtract (6) above from 10:  $10 - 9 = 1$ .

The result of the algorithm is 1.

**Invalid example**

In this example, the ISBN has been entered incorrectly as two numbers have been transposed (7 and 3) accidentally: 978095374041-1.

- 1 Add all the digits in even numbered positions together (978095374041) :  $7 + 0 + 5 + 3 + 0 + 1 = 16$ .
- 2 Multiply result (1) above by 3:  $41 \times 3 = 60$ .
- 3 Add all the digits in odd numbered positions together (978095374041) :  $9 + 8 + 9 + 3 + 4 + 4 = 37$ .
- 4 Add results (2) and (3) together:  $60 + 37 = 97$ .
- 5 Divide the result (4) above by 10:  $97 \div 10 = 9.7$ .
- 6 Calculate the remainder (modulo 10) of result (5) above:  $97 \text{ MOD } 10 = 7$ .
- 7 Subtract (6) above from 10:  $10 - 7 = 3$ .

The result of the algorithm is 3. The result 3 is compared with the check digit of 1 that was entered. They do not match. The ISBN entered is invalid.

**TASK**

Use the website [www.upcdatabase.com/checkdigit.asp](http://www.upcdatabase.com/checkdigit.asp) to generate check digits for product codes.

**DISCUSSION POINT**

The usual algorithm for UPCs is to multiply the odd digits by 3 rather than the even digits. It is only for 13 character barcodes that the even digits are multiplied by 3.

Find out how to calculate a check digit for 10 digit barcodes.

**Verification**

**Verification** is the process of checking that the data entered into the computer system matches the original source.

**Visual checking**

A method of verification can be for the user to visually check that the data entered matches the original source. This can be done by reading the data displayed on screen and comparing it with the original data. If the data matches,

then it has passed the verification process. If it does not match, then it has failed the verification process and needs to be re-entered. Visual checking does not ensure that the data entered is correct. If the original data is wrong, then the verification process may still pass. For example, if the intended data is ABCD but ABC is on the source document, then ABC will be entered into the computer and verified, but it should have been ABCD in the first place.

### Double data entry

Another method of verification is to input data into the computer system twice. The two items of data are compared by the computer system and if they match, then they are verified. If there are any differences, then one of the inputs must have been incorrect.

#### EXAMPLE

When changing a password, most systems will ask the user to enter the new password twice. This is because it is critical that the password is entered correctly in order that the user can gain access to the system in the future. If the new passwords match, then the password will be changed. If the new passwords don't match, then one of the passwords must have been entered incorrectly.

It is still possible to pass double entry verification and for the data to be incorrect. If the data is entered incorrectly twice, then the two values may match. For example, if the CAPS key is left on by mistake then both entries would match.

### The need for both validation and verification

As you will have seen in the two sections above, it is possible to enter valid data that is still incorrect. It is also possible to verify incorrect data. By using both validation and verification, the chances of entering incorrect data are reduced. If data that is incorrect passes a validation check, then the verification check is likely to spot the error.

#### EXAMPLE

The validation rule is that a person's gender must be a single letter. N is entered. This passes the validation check but is clearly incorrect. When verified using double entry, the user enters N first

followed by M the second time. The verification process has identified the error.

However, it is still possible that the user could enter N twice and both the validation and verification processes would fail.

### Proof reading

**Proof reading** is the process of checking information. For example, when this book was written it was checked for spelling errors, grammar errors, formatting and accuracy. Proof reading can take place for a document or when data is input. When proof reading a document, it is best to have a proof reader who is different from the original author of the document, as they will be able to check the work objectively and seek out errors. However, it is also possible for the original author to proof read their own document, but they may not notice some of their own errors. When data is input, it is usually proof read by the person inputting the data.

#### TASK

Check the information below by proof reading it.

I were walking along the road yesterday wen I spotted a dog without a lead I called the dog but it did not respond. the dog ran away;

#### QUESTIONS

- 1 Describe the purpose of verification.
- 2 Identify three methods of validation.
- 3 Explain using examples why validation and verification do not ensure data is correct.</NL>

#### Answers

- 1 To ensure data that is input matches the source data.
- 2 Presence, range, type, length, format, check digit, lookup.
- 3 Data can match the rules but still be incorrect (e.g. date of birth of 29/1/13 matches the rule of a data type but it may be that the date should have been 29/1/03).

Data can be checked against the source, but the source maybe incorrect (e.g. the name 'Siohan' is written on the original source and is visually checked to have been entered as 'Siohan', but the actual spelling should have been 'Siobhan').

## 1.06 Summary

Information has context and meaning so a person knows what it means. The quality of information can be affected by the accuracy, relevance, age, level of detail and completeness of the information. Proof reading is the process of checking information.

Data are raw numbers, letters, symbols, sounds or images without meaning. Knowledge allows data to be interpreted and is based on rules and facts. Static data does not normally change. Dynamic data updates as a result of the source data changing. Data collected from a direct data source (primary source) must be used for the same purpose for which it was collected. Data collected

from an indirect source (secondary source) already existed for another purpose.

Coding is the process of representing data by assigning a code to it for classification or identification. Encoding is the process of storing data in a specific format.

Encryption is when data is scrambled so that it cannot be understood.

Validation ensures that data is sensible and allowed. Validation checks include a presence check, range check, type check, length check, format check and check digit. Verification is the process of checking data has been transferred correctly. Verification can be done visually or by double data entry.

### Review questions

- 1 Give an example of data. [1]
- 2 Describe the term knowledge. [1]  
A news and sports website includes both static and dynamic data.
- 3a Compare static and dynamic data. [4]
- 3b Identify and describe two factors that affect the quality of information on the website. [4]
- 3c Give an example of coded data related to the news and sports website. [1]  
The website streams news and sports videos.
- 3d Explain why the video is compressed when it is encoded. [2]
- 3e Identify three factors that will affect the size of a video file. [1]
- 3f Identify and describe two factors that affect the file size of images on the website. [4]  
Users can pay for premium services on the website using their credit card.
- 4a Explain why the website uses https at the beginning of the website address instead of http. [4]
- 4b Describe symmetric encryption. [1]  
The journalists working for the website encrypt their emails.
- 4c Describe how asymmetric encryption is used with emails. [2]  
When the users subscribe to premium features, they have to choose a password.
- 4d Describe how verification can be used when entering the password. [1]  
There is a form on the website that can be used to submit news stories. When data is entered onto the form, it is validated.
- 5a Describe the purpose of validation. [1]
- 5b Using an example related to submitting a news story, identify and describe one method of validation. [3]



**SAMPLE**

# Chapter 6

## Using networks

2

### Learning objectives

*By the end of this chapter, you will:*

- understand the advantages and disadvantages of networking computers
- be able to describe the characteristics and purpose of different types of networks, such as LAN, WAN, VPN, client-server, peer-to-peer, intranets, extranets and the internet
- be able to define the terms internet and World Wide Web and discuss the difference between the two
- be able to discuss the benefits and drawbacks of using the internet
- understand how the internet is used for communication
- be able to discuss the advantages and disadvantages of mobile networks
- understand how to set up a video and a web conference and describe how they use networks
- be able to describe how data is transmitted in a video conference
- understand the impact of video conferencing on our lives

## 6.01 Networking computers



### KEY TERMS

**Network:** two or more computers that are linked together

**Network architecture:** the design of a network

A single computer on its own is known as a stand-alone computer. Up until the 1990s, computers in homes and organisations were largely stand-alone devices. Today, many computers are connected together.



### KEY CONCEPT

When we connect stand-alone computers together we create a **network**.

The first instance of computers being linked together was in 1969 with the Advanced Research Projects Agency Network (ARPANET). Universities connected computers together with the intention of being able to communicate and share resources. This network was the predecessor to the internet.

Networking computers has a number of advantages and disadvantages.

The advantages of networking computers will mostly outweigh the disadvantages. This is why we chose to network computers together.



### KEY CONCEPT

The way a network is designed is called its **architecture**. A network's architecture can be described by its geographical and physical nature, and by how data are stored and accessed on it.

### Local area networks (LANs) and wide area networks (WANs)



### KEY TERMS

**Local area network (LAN):** a network that covers a small geographical area

**Wide area network (WAN):** a network that covers a wide geographical area

We can describe a network by its geographical and physical architecture. This creates two categories of network:

- local area network (LAN)
- wide area network (WAN).

Advantages	Disadvantages
Users can share their data with other users on the network.	Setting up a larger network can be expensive, especially if the network requires cabling, several servers, switches and wireless access points. The staffing costs to maintain a network can also be expensive.
Users can access their data from different computers on the network.	Networks can have security risks. Many users' data may be accessed if an unauthorised user gains access to a networked file server.
File servers can be used to store data in a central location. This makes it easier to keep the data secure and to create a back-up copy.	Viruses could easily spread from one networked computer to another, if a network is not sufficiently protected.
Application servers can be used to store software in a central location. This can often reduce software costs as a site licence could be purchased, if available. A software site licence is generally cheaper than individual licences. This also means that software updates would need to be run on the server only.	A whole network could fail if one element of the network fails. This will depend on how the network is structured.
Communications can be sent from one computer to another.	A high amount of network traffic may cause a network to run slowly and cause frustrating delays.
Peripherals can be shared. This means that a number of users can share resources in a company or home, rather than requiring their own.	
An internet connection can be shared, allowing each user on the network to access the internet.	

Table 6.01 - Advantages and disadvantages of a network.

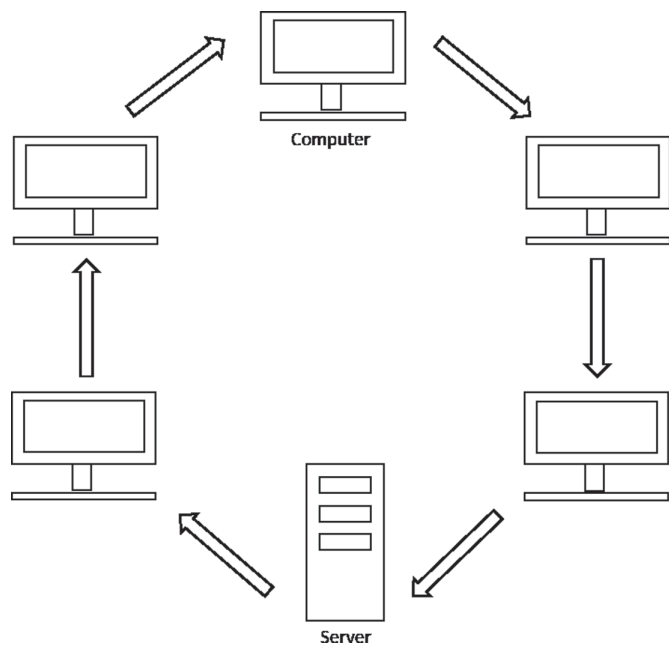


Figure 6.01 - A basic structure for a LAN.

A LAN is a network that is restricted to a small geographical area. LANs are usually found in homes, schools, universities and small businesses.

A WAN is a network that covers a larger geographical area. A WAN may be contained within a country or can spread across several countries. Organisations such as banks, police forces and hospitals use WANs to share data. The internet is a global WAN. Several LANs can be connected

together to form a WAN. A business that has several different branches often operate their networks this way.

The use of a LAN and a WAN can be compared for several different factors.

The terms LAN and WAN describe the geographical architecture of a network and how it is physically structured. They do not describe how data is stored and accessed. The two architectures for storing and accessing data on a network are:

- client-server
- peer-to-peer.

### Client-server and peer-to-peer networks

**KEY TERMS**

**Server:** a dedicated computer that manages resources and services from a central location

**Client:** a computer that is connected to a server

**Peer:** a computer in a network that is not connected to a server and does not control another computer

A client-server network uses a dedicated computer to store data, manage resources and control user access. This computer is known as a **server**. The server acts as a central point on the network that other computers

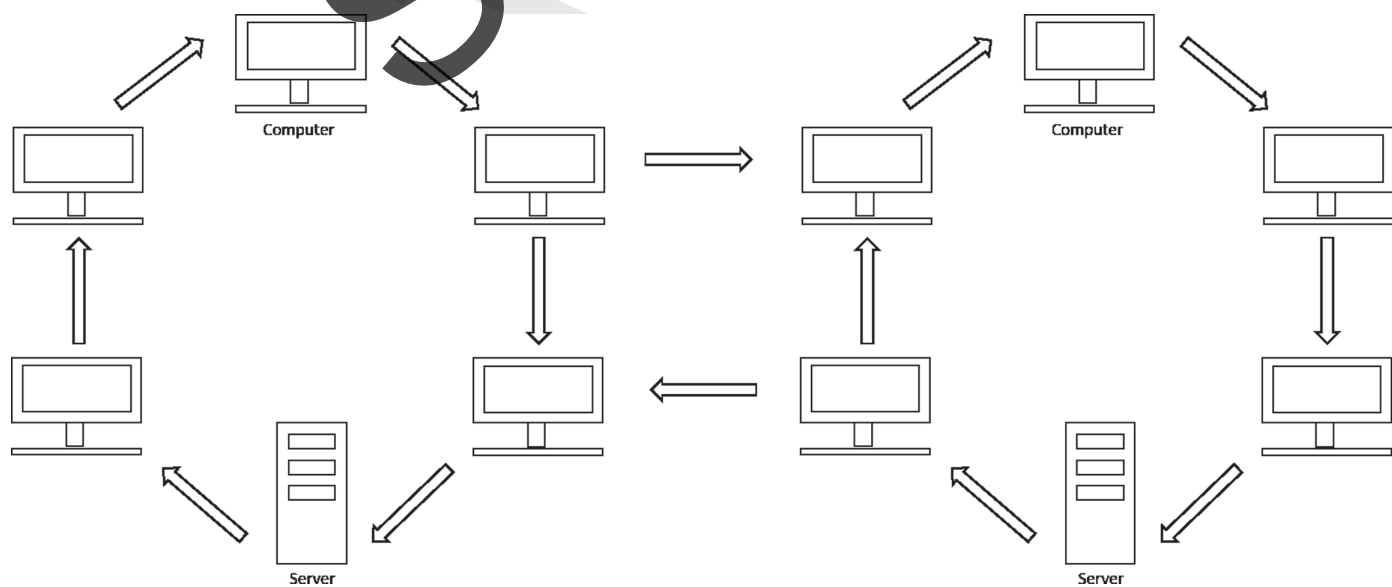


Figure 6.02 - A basic structure for a WAN.

Factors	Local area network	Wide area network
Data transfer rates	Can have up to 1 Gb per second.	Often restricted to less than 200 Mb per second.
Data transmission errors	Fewer errors as data packets are sent over short distances	Greater chance of error as packets are sent over longer distances
Connection method	Restricted to within a small location. Copper wire, fibre optic cable or radio waves are mostly used.	Expand to a wider location. Copper wire, fibre optic cable and microwaves are mostly used. They can often use public telephone network systems, leased lines, transmission towers and satellite communication.
Security	Comparatively high as security can be easier to implement over a smaller number of devices and connections.	Comparatively low as security is difficult to implement across a higher number of devices and connections.
Ownership	Mostly owned by individuals or individual organisations.	Elements of the network often use communication infrastructures owned by others.

Table 6.02 - Comparison of a LAN and a WAN.

connect to. A computer that connects to the server to access the data and resources it manages is known as a **client**.

Network servers handle a variety of functions, including:

- storing a database of usernames and passwords to control user access
- data storage
- security
- assigning levels of access that determine what resources a user may access
- monitoring network traffic
- hosting of shared applications
- managing an internet connection
- scheduling and running back-ups
- email services
- printing jobs on network printers
- domain name services.

A client-server network will usually have at least one server and several clients. A network may have more than one server, each one dedicated to handling a specific function. For example, one server may handle user access, while another may handle data storage. This often occurs in larger networks where a server has to service many clients at the same time. Spreading the load across several servers helps to keep the network running efficiently.

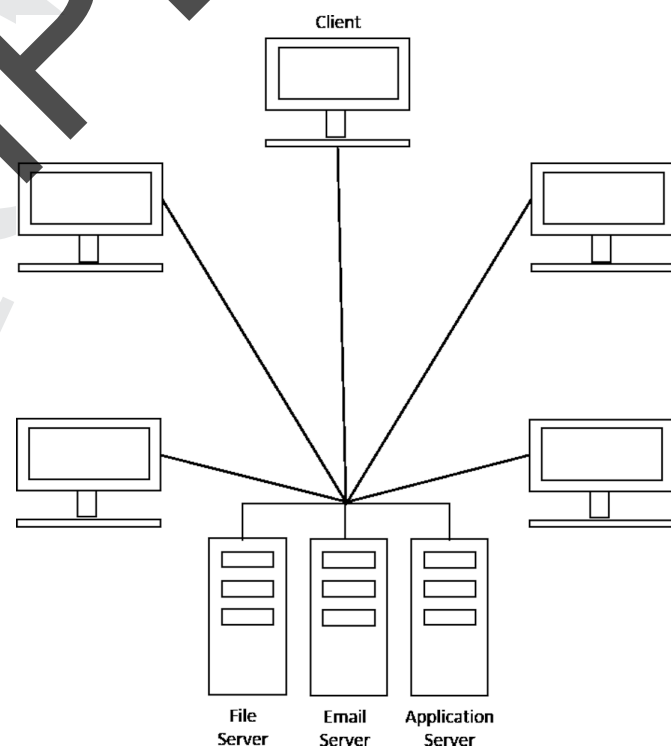


Figure 6.03 - Example of a client-server network.

A peer-to-peer network has no central server to manage the network. Each computer on the network shares its data and resources with all the others. No computer has control over another. Each computer is known as a **peer**.



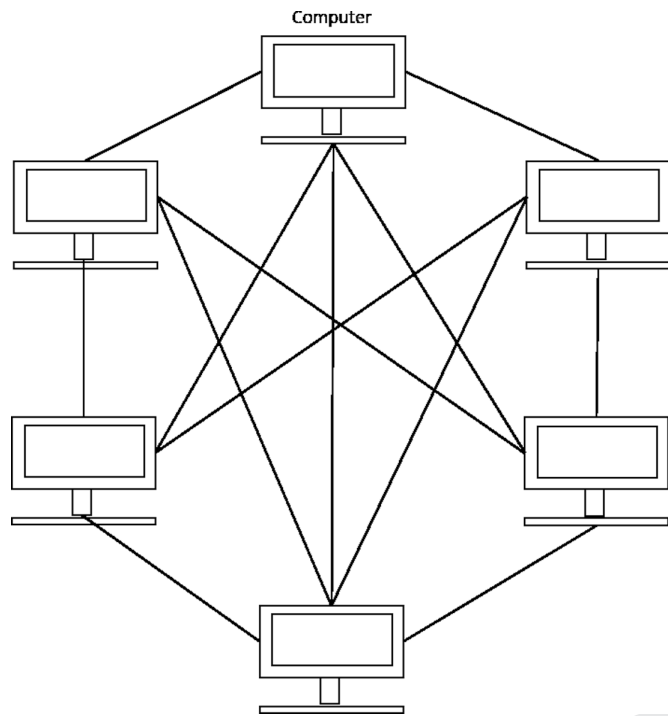


Figure 6.04 - Example of a peer-to-peer network.

Advantages of a client-server network	Disadvantages of a client-server network
Centralised database of users, usernames and passwords provides login security	Expensive to set up and maintain
User data is kept private	Failure of a central server may make the whole network fail
Levels of access can be applied to resources	Requires specialist technical staff to maintain
Users do not have to worry about security as a network manager maintains the network	
Users do not have to make back-ups as back-ups are centralised and handled by the network manager	
On larger networks several servers can be used for different functions to maintain network efficiency	

Table 6.04 - Advantages and disadvantages of a client-server network.

On a peer-to-peer network, users store data on their own computer, but may allow other users access to it. Alternatively, they may choose to keep their data private. Individual peers may provide resources that others do not. For example, a peer may have a printer attached to it and other peers may print to that printer.

Table 6.03 shows a comparison of the architecture of a client-server and a peer-to-peer network.

Client-server network	Peer-to-peer network
A dedicated computer controls the network	No individual computer controls the network
Centralised database of users, usernames and passwords	No centralised database
Centralised data storage	Data is stored on individual computers
Centralised back-ups	Relies on users remembering to back-up individual computers
Suited to large networks with many computers	Suited to small networks with few computers

Table 6.03 - Comparison of client-server and peer-to-peer networks.

Each type of architecture has its advantages and disadvantages. Peer-to-peer networks are often cheaper to install and maintain than client-server networks. Client-server networks can be expensive to purchase and often require trained staff to maintain them.

Each type of network is used in many different way, for example:

- Banks use client-server architecture to provide online banking services. When a device (such as a personal computer (PC), tablet or smartphone) connects to the bank through the internet, the device becomes a client. The bank's servers handle the client and provide access to online banking facilities.
- Companies use servers to host websites. Client computers access the web pages they contain.

Advantages of a peer-to-peer network	Disadvantages of a peer-to-peer network
Cheap to set up and maintain. Does not require expensive servers.	No centralised database of users, usernames and passwords, making it harder to maintain security.
Failure of an individual peer only removes access to the resources and data which that peer handles. The rest of the network remains unaffected.	Users have to worry about making back-ups as no centralised back-up service exists.
Does not require specialist technical staff to maintain.	The more computers there are on a peer-to-peer network, the slower it will run.

Table 6.05 - Advantages and disadvantages of a peer-to-peer network.

- On-demand television and film providers have servers that stream content to client devices.
- Many peer-to-peer networks exist over the internet that allow users to share music, images and videos. Each peer has some content and allows other users access to it.

## Virtual private networks (VPN)

### KEY TERMS

**Tunnelling protocol:** a tunnel between two points on a network that is governed by a set of rules

**Encryption:** a method of modifying data to make it unreadable if intercepted

Companies or organisations may want to allow users to access a network, often a LAN, from a remote location. They will need to use external resources to create this connection. The external resources will not be owned or maintained by the owner of the LAN, which can make access to the LAN network much less secure. This can be of great concern to the owner of the LAN and the users who want to remotely access the LAN.

The issue of security can be overcome by implementing a virtual private network (VPN).

### KEY CONCEPT

A VPN is a method of allowing a remote computer to connect securely to a private network using external resources.

Once connected to the VPN, the remote computer then has secure access to the LAN's resources, just as if it were directly connected to the LAN.

Security is implemented through the use of **tunnelling protocols** and **encryption**. For example, the secure shell (SSH) tunnel protocol uses encryption to securely transmit unencrypted data packets across an unsafe network.

To connect to a VPN, a user remotely logs on to the network. Once the user's identity has been authenticated, the VPN's tunnelling protocol sets up an encrypted link to the user's computer. Data can then be transmitted securely.

### KEY CONCEPT

VPNs are often used by organisations to provide LAN access to travelling employees, employees who telework and users that want to access LAN resources at home. Users who want to securely access data on a home computer when they are away can also use a VPN.

There are several advantages and disadvantages for an organisation using VPNs.

### TASK

Find out what kind of network you have in your school or home. What is the structure of it and how does it store data?

Does your school or home network have any servers? If it does, what are they used for?

## Intranets and extranets

An intranet is a privately owned network that uses internet technologies, for example web pages, for internal use within an organisation. Web pages hold organisations' information such as news, events, historical data and commonly used documents. Email and instant messaging can also be used for internal communication. Training videos or corporate messages can be streamed and watched by employees.

Advantages of a VPN	Disadvantages of a VPN
VPNs maintain a high level of security for data transmission over the internet.	Performance of the VPN may be outside the control of the organisation. It is dependent on the external resources that are used.
VPNs use external resources. Organisations that use VPNs can save money as they do not need to purchase or implement secure, dedicated connections, called leased lines, for users who want to access the network.	A certain amount of technical expertise is required to set up and maintain VPNs. The cost of this expertise must be factored in when deciding whether to implement a VPN.
The cost of maintaining the external resources is covered by someone else.	
Adding extra clients costs very little as the network is virtual.	

Table 6.06 - Advantages and disadvantages of a VPN.

Intranets usually use client-server architecture. They often use several servers. A web server hosts the intranet's web pages, while an email server handles email communication. Another server will handle user accounts and access.

Many organisations now use an intranet. For example, schools use a virtual learning environment (VLE) to provide teaching and learning resources to students. The VLE is an intranet.

**KEY CONCEPT**

Intranets are used to increase efficiency because all information is stored centrally and can be easily accessed.

Additionally, unlike when browsing the internet, only information relevant to the organisation is available. An intranet is for internal use only. This increases security and confidentiality as only users within the organisation can access information on the intranet.

There may be times when an organisation will want to allow users that are currently external to the organisation

to access information on its intranet. It can do this by using an extranet. An extranet is an intranet that has been opened up to allow external user access. Extranets use VPNs to allow an external user to securely access the information on them.

Extranets usually only allow certain information to be accessed. For example:

- A school's VLE may allow parents to access their child's marks, attendance data and reports.
- A doctor's surgery may be able to access a local hospital's extranet to make an appointment for a patient.
- A company may provide tracking information on deliveries it is making to customers.

Access to an extranet is normally available through the organisation's website. A user logs in using an email address and password that will notify the extranet that the user is an authorised external user.

**6.02 The internet**

**KEY TERMS**

**Packet switching:** data that is broken down and transmitted in small sections

**TCP/IP:** a communication protocol used by the internet

**Internet service provider (ISP):** A company that provides access to the internet

**KEY CONCEPT**

The internet is a global network made up of interconnected networks and individual computers. This global network allows users to access online facilities such as electronic communication, information sharing, data storage, banking, video and music streaming and social networking.

The beginnings of the internet were quite different from how it is in the present day. In the 1950s and 1960s, the United States government formed several organisations whose purpose was to develop new technology. One organisation was the Advanced Research Projects Agency (ARPA). As projects developed, employees at ARPA became frustrated at the lack of computing power available to them. At the time, only a handful of powerful computers were available across the country. These were



**Figure 6.05 - The internet connected to the world on a global scale.**

often at locations far away from the researchers who needed them. The invention of **packet switching** allowed the creation of computer networks. Taking advantage of this communication breakthrough, ARPA created ARPANET, a network of computers that could share data and computing power.

The 1970s saw the introduction of the protocol, **TCP/IP**. This allowed separate networks to connect to each other, forming the basis for the internet. The term 'internet' originates from 'inter-network', meaning a network that was formed from the joining of networks.

The 1990s saw a huge increase in the size of the internet and its purpose, as many commercial, educational and home users took advantage of the facilities available online. Today, the number of users and devices on the internet runs into the billions. It has become a backbone, the purpose of which is to allow communication and sharing of data on a global scale.

### Characteristics of the internet

Any device connected to the internet forms part of the internet, meaning the number of devices on the network runs into billions. To connect to the internet we need to use an internet service provider (ISP). ISPs use public telephone systems, underwater cables, microwave transmissions via satellites and radio waves to provide connections.

Data and information is held and distributed by content providers. A content provider is an organisation, group or company that offers content on web pages, for example text, images, sound and video, or web-based services, for example streaming media, gaming, shopping and electronic communication.

Although many of the computers connected to the internet can act as a server, most act as clients. The majority of data is held on web servers that are accessed by individual computers.

The widespread nature of the internet means that it has no centralised control. No one company or organisation runs or has authority over the internet, although various independent organisations control aspects of it:

- The Internet Corporation for Assigned Names and Numbers (ICANN) regulates domain names and their suffixes, such as .com, .co.uk and .eu.
- The Internet Engineering Task Force (IETF) devises and administers internet standards, such as protocols that govern communication.
- The World Wide Web Consortium (W3C) attempts to enforce agreement from organisations to adopt new standards.
- Telecommunications companies own the infrastructure that the internet uses for communication.

#### KEY CONCEPT

The lack of authority means that the information the internet carries is of mixed value. Information can be useful and correct, it can also be inaccurate, out of date or completely incorrect. Some information is deliberately designed to mislead.

The internet is constantly changing and evolving. For example:

- Every day newly bought devices connect to it.
- The internet is expanding geographically as remote regions are connected through the installation of cables or through satellite communication.
- The data it holds constantly changes as organisations add, edit and delete data from their websites, and social networking users add and remove data.
- Newer technologies allow the range of facilities to expand. For example, the introduction of broadband connections has encouraged the growth of internet-based, on-demand television services.

The huge scale and size of the internet means there are many benefits and drawbacks for organisations, businesses and individuals with using it.



#### DISCUSSION POINT

What would your life be like if you did not have the internet? What would be different? Would it be better or worse?

Benefits of using the internet	Drawbacks of using the internet
Information. A vast range of information is available on virtually any topic.	Incorrect information. Anyone can publish information on the internet. This means that some of the information available is inaccurate, out of date, incomplete or simply incorrect.
Global audience. The internet is worldwide, therefore contributors have a global audience.	Unsavory behaviour. One serious problem with the internet is that it allows people to interact anonymously with other people. Users are unsure who they are really communicating with. This means that people can make up fake identities in order to cyberbully, intimidate, steal and commit abuse.
Immediate transmission of information. Users can send messages immediately across the world. Postings on news sites and social media can be immediately seen by an audience.	Hacking. Some users may use the internet to attempt to gain access to other users' computers. This is often in an attempt to steal personal data or to crash websites. Protection against hacking can be expensive to an organisation.
Immediate availability of services. Users can quickly access services such as online banking, book medical appointments, pay bills and book tickets for transport and for leisure.	Viruses. An organisation or individuals will need to protect their computer against viruses. A virus is malicious software installed on a computer without consent or knowledge. Removal of a virus can be a difficult job.
International ecommerce. Users have access to online shops 24 hours a day and products from all over the world. Companies can also market products to consumers nationally and internationally.	Identity theft through phishing and pharming. Phishing uses fake emails pretending to be from an organisation. Phishing emails deceive users into giving their personal data by tricking them into following a link to a fake website. The website asks them to enter personal details. Once gathered, this data is used for criminal activities. Pharming works in a similar way. When users attempt to visit a known website, they are instead directed to a fake website that looks just like the real one. The cost of identity theft to individuals and businesses can be enormous.
Entertainment. Products such as software, music, films and books can be purchased and immediately downloaded. Users can play online games, or join in multiplayer games over the internet.	Unsuitable material. The free nature of the internet means that unsuitable content is present, such as pornography, graphic violence, promotion of terrorism, torture and physical and written abuse. Social networking allows anyone to publicly abuse friends, colleagues, celebrity figures and people in the news, often with little consequence to the offender.
Free education. Many free online tutorials and educational resources are available to help people learn new skills or knowledge.	Spamming. The ease of communication means that anyone can generate and transmit unwanted emails to millions of users. It is estimated that 69% of emails sent in 2014 were spam.
Social networking. The internet allows families, friends, organisations and businesses to regularly communicate, share and engage in social activities.	Online addiction. Some people have difficulty separating virtual environments from reality. Others find virtual worlds, such as those in massively multiplayer online role-playing games (MMORPGs) more alluring and comforting than the real world. They feel safer and happier in these environments than they do in real life. This can lead to social isolation and insecurity.
Up-to-date news. News stories on the internet can be read as soon as they are published.	

Table 6.07 - Benefits and Drawback of using the Internet.



Benefits of using the internet	Drawbacks of using the internet
Reduction in costs to businesses and organisations. Marketing over the internet, especially through social media, can be extremely cheap compared with traditional advertising and promotional methods.	
Feedback. The internet has allowed companies and organisations to easily gain feedback on products and services through rating systems, comments, online surveys and 'likes' on social media.	
Big data. The nature of the internet allows vast amounts of data to be collected by organisations. User behaviour on websites can be monitored, such as the type of products a person buys, or searches for information on. Social networking sites collect data on what news items are trending, and what users like and dislike. Apps on smartphones collect data about habits, and behaviour. All of this collected data can be analysed by researchers or sold to businesses for marketing purposes.	
New communication applications. The internet allows communication applications that previously were not possible. For example, home security cameras can be monitored while the householder is away from home. Central heating can be remotely switched on or off from a smartphone app.	

Table 6.07 - (cont.)

### 6.03 The World Wide Web

#### KEY TERMS

**Hyperlink:** a link that can be clicked to locate to another place in a document, or a different document entirely

**Web browser:** a software application for retrieving and presenting information on the World Wide Web

The term 'internet' describes the physical, global network of computers. The internet does not actually contain any information. It is a physical infrastructure.

The World Wide Web is an information system comprising documents, images, videos and sounds. It uses the internet as its base. Users access this information through websites and web pages, which are accessed through the use of a **web browser**.

The concept of the World Wide Web was developed by Tim Berners-Lee in 1989 when he worked for CERN.

Berners-Lee proposed a system for a more efficient communication system for CERN that allowed documents on an information system to be accessed via **hyperlinks**. He realised that any document could be accessed by a hyperlink, whether it is a text document, an image or a sound clip. In 1990 Berners-Lee and Robert Cailliau published a proposal to build a World Wide Web of hypertext-linked documents that could be viewed by a browser. Documents would be stored on servers, with users' computers acting as clients. From this proposal the World Wide Web was born and in 1991 it became available to the public.

Today the World Wide Web is the basis for, among other things, national and international information sharing, banking, shopping, on-demand television broadcasting and social networking. It has billions of users, including businesses, organisations, governments, schools, colleges, universities and domestic users.

## 6.04 Communication using the internet

The internet is for communication. It allows users to communicate in several ways, including:

- instant messaging (IM)
- voice over IP (VOIP)
- news services.

Instant messaging (IM) is a form of communication that allows real-time text transmission over a network, often the internet. Depending on the type of IM service, messages are either transmitted once a message is composed and sent by the user, or transmitted character by character as the user types. Some services allow files to be transferred and conversations to be saved.

IM is now a feature of many social networking sites, allowing users to chat online instead of posting messages on their profile page. Many organisations see IM as a valuable tool for customer interaction, for example support helplines. This can be especially useful where a support technician needs to give instructions to a customer. The customer can read (and reread) and follow the instructions in their own time.

Users of IM often see it as an equivalent or alternative to text messaging used on smartphones. They make use of text speak (where common words and expressions are abbreviated), for example LOL (laughing out loud), and emoticons to reduce typing and speed up replies.

An example of an IM service is the private messaging system available on Facebook. Users can send messages

to each other privately as well as posting on each other's profile.

Voice over IP (VOIP) technologies allow users to have voice communications, the same as if they were using a telephone system. Instead of using an analogue telephone transmission, the user's voice is recorded by the computer and transmitted digitally to the recipient. The VOIP system can work internally across a network, or externally making use of the internet to carry the data transmission.

VOIP has advantages over traditional analogue telephone systems. When the system is run over an organisation's LAN, calls within the organisation will be free. Extra VOIP phones can be added to a network, simply by connecting a new phone to a network point or by adding it to a wireless network. However, if a network is experiencing heavy traffic, voices can become garbled, distorted or missing altogether.

An example of a VOIP service is Skype. Users can make calls using the Internet using the Skype application.

News services are organisations that collect and broadcast news stories for subscribers. Users can sign up for news alerts using a smartphone application. Some services allow news stories to be filtered out so that only news that covers topics a user is interested in is sent to their smartphone. Others broadcast headline news to subscribers.

An example of a news service is Google News. Users can subscribe and receive news stories on a daily basis, often in the categories of their choice.

All of these services are available because of packet switching technology.

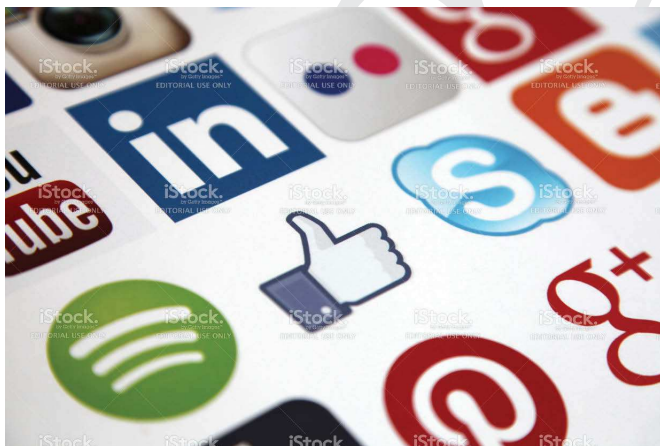


Figure 6.06 - Many social networking services now offer an instant messaging feature.

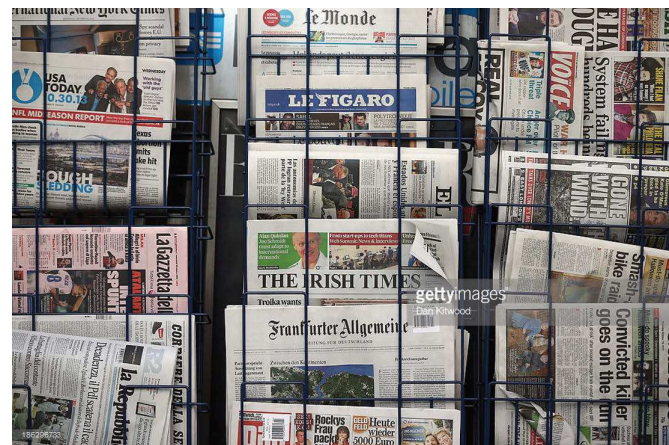


Figure 6.07 - Users can subscribe to news services to receive news stories in the category of their choice.

## 6.05 Mobile networks



### KEY TERM

**Cell:** the geographical area covered by a radio transmitter



### KEY CONCEPT

A mobile network is a wireless WAN that uses radio to transmit and receive communications from portable devices, such as smartphones. Mobile networks now cover a considerable amount of the inhabited areas of world.

Mobile networks are broken down into small areas called **cells**. At the heart of each cell is a radio base station that transmits and receives messages. The base stations connect to public telecommunications services allowing access to the internet.

Cells vary in size:

- picocells cover an area of less than 200 metres
- microcells cover an area up to 2 kilometres
- macrocells cover larger regions

There have been several generations of mobile networks, each providing faster access speeds and greater reliability:

- 1G networks. These were the first generation mobile networks that used analogue signals. These networks were largely limited to voice and text message communications.
- 2G networks. The second generation mobile networks switched from analogue to digital transmission, improving signal quality. 2G networks were able to connect with each other, allowing a phone to use other networks.
- 3G networks. Third generation networks increased data transmission speeds up to 2 Mbps, allowing internet access, video transmission and online gaming.
- 4G networks. Fourth generation networks are the current, latest generation of mobile network technology. In theory, they allow data transmission speeds of up to 1 Gbps, allowing greater use of video streaming facilities.

Mobile networks have their advantages and disadvantages.

Advantages of mobile networks	Disadvantages of mobile networks
Mobile networks have enabled users to communicate with others and access the internet while on the move, often through the use of smartphones.	Quality of reception can vary and can be poor towards the edge of a cell, leading to interrupted or delayed transmissions.
Breaking the network down into cells allows for lower power radio transmitters to be used, bringing energy and cost savings.	Frequency use must be carefully planned to avoid interference between cells.
There are only so many radio frequencies that can be used for mobile communications. These frequencies can be split among neighbouring cells, and reused in more distant cells, thereby increasing the number of communications that can take place at the same time.	
Larger geographical areas can be covered than by using a single transmitter. Even high power transmitters are limited in range. By using several, low power transmitters a wider area can be covered.	
The use of multiple transmitters means the network is more robust. Failure of one base station only affects one cell, leaving other areas of the network unaffected.	

Table 6.08 - Advantages and disadvantages of mobile networks.

Mobile networks have grown hugely in popularity, with many users now accessing the internet through 3G or 4G connections. The rapid growth in use of social media networks has led to many people checking profiles and sending posts frequently during the day, often via a mobile network.

### TASK

Find out what further developments are in progress for mobile networks.



Figure 6.08 - People can use mobile devices, such as tablets, for video conferencing calls.

## 6.06 Video and web conferencing

One form of communication that is slowly gaining in popularity is video conferencing. Video conferencing allows users to see and hear each other as they talk. Web conferencing is a similar communication medium. However, whereas video conferencing allows two-way communication, web conferencing traditionally allowed one-way communication. A user would broadcast a video to many watching users, for example a company delivering a training session or a university lecturer presenting a lecture. However, web conferencing technology has expanded to allow two-way communication.

### How to set up a video conference

Video conferencing requires certain hardware and software:

- a desktop computer or laptop
- a webcam to record the user's image
- a microphone to record the user's voice
- a monitor to display the other participant's image
- speakers to produce the other participant's voice
- network (and internet) access to connect to the other participant
- video conferencing software to conduct and control the video conference.

Many modern devices such as laptops, smartphones and tablets already have this hardware. Smartphones and tablets are often supplied with video conferencing software installed.

Before a user can participate in a video conference, they must first register with a video conferencing service. The service will have software that the user needs to install. With the software running and the device connected

to a network or the internet, the user selects another registered user and initiates a conference call. The other participant can accept or reject the call. Similarly, the user can accept or reject video calls from other registered users.

### How to set up a web conference

Setting up a web conference is a similar process to a video conference. Web conference software is available, but quite often the same software that is used for video conferencing can also be used to set up a web conference.

Before initiating the web conference, the user sends an invitation to interested participants, stating the time and date of the conference. The invitation often contains an access code that allows the interested participant to log on to the conference, along with a link that specifies the network address at which the conference can be accessed. Using access codes helps keep unwanted users from participating.

At the appointed time and date, users follow the link and log on using the access code. They can now watch and participate in the conference.

### Using networks in video and web conferencing

The wide-scale use of video conferencing and web conferencing would not be as possible without the use of networks. The use of packet switching enables video and sound transmissions to be broken down into packets, just like any other type of data, to be transmitted. The resulting packets can be easily and swiftly transmitted across networks and the internet.



Video and web conferences held internally within organisations will make use of that organisation's LAN. The LAN may be wired or wireless. To connect externally, the conferences will also need to make use of a WAN or the internet.

Video and web conferencing requires a reasonable data transmission speed. The higher the required quality of video and sound, the higher the transmission speed that is required. As a result, for video or web conference over a WAN, one of four types of connection is usually used:

- Integrated Services Digital Network (ISDN)
- Asynchronous Digital Subscriber Lines (ADSL)
- Synchronous Digital Subscriber Lines (SDSL)
- 3G/4G mobile networks.

Integrated Services Digital Network (ISDN) is a networking technology that allows data to be transmitted digitally over analogue copper wire. This is usually with transmission speeds of up to 64 Kbps per channel. Two channels could also be used in parallel to send data at up to 128 Kbps. ISDN was popular in the 1990s and early 2000s but has largely been replaced by much faster ADSL and SDSL networks. However, ISDN is still used in areas where ADSL and SDSL technology has yet to be implemented.

Asynchronous Digital Subscriber Line (ADSL) technology, also known as broadband, is now the most common type of connection to the internet. It has typical transmission speeds of around 100 Mbps. ADSL is used in homes and businesses to deliver reliable, high-speed connections. Asynchronous means that download and upload speeds are different. This means that ADSL works by unevenly splitting the balance between download and upload speeds. Most homes and businesses will download more data from the internet than upload to it. As a result, ADSL download speeds are usually far greater than upload speeds.

Synchronous Digital Subscriber Line (SDSL) technology uses the same data transmission technology as ADSL. However, this time the download and upload transmission rates are balanced and equal. SDSL is used by organisations that upload and download equivalent amounts of data.

Mobile devices such as smartphones make use of 3G/4G technology to allow users to use video conferencing when travelling, or when access to a LAN is not possible.

For a video conference to successfully take place, the video and audio signals must be transmitted over a

network. Video signals require high volumes of data transmission. The higher the resolution of the video, the more data that needs transmission, and the greater the traffic that is placed on a network. High traffic can lead to disruption or interruption of the video conference.

To overcome this problem, video conferencing uses software called codecs. Codec is an abbreviation of the words 'coder' and 'decoder'. At the sender's end, the 'coder' element of the codec takes the video and audio signals and compresses them to reduce their file size. Reducing the file size means less data needs to be transmitted, reducing network traffic and the possibility of disruption to the conference. At the receiver's end, the 'decoder' decompresses the video and audio signals so that they can be played.

The same codec must be used at both the source and the destination of the data transmission. Different codecs compress and decompress in different ways, therefore are incompatible with each other. As a result, several codec standards exist. The H.264 codec is one of the most common and is used not only for video conferencing, but also with DVDs and YouTube videos. Other codecs for video conferencing include H.261 and H.263+.

## The impact of video conferencing

### KEY TERMS

**Digital divide:** the separation between those that have access to technologies and the internet and those that do not

**Teleworking:** working from home using technologies to keep in contact with an employer

### KEY CONCEPT

Improvements in internet and communication technologies have led to affordable high-speed connections at home and at work. These high-speed connections can usually handle video transmissions. The introduction of 4G mobile networks meant that mobile devices are also capable of high-speed connections.

Additionally, many computing devices now feature inbuilt video conferencing equipment. Smartphone, laptops and tablets generally have microphones, speakers and cameras, along with the network connectivity needed for video conferencing. Free software allows cheap video conferencing.



As a result, an increasing number of users have now taken to using video conferencing as a method of communication and are enjoying the benefits it brings. Those users who live a distance from friends and relatives have especially taken to this technology as it allows them to see, as well as hear, their loved ones. Users with hearing and speaking impairments enjoy video conferencing as it enables them to communicate using sign language.

A growing example of a video conferencing software that people use is Apple's FaceTime. Users that have an iPhone can use the inbuilt FaceTime software application to communicate with each other using video conferencing

However, the general public has yet to fully embrace this technology. Many users feel self-conscious and uncomfortable in front of a camera. Even though connection speeds have improved greatly, video conferencing does require high-speed data transmission. High levels of network traffic can more frequently interfere with, degrade or interrupt video transmissions. This leads users to return to more reliable voice-only methods.

It could be argued that video conferencing is helping to widen the **digital divide**. Many older users are wary of technology. They are afraid of the complexity of video conferencing, opting for more traditional voice communication methods that are familiar and comfortable.

Video conferencing has not yet directly impacted on law. However, legislation has been implemented in several countries that allows a witness to give testimony via video conferencing. This process is reserved for witnesses who may be psychologically upset by having to attend, especially those that may have to face someone who has subjected them to physical assault or abuse.

In some countries, such as the United States, video conferencing allows defendants to make initial court appearances without having to leave jail. This saves travel expense and increases safety to the public as suspects remain in custody.

In the UK, several police forces have implemented video conferencing as a way of conducting interviews quicker. An interviewer may be based at one police station, but can interview suspects and witnesses at other stations, without having to travel to each site.

Video conferencing has had an impact on education and the way that students learn. Rather than listening to a teacher, or reading a book or web page, video conferencing is used to engage students. Video



Figure 6.09 - Video conferencing is used in some courts to allow witnesses to testify without being in the courtroom.

conferencing allows students to participate in conversations and share ideas with fellow students at schools and universities in other parts of the world. This can lead to unexplored perspectives about subjects being shared and argued in live debates, resulting in wider, more considered points of view, especially when the students are from differing cultures.

Speaking to industry and subject experts often engages and interests students in a topic. Video conferencing allows experts to give talks and lectures without having to travel to a school. Sometimes a talk can be given to more than one school at a time, allowing many students to benefit from the experience. This is especially useful to schools in remote, isolated areas where travel would be difficult or expensive.

Universities make heavy use of video conferencing. The technology allows lecturers to teach courses at several universities, while being based at one. Sharing of such expertise is becoming more common, especially in subjects where finding a lecturer for a course has been difficult. Additionally, video conferencing allows lecturers to hold a lecture for students while away on research or at conferences.

Medicine and medical care are areas where video conferencing continues to be of growing usefulness. Consultants and physicians in different hospitals are able to discuss a patient's case and help to offer a diagnosis. Extra peripherals such as ultrasound imaging devices and video endoscopes can be connected and images from them shared in the video conference, helping consultants at other locations to help assess a patient's condition. This has the dual benefit of reducing travelling costs for consultants and increasing efficiency as consultants can see more cases in any given time period.



**Figure 6.10 - Medical students using video conferencing software to see and interact with a live, surgical operation.**

In some countries medical practices have introduced video conferencing to help doctors and nurses keep in touch with patients who have physical impairments, and for those in remote areas where travelling may be difficult or dangerous. In such situations video conferencing can bring comfort to a patient, especially the elderly, who might not otherwise have much human contact.

Sometimes the best way to do business is in a face-to-face meeting. Email and telephone provide useful media for arranging and making business deals, but when it comes to agreements of great value, people tend to prefer to do business in person. Meeting someone in person helps to promote goodwill, friendship and trust.

However, face-to-face meetings are not always possible. For example, time constraints or cost of travel may make them impossible. Video conferencing has provided a possible solution to these situations. It allows people to have face-to-face meetings without having to leave the office. This saves travel expenses and it also allows for more efficient use of an employee's time, as hours are spent working and not travelling.

Video conferencing has also helped to facilitate the growth in **teleworking**. Employees who work regularly from home can sometimes feel cut off from their employer. Teleworkers that use video conferencing tend to feel happier and less isolated from their employer and co-workers. Companies that have regional offices are finding that video conferencing brings unexpected benefits. Many companies that regularly use video conferencing to communicate with their regional offices find that employees in those offices feel more involved with the organisation, increasing employee satisfaction and productivity.

Companies are also using video conferencing for staff training. Some companies have several offices, and some, like banks, have many branches. A trainer at a company's head office can deliver a training programme via a video conference to staff at different offices or branches simultaneously.

However, video conferencing can also cause anxiety and stress among some workers. Some people are less confident in front of a camera. As a result, a video conference made by an employee with senior personnel can cause more stress than a face-to-face meeting.

One area in which video conferencing has seen a large-scale uptake is in the media sector. Video conferencing allows journalists to deliver live, on-the-spot reports in situations where a live broadcast might not otherwise be possible.

Another aspect to consider is press video conferencing. With businesses and organisations taking advantage of the global audience the internet permits, it is increasingly becoming normal to hold international press conferences. These conferences are often held to launch new products or to make statements about political affairs. Press video conferencing allows journalists to participate without actually having to attend.



#### DISCUSSION POINT

Can you think any other ways that video conferencing could be used in our lives?

## 6.07 Summary

Stand-alone computers that are connected together create what is called a network. There are different types of networks that can be categorised by their geographical and physical structure or their method for data storage.

A LAN is a network that is geographically small in area. They are usually found in homes, schools and small businesses.

A WAN is a network that is geographically large in area. The largest and most common WAN is the internet. The internet is the network infrastructure, whereas the World Wide Web is the web pages of content we can view using the internet, an ISP and a web browser. The internet is used for communication in various different ways. These include IM, VOIP, news services, video conferencing

and web conferencing. Video conferencing has had an impact on many areas of society, such as personal communications, court cases, business, education, medicine and media. These impacts have been both positive and negative.

A client-server network has servers that run the network centrally. Each computer that is part of the network is

called a client. A peer-to-peer network does not have servers. No single computer has control over another in this network.

A VPN is a method of allowing a remote computer to connect securely to a private network using external resources.

### Review Questions

1. Describe the difference between a LAN and a WAN. [4]
2. Explain what is meant by an intranet and how it could be used in an organisation. [4]
3. Explain the advantages and disadvantages of a VPN. [4]
4. Describe the difference between the internet and the World Wide Web. [2]
5. Identify four pieces of hardware that could be used when video conferencing. [4]
6. Discuss the impact of video conferencing on education. [6]

# Glossary

SAMPLE

**Actuator:** this is a type of motor that controls a mechanism or system

**Alignment:** positioning text so that it is in line, for example on the left, right or centre

**Alpha testing:** initial testing of the software by a limited group of people

**Analogue:** this is the smooth stream of data that we process on a daily basis

**Animation:** a series of images are played one after each other to simulate movement

**Array:** a data structure that can store multiple items under one identifier

**Artificial intelligence:** computer systems that perform tasks that normally require human intelligence

**Assignment:** providing a value to a variable

**Attribute:** a category of information within an entity

**Bandwidth:** the range of frequencies available for a communication method which determines the transmission rate; the amount of data that can be transmitted at one time

**Beta testing:** a sample of users test a pre-release version of the software

**Biometrics:** the measurement of a person's physical characteristics; unique physical characteristic of a person that can be used by a computer for identification purposes

**Bitmap:** an image made up of small squares, called pixels. Each individual pixel can only be one colour

**Bit rate:** the number of bits that can be transferred per second

**Black box testing:** testing of inputs and outputs to a system or part of a system with no consideration for the workings of the system

**Bluetooth:** a short range wireless communication standard

**Bridge:** connects two LANs

**Broadband:** a method of faster data transmission that can carry several channels of data at once

**CAD:** computer-aided design

**CAM:** computer-aided manufacture

**Cell:** a rectangle within a spreadsheet where data can be positioned; the geographical area covered by a radio transmitter

**Chaining:** combining together instructions

**Circuit switching:** a single communication path is opened for the duration of data transmission

**Client:** a computer that is connected to a server

**Closure:** completion of a project

**Cloud computing:** using remote servers hosted on the internet to store data

**Coding:** representing data by assigning a code to it for classification or identification

**Collaboration:** working together

**Compiler:** translates high-level programming language into an executable file in machine code

**Composite key:** two or more fields that form the primary key

**Conception:** start of a project

**Condition:** a statement that can be evaluated as true or false, for example  $5 > 6$

**Confidential:** needs to be kept secret

**Construct:** a control structure, such as a loop or a conditional statement

**Critical path:** the tasks which must be completed on time for a project to complete on time

**Custom-written:** software that is written especially to meet the requirements of a client

**Data:** raw numbers, letters, symbols, sounds or images without meaning

**Database:** a structured method of storing data

**Database management system:** software used to manage a database

**Data dictionary:** metadata (information) about the database

**Data type:** the type of data being stored, for example a number, string (text)

**Demographic:** a particular section of a population

**Design specification:** illustration of how the system will look, what the data structures will be and how the system will work

**Device:** a hardware component of a computer system formed of electronic components

**DFD:** data flow diagram which shows how data moves around a system