

Introducing computing and IT



About this free course

This free course is an adapted extract from the Open University courses TM111 *Introducing Computing & Information Technology 1* <http://www.open.ac.uk/courses/modules/tm111> and TM112 *Introducing Computing & Information Technology 2* <http://www.open.ac.uk/courses/modules/tm112>.

This version of the content may include video, images and interactive content that may not be optimised for your device.

You can experience this free course as it was originally designed on OpenLearn, the home of free learning from The Open University:

<https://www.open.edu/openlearn/science-maths-technology/computing-ict/introducing-computing-and-it/content-section-0>

There you'll also be able to track your progress via your activity record, which you can use to demonstrate your learning.

Copyright © 2019 The Open University

Intellectual property

Unless otherwise stated, this resource is released under the terms of the Creative Commons Licence v4.0 http://creativecommons.org/licenses/by-nc-sa/4.0/deed.en_GB. Within that The Open University interprets this licence in the following way:

www.open.edu/openlearn/about-openlearn/frequently-asked-questions-on-openlearn. Copyright and rights falling outside the terms of the Creative Commons Licence are retained or controlled by The Open University. Please read the full text before using any of the content.

We believe the primary barrier to accessing high-quality educational experiences is cost, which is why we aim to publish as much free content as possible under an open licence. If it proves difficult to release content under our preferred Creative Commons licence (e.g. because we can't afford or gain the clearances or find suitable alternatives), we will still release the materials for free under a personal end-user licence.

This is because the learning experience will always be the same high quality offering and that should always be seen as positive – even if at times the licensing is different to Creative Commons.

When using the content you must attribute us (The Open University) (the OU) and any identified author in accordance with the terms of the Creative Commons Licence.

The Acknowledgements section is used to list, amongst other things, third party (Proprietary), licensed content which is not subject to Creative Commons licensing. Proprietary content must be used (retained) intact and in context to the content at all times.

The Acknowledgements section is also used to bring to your attention any other Special Restrictions which may apply to the content. For example there may be times when the Creative Commons Non-Commercial Sharealike licence does not apply to any of the content even if owned by us (The Open University). In these instances, unless stated otherwise, the content may be used for personal and non-commercial use.

We have also identified as Proprietary other material included in the content which is not subject to Creative Commons Licence. These are OU logos, trading names and may extend to certain photographic and video images and sound recordings and any other material as may be brought to your attention.

Unauthorised use of any of the content may constitute a breach of the terms and conditions and/or intellectual property laws.

We reserve the right to alter, amend or bring to an end any terms and conditions provided here without notice.

All rights falling outside the terms of the Creative Commons licence are retained or controlled by The Open University.

Head of Intellectual Property, The Open University

Contents

Introduction	5
Learning Outcomes	6
1 Computers: from rarity to ubiquity	7
1.1 The telephone	7
1.2 The computer	8
1.3 The information society	9
1.4 Summary of Section 1	13
2 Some aspects of our information society	14
2.1 Business	14
2.2 Communities	18
2.3 Information	19
2.4 Entertainment	20
2.5 Public information and services	21
2.6 Communicating on the move	23
2.7 Summary of Section 2	24
3 Participating in a digital world	25
3.1 Netiquette: respecting other people online	25
3.2 Ethical and legal considerations	30
3.3 Copyright	31
3.4 Good academic practice	32
3.5 Summary of Section 3	32
4 Are search engines bad for your health?	33
4.1 Childhood vaccinations	34
4.2 The number of myths	36
4.3 Countering myths	36
4.4 Explicit recommendations	37
4.5 Summary of Section 4	38
5 Hard disk drives: a conversation	39
5.1 The structure of a hard disk drive	41
5.2 Storing files on a hard drive	42
5.3 Formatting a hard disk	43
5.4 Deleting data from an HDD	45
5.5 Fragmentation	47
5.6 Summary of Section 5	50
6 Solid-state drives: The conversation continued	51
6.1 Comparing SSDs and HDDs	52
6.2 How flash memory works	53
6.3 Writing data to an SSD	54
6.4 Deleting data from an SSD	55
6.5 Summary of Section 6	55

Conclusion	56
Glossary	56
References	60
Acknowledgements	61

Introduction

The profound technological, economic, political and ethical changes brought about by computing and information technology (IT) will affect every one of us. In this course we talk about *digital* technologies in relation to computing and IT. A fundamental idea is that a digital technology is any technology based on representing data as sequences of numbers. Computers use digital technologies so, in this course, the term 'digital world' refers to the influence of computers, in a wide variety of forms, on the world.

In this course you will learn a little about the historical development of computers and their role in today's society. You will then consider examples of digital technologies in the world around you and their influence on your life. Through a case study of online research into information on childhood vaccination, you will see how computer algorithms can shape our beliefs and actions. You will also be introduced to the field of digital forensics and learn how data are stored on and deleted from hard disk and solid-state drives.

This OpenLearn course is an adapted extract from the Open University courses [TM111 *Introduction to computing and information technology 1*](#) and [TM112 *Introduction to computing and information technology 2*](#).

Note: The word 'data' is plural ('datum' refers to an item of data), but is often used in the singular, e.g. 'The data is'. You will find both uses in this course, depending on the context.

Learning Outcomes

After studying this course, you should be able to:

- understand the idea of a 'digital world'
- explain the role of computing and information technology (IT) in a digital world
- understand the importance of keeping safe online
- understand how the use of online search engines can influence your beliefs and actions
- describe how data are stored on and deleted from a hard disk drive (HDD)
- describe how data are stored on and deleted from a solid state drive (SSD).

1 Computers: from rarity to ubiquity

Timings

This first section of the course should take you around one hour to complete.

The theme of this section is the development of digital technologies, from the physically large and highly expensive equipment of the 1950s to the omnipresent networks and computer-based devices which our digital lives are founded on today. This section will:

- compare the development of the computer to the development of the telephone
- describe some of the digital technologies that form an integral part of many people's lives
- introduce the term *information society*.

1.1 The telephone

In the 1950s, telephones represented the cutting edge of technology (Figure 1). However, they were not only large and expensive to purchase but also inconvenient to use. If you wanted to make a call over a distance of more than 15 miles, you first had to call the operator who would then make the connection.



Figure 1 A 1950s telephone

Telephones were also expensive to use. In the UK, calls were charged in units of three minutes, each unit costing the equivalent of between 5 and 20 pence in today's currency. (In comparison, in 1959, a pint of milk cost about 3 pence.) Few people had a home telephone and even fewer made long-distance calls. The high cost of using a telephone ensured that anyone who had one tended to use it infrequently, and then only for short calls.

In 1959, the General Post Office, which ran almost all of the UK telephone systems at the time, introduced subscriber trunk dialling (STD). This advance in technology allowed users to make long-distance calls directly, without an operator, and to be charged only for the actual duration of the call. The telephone became easier and much cheaper to use. As a result, more people began to use it and for longer calls.

Over time, the telephone has become part of the background of our lives, until today when it is extremely unusual to find someone who has neither a home telephone nor a mobile phone. Just as the telephone changed from a status symbol to become simply another piece of the modern world, so the computer has made a similar transition.

1.2 The computer

There were very few computers in the 1950s, and those in existence were treated as objects of wonder with almost mythical powers. They were nothing like the computers of today. For one thing they were huge, as shown in Figure 2. They were also delicate and consumed a lot of electricity, wasting much of it as heat.



Figure 2 The computer scientist Grace Hopper (1906–1992) beside a 1950s computer

Today, however, a computer is just another item stocked in supermarkets alongside toothpaste and dog food. And as computers have become cheaper and smaller, they have been incorporated into a kaleidoscopic range of devices that bear no resemblance to what was once thought of as a computer. Powerful computers now sit at the heart of objects as diverse as smartphones and games consoles, cars and vacuum cleaners. The cost of computer power continues to decrease, making it possible to incorporate computer technologies into almost any object, no matter how small, cheap or disposable. And these **smart devices** are ‘talking’ to one another, not just within a single room or building but around the world via the **internet**, using the **world wide web** (see Box 1). Thus, even as

the computer vanishes from sight, it becomes vastly more powerful and ever-present – to use a term you'll become very familiar with, it is now **ubiquitous**.

In the smartphone, two previously separate technologies – computers and mobile telephone systems – have become converged in one device. This **convergence** of computing and telephone systems has resulted in new goods and services for people to buy and use. For example, people who were not previously interested in computers are happy to take photographs on their mobile phones and share them. In this situation, the computer in the mobile phone handles the processing of the photograph and enables it to be shared by email or uploaded to social media sites.

Activity 1 (exploratory) Growth in use of technologies

Can you think of another technology that has made the transition from novel to commonplace, like telephones and computers?

Discussion

There are many possible answers to this question. We thought of washing machines, which have advanced from hand-driven drums to the automatic machines of today.

Box 1 The internet and the web

You will have come across the terms *internet* and *web*. Although in everyday life people tend to use these terms interchangeably, in reality they are two separate (although related) entities.

The internet is a global network of networks: an *internetwork* (hence its name). It is the infrastructure that connects computers together.

The web (short for *world wide web*), on the other hand, is a service that links files across computers, allowing us to access and share information. Thus the web is a software system that has been built upon the **hardware** of the internet.

Apart from anything else, this means that it is technically incorrect to refer to 'searching' or 'browsing' the internet. When you carry out an online search, you are in fact searching the web!

1.3 The information society

As a result of advances in computing and information technology (IT), our notions of time and location are changing – distance is no longer a barrier to commercial or social contact for those of us connected to a suitable **network**. Some people may find it difficult to imagine not having access to the information and services that play a crucial part in their daily lives. Others may feel that they have no part to play in the digital world because their network access is very limited or even non-existent. Some simply don't care about the digital world, viewing it perhaps as a waste of time. Yet whether we are aware of it or not, digital information is flowing constantly around us.

Consider a computer that is connected to the internet – the one you are using to study this course, for example. This may be a computer you use at home, in a library or at work; you may use it on the move or in a fixed location. Whatever the case, this computer is part of a

complex system, consisting of wires and optical fibres, microwaves and lasers, switches and satellites, that encompasses almost every part of the world. The oceans are wrapped in more than a quarter of a million miles of fibre-optic **cable** with several strands of glass running through it. Each of these strands can carry thousands of simultaneous telephone conversations, numerous television (TV) channels, or a range of other forms of digital content (such as web pages).

This modern communications network enables us to use a mobile phone in the depths of Siberia or take a satellite telephone to Antarctica (Figure 3), watch TV in the middle of the Atlantic, do our banking on the train, or play games with a person on the other side of the world. It is one of the greatest technological achievements of the last 30 years and it is so reliable, so omnipresent, that we very rarely stop to think about what actually happens when we dial a telephone number, click on a web link or switch TV channels. Or rather, we tend not to think about it until something disrupts the network – whether it is a widespread problem such as a power cut, or something more localised such as finding ourselves in a rural area with slow broadband access or no mobile phone signal.



Figure 3 Using a satellite phone in Antarctica

The end of the twentieth century and the beginning of the twenty-first century are often compared to other historical periods of great change, such as the Industrial Revolution, because of the huge technological changes that are happening in many areas of our lives. These developments are taking place in conjunction with correspondingly large social and economic changes, often characterised by the terms **information society** and **network society**. Such notions are frequently referred to by policy makers when driving forward changes in our technological infrastructure. Politicians often refer to the inevitability of technological change in our information society and stress the need to be at the forefront of these changes in order to secure future prosperity. For example, by developing the broadband network infrastructure, by making public services available online, and by equipping schools and local communities with computers.

Another term you may have encountered is that of the **knowledge society** which refers to the way in which new information systems can transform human societies. Included in the knowledge society is the idea of the ‘learning society’. The pace of change is so rapid today that learning can no longer be confined to our school years and early adulthood.

Everyone must continue to learn throughout their adult lives in order to benefit from the economic opportunities that rapid development makes possible.

Activity 2 (exploratory) Impact of technological changes

Can you think of an example where changes in technology have resulted in changes to your work, social or family life?

- Have those changes improved your life?
- Have they created any problems?

Discussion

The biggest change for me has been the ability to work from home. I'm currently at home typing this text on a computer while listening to some music, which is also stored on the computer. I'll shortly email my document to a colleague, who will be able to read it a few seconds later. I can also share my document through an online file storage system.

Being able to work from home has its advantages and disadvantages. I find it easier to concentrate when I'm working at home, rather than in a busy open-plan office. It also saves me time and money that used to be spent on travelling. On the down side, the boundaries between my home life and my work life have become very blurred, as my family will confirm.

The rise of texting and messaging

You might wonder about how these technologies change the way we live, work and relax, but sometimes it works the other way round. Unintended uses sometimes develop alongside the intended uses of emerging technologies in our information society. A classic example is the text-messaging facility on mobile phones, often referred to as SMS (which stands for *short message service*). This was originally a minor feature designed to be used by engineers testing equipment – it was not expected to be used by phone owners at all. Yet in the first decade of the twenty-first century, SMS messages were one of the most profitable parts of the mobile phone business (International Telecommunication Union, 2006). SMS resulted in a whole new method of communication and form of popular culture, different ways of interacting with radio and TV, and even a new language form: texting. Income from SMS messages peaked at around US\$104 billion in 2011 (Associated Press, 2012).

In more recent times, **instant messaging** (IM) 'apps' used on smartphones, such as WhatsApp, Snapchat and Telegram, have overtaken SMS in popularity. By 2015, the IM company WhatsApp was handling 30 billion messages every day, compared with the 20 billion SMS messages sent daily (Evans, 2015). This difference in usage is partly because there is no extra cost involved in using IM services. SMS messages can be expensive on some contracts and require a mobile phone signal, but they are still in use. Multimedia messaging services (MMS) allow for text, audio and images to be sent via the mobile phone signal. The total number of SMS and MMS messages sent in the first three months of 2016 was 24.1 billion, a decrease of 1.8 billion messages (6.8%) compared to a year previously (Ofcom, 2016). These services and applications are blurring the boundaries between texting and online communication.

Note: 'billion' is a word that in the past had different meanings in the UK and the USA. However, the two countries now agree that a billion is one thousand million (1000 000 000) and this course will follow that convention.

Now read Box 2 before attempting the next activity.

Box 2 Interpreting charts

A bar chart is a way of representing numbers pictorially and can illustrate patterns or trends in the data. In a bar chart, the numerical values are represented by the height or length of the rectangles. Their particular strength is enabling us to make comparisons.

The chart in Figure 4 represents the percentage of internet users who used instant messaging in 2005 and in 2014 according to age group. All charts and graphs follow a set of conventions. The charts should include a title explaining what is shown or being represented. The two values being described in Figure 4 are the percentage of internet users and the age group. Each axis should be labelled clearly to indicate what units are being used and have a marked scale.

In this example, the vertical axis represents percentage of internet users on a scale of 0% to 90% and the horizontal axis shows the age group. As with all types of graph, it does not offer any explanations about why the data is how it is. You will have noticed that use varies by age group and that there has been a considerable increase in use over time across all age groups.

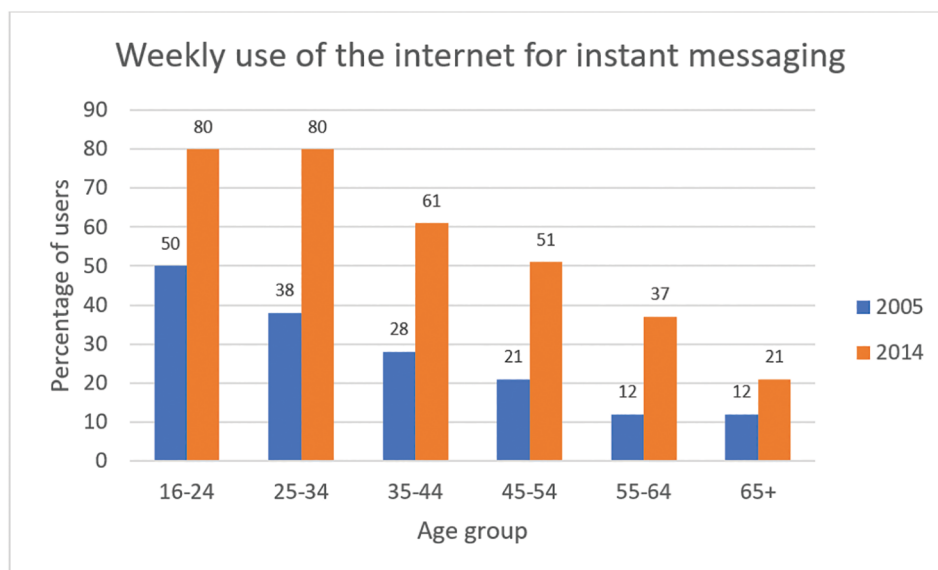


Figure 4 Weekly use of the internet for instant messaging, by age group: 2005 versus 2014 (Source: Ofcom, 2015)

Activity 3 (self-assessment) Interpreting a chart about internet use

Study the chart in Figure 4 and then answer the following questions.

- Which age group shows the smallest increase in the percentage of internet users who instant messaged between 2005 and 2014?

- b. Which age group shows the biggest increase in the percentage of users between 2005 and 2014?
- c. Can you offer any explanation for the increase in use among all of the age groups?

Discussion

- (a) The 65+ age group shows the smallest increase in the use of instant messaging, that is 9 percentage points.
- (b) The 25–34 age group shows the largest increase in the use of instant messaging, that is 42 percentage points.
- (c) One reason for the increase in use is that smartphones have become much more ubiquitous. While they are still expensive, the handsets are cheaper and more intuitive to use and the 'bundles' of telephone and internet access have become more affordable.

1.4 Summary of Section 1

After briefly considering the rapid development of computers, this section outlined some of the ways in which digital technologies pervade the world around us, giving rise to the social and economic changes that characterise our information society.

The concept of an information society is a key one in this course and the next section will give you a better idea of what it involves.

2 Some aspects of our information society

Timing

This section of the course should take you around three hours to complete. If you don't have time to work through it all at once, there are break points where you can stop and return later.

Section 1 introduced the concept of a digital world. This section will outline some key aspects of such a world. You will look at how it affects business, information, entertainment, public services and communication.

One of the things you will notice while studying computing and IT is the speed of change. Some of the examples you read about in this course may no longer exist, or are nothing like as important as when this material was written in 2019. This is inevitable, but you shouldn't be concerned by it; your focus should be on the general principles, which remain valid. For example, some companies you read about may cease to exist, and some may be more or less prominent. Yet the likelihood is that, although companies come and go, there are still ways to share photos (as Flickr and Instagram allow at the moment), to edit documents online (as is possible with Microsoft Office Live and Google Docs), to store files online (as file-hosting services such as Dropbox allow), and so on.

2.1 Business

Every time you use a debit or credit card in a shop, the shop till communicates with a card terminal that transmits your identification details from your card to your bank or credit card company for verification. Your balance is then adjusted according to your purchase. A similar chain of events is initiated if you shop online (buying a ticket for an airline or a train, perhaps) or over the phone (when booking a cinema ticket, for example). Many banks also provide online banking services, reducing the need for customers to visit a branch. **Automated teller machines (ATMs)** allow you to check your bank balance and withdraw cash wherever you are in the world.

In each of the above situations – using a debit or credit card, shopping or managing your money online or over the phone, or using an ATM – the machines involved are connected via a network to a central computer, which has records of your account in an electronic filing system known as a **database**.

Financial services have undergone huge changes in recent years as a result of developments in the digital technologies driving them. The examples just described show how convenient and accessible such services have become. Yet at the same time, issues of identity and security have become a concern. New ways of communicating have also created new types of crime, including identity theft and financial fraud. In turn, these problems have fostered the development of new security industries that try to inform us and sell us solutions to reduce the chances of us becoming victims of online crime. Also, this move towards online services means that banks have closed many branches, leading to job losses. This is particularly difficult for customers who prefer to use branches, especially those living in rural areas.

Commerce

Advances in digital technologies have led to changes in many areas of commerce, with some existing kinds of business being transformed by the opportunities these developments offer. One of the most obvious changes is the emergence of retailers such as Amazon, which have an online shop but no physical one that customers can visit. Yet the internet not only benefits the largest companies but also allows even the smallest retailers to advertise their services to a global audience. Incredibly specialised companies can flourish by relying on the internet's immense reach to deliver potential customers.

Of course, this growth has resulted in casualties in traditional (sometimes called 'bricks and mortar') retailing. High-street shops specialising in items such as books, DVDs, music and games have all lost business to online retailers. This has driven some shops out of business but a number of high-street stores have also opened successful online stores. Online retailers have lower costs – they don't pay expensive high-street rents and can easily be based in countries with low tax regimes – and they can pass on these savings to customers. There are other concerns about these large online retailers, such as the employment terms and conditions of warehouse workers.

Although the low costs offered by online stores can be very attractive to customers, they might be counterbalanced to some extent by the less immediate and less tangible nature of the shopping experience. However, online retailers make up for this by providing a variety of services that reassure and inform their customers. One such service is the tracking of goods online – although this is a development that can produce its own frustrations, as illustrated in Figure 5. Interestingly, since 2015, Amazon has set up a small number of 'till-free' bricks-and-mortar shops in the USA, including Amazon Books and some convenience shops also, and is now considering shops in London.



Figure 5 A light-hearted look at the pros and cons of online tracking

Activity 4 (exploratory) The advantages and disadvantages of online shopping

Add some more entries to the following table of advantages and disadvantages of online shopping. Can you think of ways in which online retailers may try to address the disadvantages?

	Advantages	Disadvantages
Buyer	More choice Can often track goods	<i>Provide your answer...</i>
Seller	<i>Provide your answer...</i>	<i>Provide your answer...</i>

Discussion

We thought of the following but you may have come up with others.

	Advantages	Disadvantages
Buyer	More choice	Can't try goods (e.g. shoes, clothes) first
	Can often track goods	Harder to return goods if faulty
	Often cheaper than a local shop	Harder to get help and advice before or after buying
	Some items can be downloaded immediately after buying	May be worried about online fraud
Seller	No need for a physical shop	Reliant on delivery services
	Lower running costs	Likely to get more returned goods
	Don't need to keep everything in stock – can arrange delivery from suppliers	Need to respond to email or telephone queries
	Can supply music, books, software, etc. for download rather than having to supply a physical item	Some potential customers will be reluctant to buy online

Online retailers may try to address the disadvantages listed above in various ways. For instance, they might provide:

- free collection of returned items
- links to online reviews of products to help advise prospective buyers
- telephone helplines
- support forums to help customers before and after buying
- advice on how to shop safely online.

As well as direct retailing, other types of businesses have also moved online – auction sites such as eBay fall into this category. Some online businesses are less conventional and, as a result, it's often harder to see how they find the money they need to survive. For example, there are many collaborative projects that produce free products including

software, online encyclopedias and educational resources. These often rely on volunteers contributing their time, with money being provided by advertising, sponsorship or donations.

Work

Technology has changed the way in which other businesses operate too. Greater quantities of information are exchanged between numerous locations over public and private networks. Vast amounts of data are stored on computers and accessed remotely from a variety of devices.

Just as individuals buy from companies online, many companies now sell to each other online as well, for the same reasons of reduced cost and wider choice that attract individuals to online stores. Manufacturing tasks that used to take days can now be completed in minutes using computer-operated machine tools working in automated production lines.

The way we use technology has also affected our individual working lives. For example, telephone and online banking mean that banks no longer need large numbers of counter staff, and the role of travel agents has changed as more people book their holidays directly from the vendor by going online. Some companies have responded by reducing their number of employees, while others have retrained their staff to provide more specialised services to their customers. More generally, many people working in an office environment are expected to learn how to use new software applications in order to do their jobs.

2.2 Communities

Break point

This would be a good point to take a break if you need to do something else before returning later.

As well as revolutionising the commercial world, the internet has had an enormous impact on the way we communicate. While there are still people in many parts of the world who do not have internet access, many of us have access at home or at work. As a result, we have the opportunity to communicate with others using email, instant messaging and **online discussion groups** (in online **forums**). Existing communities have created new ways of communicating, and new online communities have developed. **Social networking** plays an increasingly significant role in the lives of many people.

The recent pace of communication change has amazed many of us, and there is no reason to think that things will slow down. Figure 6 illustrates the relative size and significance of online communities in 2010. The first version of this map appeared on the web in early 2007 when MySpace was the most important social networking site and the micro-blogging site Twitter didn't exist. Over time, new online communities may appear and grow and the current ones may decline, or others which were declining may increase in size and importance.

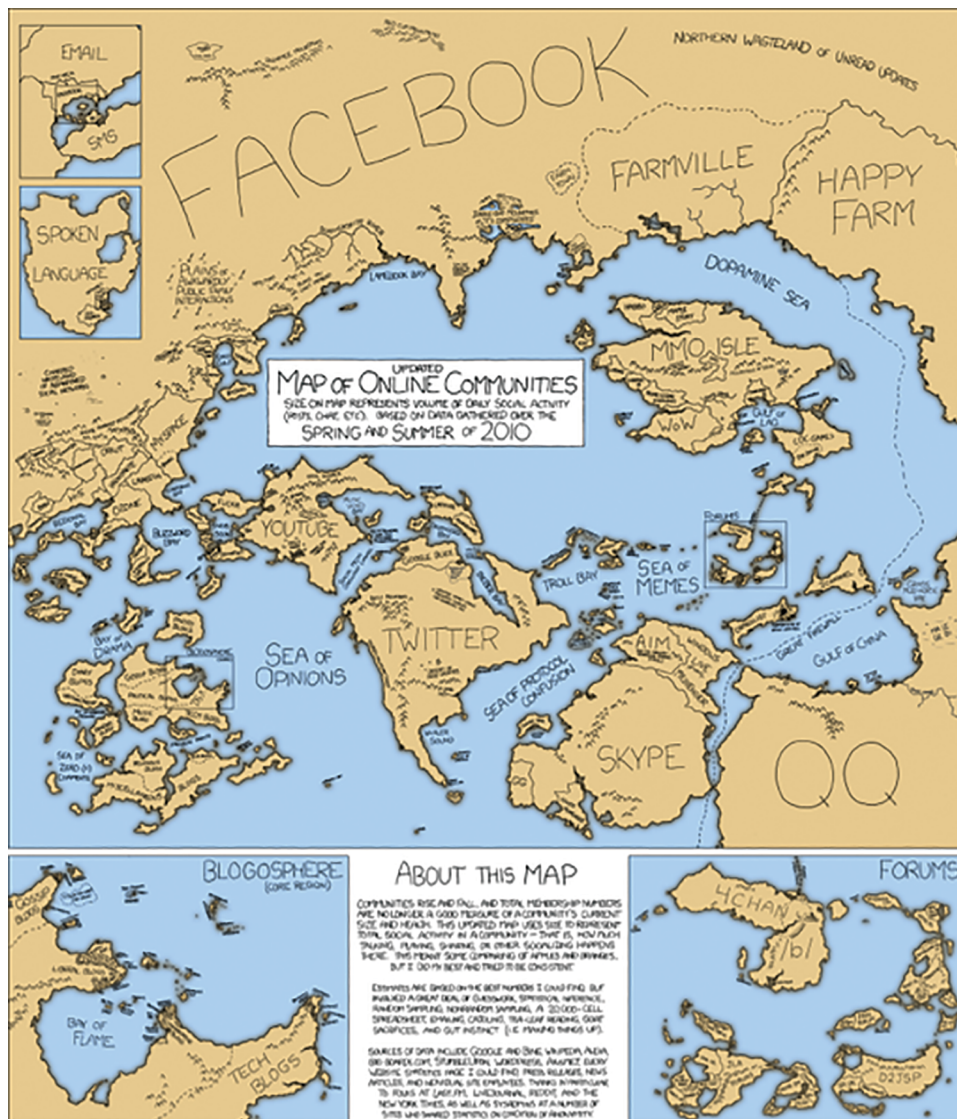


Figure 6 View of the world's online communities in 2010 (go to <https://xkcd.com/802/> for a version you can zoom in on)

2.3 Information

The internet has had a huge impact on the availability of information of all kinds. Material on the web reflects widely differing viewpoints, from official news bulletins to unofficial rumours, and comes from widely differing sources, from commercial megastores to community groups. Since no individual government, company or person has control over it, the internet has paved the way to unfettered publishing of information of all kinds, raising questions about the authority and regulation of this information. Some governments try to exert control over the information their citizens can access and create, with varying degrees of success.

For example, the widely used online encyclopedia, Wikipedia contains millions of articles, many of high-quality scholarship, but there are also articles that are of suspect quality and may even be deliberately misleading. Its greatest strength is also its greatest weakness in that anyone with an internet connection can edit its pages, except for those that are

protected. The encyclopedia itself advises users to treat with caution any one single source (whatever the media) or multiple works that derive from a single source.

The ever-increasing volume of online information available on the web means that it is important to think critically about what you find, especially if you are going to use it for study or work purposes. Finding and evaluating information online are important digital and information literacy (DIL) skills.

2.4 Entertainment

The world of entertainment is constantly evolving as new ways of creating and distributing the media we watch and listen to are developed. Digital broadcasting has changed the way we experience TV and radio, with increasingly interactive and participative programmes. Digital cameras, printers and scanners, together with desktop publishing and photo-editing software, enable greater numbers of people to experiment with image production, while online image- and video-sharing sites allow anyone with access to a relatively basic mobile phone or digital camera to share photos and videos with the rest of the online world. New digital technologies have also been at the forefront of changes in the production and distribution of music, and computer gaming has developed hand-in-hand with the evolution of graphical interfaces.

However, our increased exposure to digital entertainment has resulted in increased conflict between the rights of the consumer and the rights of the producer of the media. It is now much easier for the products of the media industries established during the twentieth century – film, music, and so on – to be illegally copied and distributed in a form that is indistinguishable from the original. **Copyright** holders are taking steps to prevent this by developing a range of **digital rights management (DRM)** techniques that make it much harder to create copies, as well as by trying to persuade users of the benefits of the original product.

Such attempts at persuasion can look very threatening, as I noticed on a recently purchased CD which has the following printed on the back cover:

FBI Anti-piracy Warning:

Unauthorized copying is punishable under federal law.

Several questions came to mind when I read this. Does that mean I can't legally put the music on my MP3 player? Copying from one format to another is legal in the USA but, although the law is under review, it is not legal in the UK at present (although people in the UK are not deterred from doing it). Will I be in trouble only if I distribute copies of the CD to other people? Making copies and distributing them to friends is illegal on both sides of the Atlantic. These are all valid concerns that demonstrate some of the problems surrounding the use of copyrighted material.

Copyright is very much a 'live' issue in the UK, the EU and the USA at present and is discussed in Section 3.3. There are many other issues arising from this, and it is very easy to make the digital future sound bleak. You have probably heard predictions to the effect that illegally copying media and making information freely available on the web puts whole businesses and hundreds of thousands of jobs in the established media industries at risk. However, as in other areas of the digital world, there are also opportunities for these businesses if they can adapt to the new environment and modify their business models to survive and grow in different directions.

Activity 5 (exploratory) Issues around the growth of digital entertainment

Can you think of any other problems connected with the growth of digital entertainment?

Discussion

There are many possibilities. For example, the advent of digital television affected even people who didn't particularly welcome it. Around the world, analogue signals are gradually being switched off as new digital signals are introduced, as happened in the UK. This meant that people had to buy new digital televisions as the old analogue versions became obsolete – quite an expensive business! At the time of writing, the UK government plans to switch off analogue radio at some stage, which means that people may need to upgrade to DAB (digital audio broadcasting) radio equipment.

You might also have thought of more technical problems, such as how to transmit the large quantities of digital data required for some forms of entertainment – video, for example – in an acceptable time and retaining acceptable quality.

2.5 Public information and services

Public bodies such as governments and transport agencies are increasingly providing services online, allowing us to organise various aspects of our daily lives more easily. These services range from simple information displays, which let us check things such as weather forecasts and transport timetables, to interactive sites that allow us to make bookings or enquiries.

In many parts of the world, medical records are increasingly moving away from paper and X-ray film towards becoming completely digital. This has several advantages, especially in allowing patients' records to be easily shared between departments within a hospital, and sometimes more widely with doctors' surgeries and other health workers.

In most parts of the UK, doctors and nurses walk around wards with tablet computers and other forms of handheld devices, instead of files and clipboards. In remote rural areas of some countries, doctors can use computer networks or even mobile phones to make a diagnosis if they are unable to see the patient in person. However, this is by no means universal and, even where such facilities exist, they aren't always available.

Information for travellers is also increasingly being made available digitally: for example, live online updates on road congestion and public transport, and arrivals information in stations and airports. Similarly, it is now common to book plane journeys online – in fact, some airlines now only accept online bookings and will only issue electronic tickets. Many of them strongly encourage, or even require, passengers to check in online as well.

In addition, many countries provide online access to at least some of their government services. For example, you might be able to renew or apply for a passport, book a driving test, claim benefits, or fill in your tax return online. Local authorities also provide digital information services – you might be able to reserve or renew a library book online, for instance – and there are numerous opportunities to learn online.

Changing internet use over time

The Office for National Statistics has been collecting statistics on the use of the internet since 1998 and has produced an annual report since 2006 based on a survey of people in the UK.

The 2019 report for England, Scotland and Wales showed a huge increase in the use of the internet and social media since their first annual report. These were the key findings:

- 87% of all adults used the internet daily or almost every day in 2019.
- 93% of households had access to the internet in 2019. This was an increase of 23 percentage points in the last decade.
- Of those households who did not have access to the internet, 61% of households did not feel that they needed the internet.
- In 2019, for the first time, more than half of adults aged 65 years and over shopped online, at 54%.
- The percentage of adults who make video or voice calls over the internet has more than trebled over the past decade, to 50% in 2019.
- In 2019, 7% of adults suffered fraudulent debit or credit card use from using the internet in the last 12 months.

(Source: Office for National Statistics, 2019)

Personally, I was surprised that only 93% of the population have internet access (the second bullet point) because this suggests that 7% of households still do not have access (although 61% of these households felt that they did not need the internet). Also, the growth in the use of internet shopping amongst the over-65s was very interesting. You may recall from Figure 4 that internet use for instant messaging was much lower for that age group in 2014.

As this section has shown, so much of our world is digital and this raises questions about inequalities in terms of access to goods and services including public services.

Security and risk

The twentieth century saw a dramatic change in the role of the state in many countries. During most of the nineteenth century, an individual might only have come into contact with the state for the purposes of taxation, marriage and death. At the end of that century and the beginning of the next, however, a series of social revolutions saw the state becoming involved in our health care, pensions and education. Unsurprisingly, each of these developments was accompanied by a significant increase in the amount of personal information stored about every one of us.

Computer technologies were developed especially to serve the enormous projects involved. IBM became a highly successful company due to its work on censuses in the USA and Europe, while the world's first business computer, LEO, was used for a variety of tasks including the calculation of tax tables for the British Treasury in the 1950s.

With the vast amount of personal information being held about us in various places, it is becoming increasingly important for us to be able to prove our identities – not just for travel but for other activities such as purchasing expensive or restricted items, paying bills and opening bank accounts. The UK is unusual in Europe in that (at the time of writing in 2019) it does not have a compulsory identity card system, despite the fact that identity cards were put in place during both world wars. However, effectively, driving licences,

passports and some types of travel pass have become forms of identity card. In several countries, identity card or passport schemes are being upgraded with new **biometric** technologies such as iris or face recognition, which claim to identify individuals uniquely. As well as the personal information that we know about, there may also exist information about us of which we are unaware. Since the terrorist attacks on the USA in 2001, much of the western world has become far more security conscious, and governments and companies alike have developed and deployed technological countermeasures. These range from smart video surveillance systems that can identify an individual in a crowd and track their movements, through the biometric technologies mentioned above, to the searching of databases for suspicious activity.

Activity 6 (exploratory) Experience of technological security issues

Can you recall an occasion when you have been personally aware of technological security measures?

Discussion

My own experience of security measures at airports involves checking in online and giving details of my own passport and that of my travelling companion several weeks before travelling to mainland Europe. A colleague who visited the USA recently went through a range of airport security screenings. In addition to having his belongings checked, he was photographed at least twice (as well as being under almost constant video surveillance in the airports), had his fingerprints scanned electronically, and was required to fill in numerous online and paper forms.

The promise is that such technologies will make us safer but could they turn the world we live in into a society strangely reminiscent of the nightmare vision of 'Big Brother' in George Orwell's novel *Nineteen Eighty-Four*?

As you've seen so far in this section, there is plenty of opportunity for digital information about each of us to be created. Some of this information we might intentionally give out ourselves – on social networking sites, for example. Other information about us may, as described above, be gathered more surreptitiously by various agencies.

In general, we have little control over how digital information about us is used or who receives it. We might assume that information gathered legally by a government agency, for instance, will be handled appropriately and used only for our benefit. Yet there have been many examples of governments and private organisations 'losing' confidential data by storing or transferring it insecurely. For example, in 2015, the communications company TalkTalk was forced to issue a warning to its customers that their data had been hacked and that criminals were using the stolen information to trick people into handing over their bank details.

2.6 Communicating on the move

Advances in **digital technology** have, in a very short space of time, revolutionised the way many of us live our lives. Nowhere is this more evident than in our ability to communicate as we travel.

When I first moved away from home, communicating with my parents involved a shared payphone in the hallway or a telephone box at the corner of the street. Today, I have a

mobile phone with which I can send and pick up emails quickly and cheaply wherever I am. As well as email, my phone enables me to communicate in several other ways – instant messaging and social networking, for example, not to mention voice calls. I can also use the built-in **global positioning system (GPS)** to find out where I am (Figure 7), I can even plot my location on a website to let my friends and family see where I am (or at least see where my phone is).

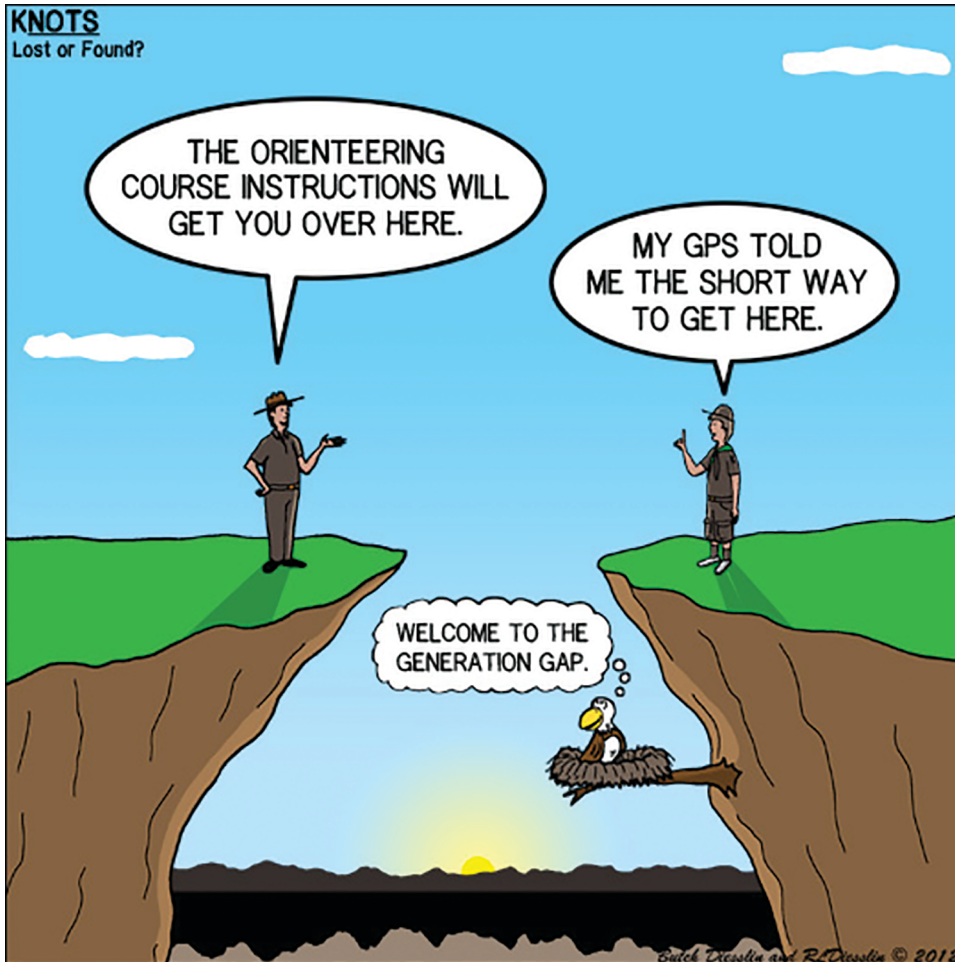


Figure 7 The disadvantages of GPS?

2.7 Summary of Section 2

In this section you learned about some of the key aspects of an information society, online social networking being one of them. The next section will outline some ways in which you can participate in online discussions effectively.

3 Participating in a digital world

Timing

This section of the course should take you around two hours to complete. If you don't have time to work through it all at once, there are break points where you can stop and return later.

Section 2 mentioned the rise of social networking – one aspect of our increasingly digital world. You may already communicate online to some extent in your daily life.

When we have a face-to-face conversation, we don't just rely on the spoken words to establish the other person's meaning. Unconsciously, we are also monitoring the tone of their voice, their facial expression and their body language. Telephone conversations are a little more ambiguous because we can no longer see the other person; email and online discussions are harder again, since all we have is text. It is extremely easy to misinterpret words on a page, so the writer must take great care before pressing the button that sends their message to the world.

This section provides a quick guide to good practice in contributing to online discussions. It should help you to work and socialise with other people online – in your studies, your social life and your working life. Much of the content is summed up by the familiar tenet known as the Golden Rule: a concept common to many ethical codes, which simply states that we should treat others as we would want them to treat us. Just as this Golden Rule is relevant to good manners – 'etiquette' – when talking face-to-face, so it is relevant to online communication. To help us apply it, it has been developed into guidelines for online behaviour called 'net etiquette' or, more commonly, just **netiquette**.

Netiquette is intended to make us all think about how we behave online and to make us aware of the effect our words could have on other people reading them. If it seems that there are far too many rules to follow, be reassured that they aren't hard-and-fast commands that you must remember and obey. Netiquette does not encompass every situation you may find yourself in – it's perfectly possible to obey all the guidance below and still annoy someone – but it will give you a good foundation for your participation in online discussions.

Most of what follows is common sense and good manners. Some of it may be familiar to you but please take time to read it if you don't have much experience of using online discussion groups to work with other people. There's a big difference between working in an online community and socialising online, so even if you are experienced at the latter, you should find this section useful.

3.1 Netiquette: respecting other people online

As children, we quickly learn many rules about how to interact with other people. Some of these rules are common sense, such as 'don't interrupt a speaker' and 'say please and thank you', and are necessary if we are to reduce the likelihood of arguments or causing offence.

Thank, acknowledge and support people

People can't see you nod, smile or frown as you read their messages. If they get no response, they may feel ignored and be discouraged from contributing further. Why not send a short reply to keep the conversation going? This can make a big difference in a small group setting. However, do bear in mind that in a large, busy forum too many messages like this could be a nuisance.

Acknowledge before differing

Before you disagree with someone, try to summarise the other person's point in your own words. Then they know you are trying to understand them and will be more likely to take your view seriously. Otherwise, you risk talking at each other rather than to each other. You should also recognise that other people are entitled to their point of view, even if you consider them to be entirely wrong.

Make clear your perspective

Try to speak personally. That means avoiding statements like 'This is the way it is ...' or 'It is a fact that ...'. These sound dogmatic and leave no room for anyone else's perspective. Why not start by saying 'I think ...' or 'I feel ...'? If you are presenting someone else's views then say so, perhaps by using a quotation and an acknowledgement.

Other cues

Emotions can be easily misunderstood when you can't see faces or body language. People may not realise you are joking, and irony and satire are easily missed – these are all good reasons to think before you send a message. To compensate for these restrictions, early internet users came up with the idea of the smiley face – :) or :-)) – which then grew into a whole family of **emoticons**.

Small digital images called *emojis* are commonly used in electronic communications, such as text messages and emails, to represent ideas and emotions. Incidentally, the word 'emoji' comes from the Japanese term for 'picture characters' so, although the word is similar to 'emoticon', this is a coincidence. Some examples of emojis are shown in Figure 8 below.



Figure 8 Examples of emojis

Remember that the systems which many forums are based on only support plain text, so you can't always rely on fonts and colours to add meaning. Even if you are using a forum that allows so-called 'rich text', it's possible that other users will be picking up messages as plain text emails or as text message alerts on their mobile phones and will not see your formatting. AND DON'T WRITE IN CAPITAL LETTERS – IT WILL COME OVER AS SHOUTING!

If you read something that offends or upsets you, it is very tempting to dash off a reply immediately. However, messages written in the heat of the moment can often cause offence themselves. It's much better to save your message as a draft and take a break or sleep on it. That gives you a chance to come back to your message when you're feeling calmer and ask yourself 'How would I feel if someone sent that message to me?'. If you decide it will make things worse then make sure you edit it before you send it.

The best advice is to try to be aware of your audience before you post. The internet is a global phenomenon; people from widely differing cultures and backgrounds may read what you write online, and what you find funny may be offensive to them. It may take time to work out what sort of 'audience' can be found in a particular forum. Some are very permissive and allow almost any sort of behaviour, while most (like those at The Open University) will not tolerate bad behaviour or abuse.

Activity 7 (exploratory) Using emoticons to express feelings

Look at the email exchange shown in Figure 9.

- What emotions are being expressed through smileys and typography?
- Would Jon and Sue still be on speaking terms if they hadn't used these devices?

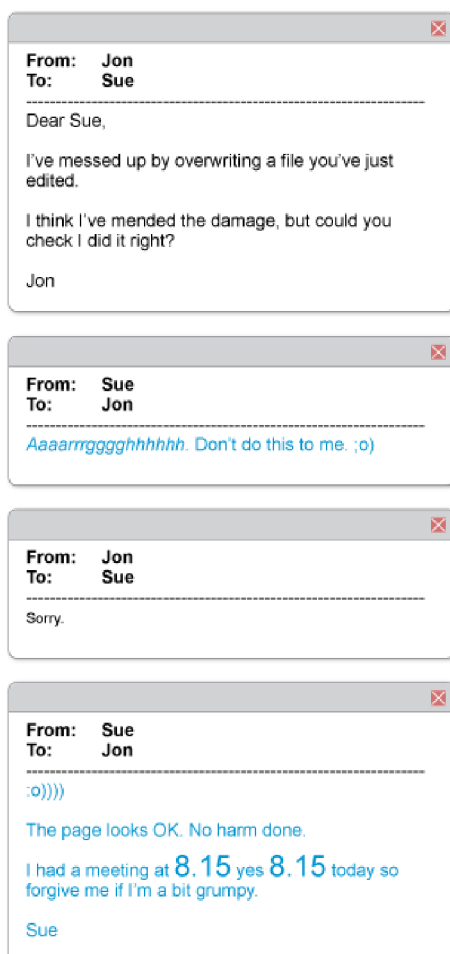


Figure 9 Email exchange between Jon and Sue

Discussion

In Sue's first reply to Jon, she expresses her frustration by typing 'Aaaarrggggghhhhhh' but she ends that message with a winking smiley. Jon's reply then says 'sorry' in a very small voice! Sue's final reply starts with a happy smiley to show that everything's OK. She uses a large font when she mentions the annoyingly early meeting time.

I feel sure that Jon and Sue would still be friendly after this email exchange. But I have seen email exchanges between colleagues that had the opposite effect, when the participants didn't take care about how they expressed themselves in their messages.

Moderation

Moderators are forum participants who are responsible for keeping order. They may have capabilities within the forum greater than those of other participants: for example, they can sometimes add new participants and suspend people who are abusive. They also work to keep the discussions friendly and relevant to the forum.

A forum with a moderator is said to be *moderated*. Forums without moderators are *unmoderated* and are generally places where newcomers should tread very carefully. Moderators tend to introduce themselves early on in forum discussions, so it's usually clear whether a forum is moderated or not.

Some other advice

- Keep to the subject and pick the right forum for your contribution.
- Before you write a message, check any rules about what is and is not considered acceptable in the forum. Many discussion forums have rules, aside from netiquette, about things such as links to commercial sites.
- Take a little time to use the forum's search facilities to see if your question or topic has already been discussed or covered in a set of frequently asked questions (FAQs). If it has, you should at least scan the existing messages to see if your points have been addressed.
- Don't feel you have to post immediately. Take your time to see what is being discussed and get a feel for the group you're joining. This very sensible behaviour has the unfortunate name of *lurking* but is quite acceptable online. If you want to post, many discussion groups have a forum devoted to new users where they can introduce themselves to other readers. These are always good places to get started.
- Try to keep your messages short and to the point. People don't want to read long, rambling messages, especially if they can't work out what response you're looking for.
- Write a concise subject line (title) for your message – people often won't spend time reading messages unless the subject line looks relevant.
- Keep to one subject (topic of discussion) per message. If you want to cover another subject, do it in another message.
- When replying to a message, quoting part of that earlier message can be helpful so that readers can easily see what you are referring to. Add your response *after* the quoted material, not before it. And keep your quotation short and to the point, otherwise the resulting messages will get longer and longer.
- If you ask a question and it is answered, thank the person who responded. It's not only polite, it also shows that the discussion has come to an end.
- If you've reached a point where you disagree with someone and neither of you is going to change your opinion (Figure 10), realise the conversation is over, agree to disagree, and move on.



Figure 10 Online discussions can become addictive!

Activity 8 (self-assessment) Netiquette

The following quiz on netiquette will test your understanding of what you have studied so far in this section.

- Writing in capitals has the same effect as shouting.
☐ True
☐ False
- You can write what you like as long as you put a smiley face after it.
☐ True
☐ False
- You should always quote the whole of someone's message when replying.
☐ True
☐ False
- Lurking is considered to be bad netiquette.
☐ True
☐ False

- e. If you disagree with someone, you should acknowledge what they have said first.
 - ☐ True
 - ☐ False
- f. If you have several ideas to discuss, it is better to put each one in a different short message than combine them into one long message.
 - ☐ True
 - ☐ False
- g. It is very important to keep a discussion going until you have convinced everyone to agree with your point of view.
 - ☐ True
 - ☐ False

3.2 Ethical and legal considerations

Break point

This would be a good point to take a break if you need to do something else before returning later.

In using a computer for communications, you have many rights of free expression, but you also have certain responsibilities to respect others and to have some awareness of privacy and confidentiality in relation to online communications. You should be aware of the following points.

- An email is generally considered to be equivalent to a private letter, and should not be quoted or forwarded to anyone else without the permission of the original sender. This can be particularly poorly observed in companies (even those whose employees are told to assume that all online communications are for the recipient's eyes only, unless otherwise stated).
- Besides the informal rules of netiquette, most forums have a code of conduct and conditions of use that govern acceptable behaviour. For example, the use of the online forums provided by The Open University is covered by its Computing Code of Conduct. You will usually find a forum's terms of use linked from its home page, or listed in the code of conduct that you are asked to agree to when you first register for an account.
- Considerations of copyright and **plagiarism** (cheating by using another person's work as if it were your own) apply to online discussions. If you are quoting something written by someone else, put it in quotation marks and acknowledge the source.
- Some forums are not wholly public, in which case messages should not be copied outside the forum. The forum's terms of use may specify this.

3.3 Copyright

One of the reasons the web has grown so quickly, and one of its most fascinating aspects, is that almost anyone can publish almost anything on it. It is very easy to find information, images, audio and video files on the web, which you can then save and incorporate into your own material.

Copying is so easy that people often make the mistake of assuming that everything on the web is freely available. This is not the case: most information you will come across is likely to be covered by copyright law. This applies not only to online TV programmes, music, photographs, books, and so on, but also to information – for example, in the form of online academic papers or simply in someone's blog.

Material that is not subject to copyright is said to belong in the **public domain** and can be used by anyone. Older works of art, music and literature, such as the works of Shakespeare and Beethoven, are in the public domain. However, individual printings, adaptations or recordings of those works are copyrighted to the publisher or performer. The situation can become very complicated because the duration and extent of copyright differs between countries. For example, material may be in the public domain in the USA but still under copyright in the UK. As a result, it is always wise to assume that any third-party material (that is, material originating from someone else) is still protected by copyright unless you're sure it's in the public domain.

Copyright holders can prosecute individuals and organisations for infringing their rights. In recent years, music and film companies have sued individuals for very large sums of money. Below are some general points you should bear in mind.

- You should seek the author's permission if you want to use any copyrighted material. Just because something is on the web does not mean that it is freely available for you to use.
- Many websites have usage policies explaining how their material can be used. Some are more restrictive than others, so make sure you find and follow the relevant policies.
- Information published online may have been put there by someone who is not the copyright holder.
- You can use copyrighted print material for your own study purposes: the generally accepted guideline from copyright legislation is that you can download and print or photocopy limited extracts of any one book without seeking permission, but you must give a full reference to show where it has come from.

This may all seem rather intimidating, but it's not quite as bad as you might fear – fortunately, copyright law makes some concessions to students and learners. As a result, students don't have to ask permission to use copyrighted material in answering an assignment or exam question, although they must still include references. However, if they want to reuse the same material for any other purpose at a later date, normal copyright law applies and they must seek permission.

3.4 Good academic practice

Many courses will encourage students and learners to use the web as a resource. Students use it as part of their studies and in their assignments or projects, and they should find it a great help in understanding and practising the things they learn.

However, using information found on the web in this way can cause problems unless you take a little care. When using material written by other people, you can quote their words but it is good academic practice that such quotations should always be *limited* and *acknowledged*. This applies whether you're quoting from this course or from other sources such as websites, journals or newspapers.

It can be very tempting to copy and paste large chunks of text into your notes – and possibly then into assignment answers – without giving a reference. However, that is very bad academic practice. It's far better to use quotations sparingly and to rewrite most of the material in your own words. This allows you to show that you've understood the material and it also helps you to remember it.

In addition, it's good academic practice to give a reference to the source of any third-party material you include in your own work. Not doing so is not only impolite, as you're failing to acknowledge the help that someone else's work has given you; it's also regarded as plagiarism and is never acceptable.

3.5 Summary of Section 3

This section discussed some of the ways in which you can make good use of the web, both for interacting with other people and for finding information. When working online, it is also important to consider how to protect yourself and your computer, and that's what we turn to next.

4 Are search engines bad for your health?

Timing

This section of the course should take you around two hours to complete. If you don't have time to work through it all at once, there are break points where you can stop and return later.

In countries without systemic surveillance, the internet appears to offer perfect freedom of expression. While previous generations of authors relied on finding editors and publishers willing to take the risk of distributing their material, today anyone connected to the internet can create new content, comment on existing material and share a wide range of media on any number of free services. Their contributions are then made accessible through search engines and by further sharing and commenting. The result has been an explosion of content.

When I was writing the paragraph above, I idly entered *Freedom of speech* into Google; it told me there were nearly 250 million relevant results (see Figure 11 below)! It will have undoubtedly changed by the time you read this.

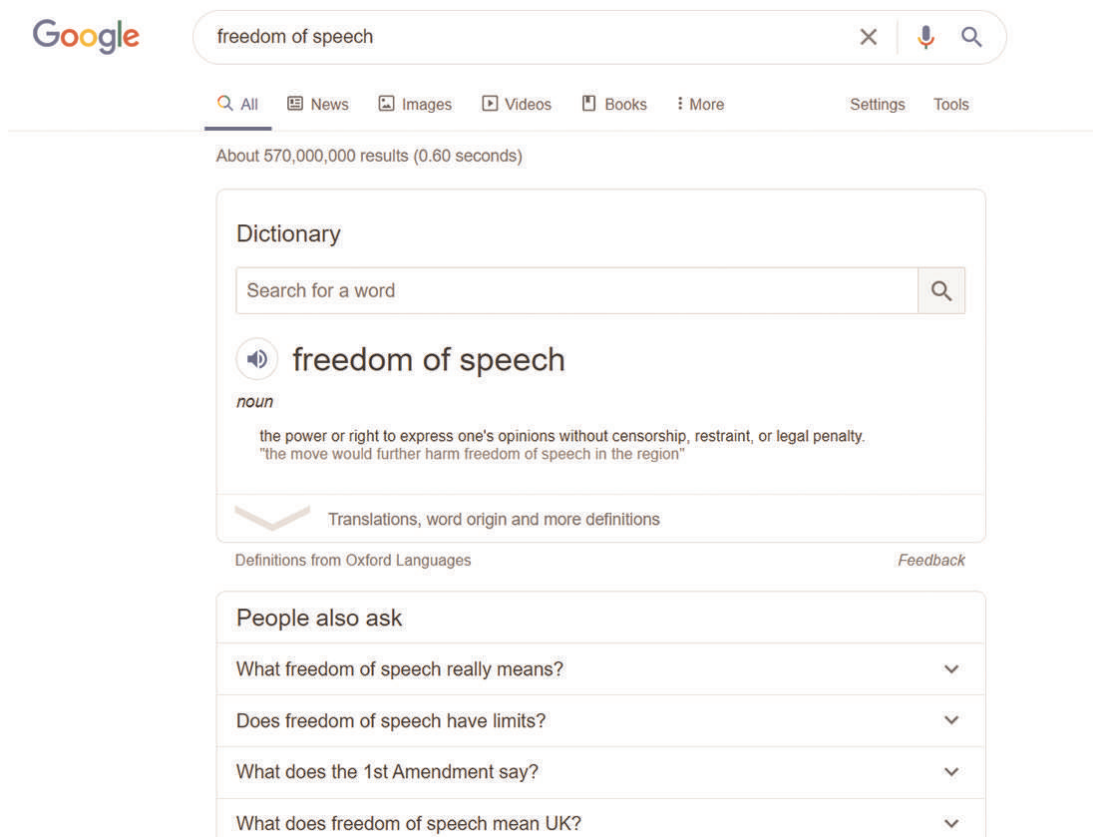


Figure 11 A Google search for 'freedom of speech' produced 248 million results when this course was written. But how many of them are actually useful? Interestingly, the first true result is a link to a company which is not flagged as advertising.

However, I have many questions about these results, including the following.

- Are there really 250 million different places where I can learn about freedom of information?
- Who created all of these documents? There can't be more than 250 million experts in the field of freedom of information.
- Do these links point to high-quality, factual resources, to incorrect information, or to advertising? In fact, the first result below the definition of freedom of speech is a link to a company selling speech recognition software. Very useful, but nothing to do with the concept of freedom of speech.
- If the material is of varying quality, were the best resources prioritised so they appear near the top of the results, rather than buried among millions of others?
- Or have some resources been given greater visibility because their creators have paid to be promoted to the front page of the search results?
- And were any results suppressed because they were considered unacceptable?

There is little chance that I could ever answer these questions in a satisfactory manner for this single search query, let alone the countless requests for information made to search engines every day.

Fortunately, researchers have been examining how people use information from search engines. You will now look at an example regarding children's health.

4.1 Childhood vaccinations

Childhood vaccinations are a routine part of life but, in recent years, there has been growing public concern over the possible adverse consequences of vaccines. These worries began with the publication of a paper by the prestigious journal *The Lancet* in 1998, studying the onset of developmental disorders, including autism, in 12 children.

Follow-up studies failed to substantiate the *Lancet* paper. It became clear that the paper's author, Dr Andrew Wakefield, had failed to follow correct ethical and scientific processes, had faked some findings, and had direct financial interests in producing evidence that would allow cases to be brought against MMR vaccine manufacturers. Lacking any scientific justification, the *Lancet* paper was formally retracted in 2002. Wakefield was struck off the [UK medical register](#) and barred from practising medicine in the UK (General Medical Council, 2010).

Note

Since the publication of Wakefield's discredited paper, researchers have repeatedly re-examined the risks of vaccination and checked the quality of previous research to determine whether their findings are still sound (Demicheli et al., 2012; Maglione et al., 2014; Taylor et al., 2014). The scientific consensus is that there is no evidence of a link between vaccination and developmental conditions such as autism. Vaccination is safe and effective against a range of childhood diseases that cause severe injury and death. This is based on a number of very large studies, encompassing many millions of vaccinations (compared with the *Lancet* paper, which used just 12 people).

While there is no evidence for any link, a belief that vaccinations cause certain chronic childhood conditions persists – especially online, where many sites continue to spread gravely misleading information.

Researchers at the University of California studied how people encounter misleading information when searching for information about vaccination (Ruiz and Bell, 2014). Nine search terms were selected from the most common vaccination-related queries submitted to Google. Three queries were positive (*vaccine benefits*, *vaccination benefits* and *MMR benefits*). Three were neutral (*vaccine*, *vaccination*, *MMR*) and three were negative (*vaccine risks*, *vaccination risks* and *MMR risks*). The first page of Google results, containing ten results, was retrieved for each search term.

Activity 9 (self-assessment) People's search patterns

Figure 12 shows the percentages of people who clicked on Google search results (the click-through rate): the top result, the second result, and so on. Does it support the researchers' decision to restrict their interest to only the first page of search results?

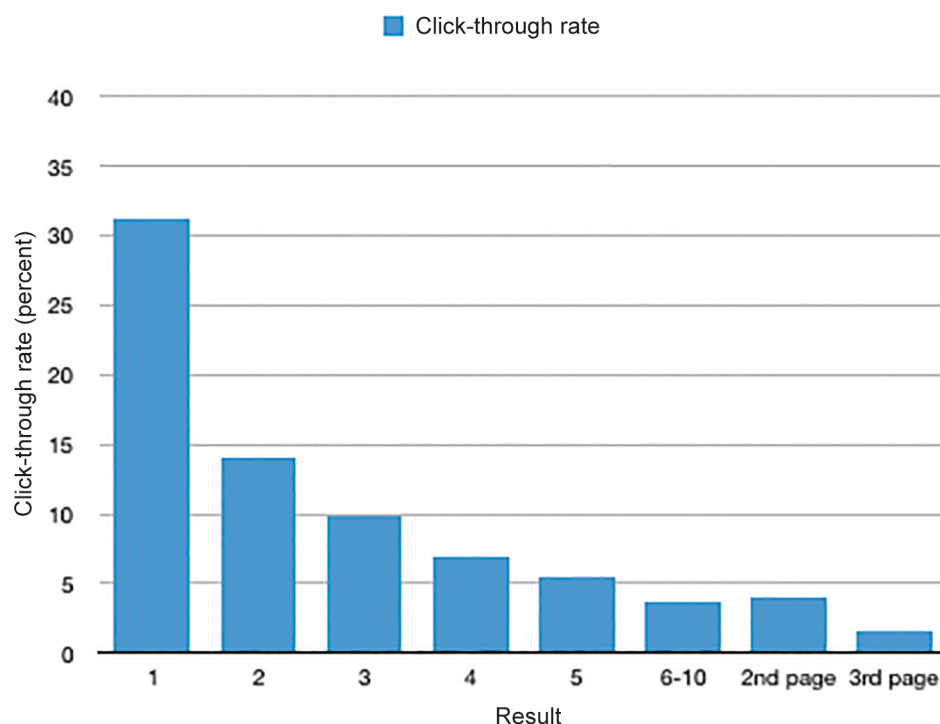


Figure 12 The click-through rate for Google searches during July 2014. Search rankings are shown on the horizontal axis. The first search result is clicked by more than 30% of all users. (Source: Petrescu, 2014)

Answer

Essentially, yes. Most people in this study (71%) only followed links on the first page of results. This finding is in keeping with other studies which show that the first page of search results attracts a disproportionate number of clicks. Lower-ranking results are rarely used by searchers, even though they may also contain useful information.

The first page of ten search results for each query was stored by the researchers. Of the 90 possible results, 84 contained information that could be used in the study. (The other results were variously indexes of other pages, book previews or discussed adult vaccines.)

The 84 pages were then analysed to identify key aspects of their content. The researchers were interested in what they termed *myths* and *recommendations*.

1. A **myth** is a falsehood about childhood vaccination, such as *vaccines cause autism*.
2. A **recommendation** is a call to action – it can be either positive (e.g. *make sure your vaccinations are up-to-date*) or negative (e.g. *do not harm your child by vaccinating them*).

The researchers then checked whether a site **perpetuated** a myth by not providing evidence that it was actually a myth, or whether a site **countered** a myth with further information showing that the myth was untrue or could not be substantiated. Their findings included the following.

4.2 The number of myths

While most sites contained factual information, myths were found on 16.7% of the websites studied. Negative search terms returned 28 useful results; on average, each of these results contained 1.93 myths about vaccination. Each of the 27 sites found using neutral search terms contained an average of 0.53 myths, while positive search terms resulted in pages with 0.40 myths apiece.

Activity 10 (exploratory) Likelihood of encountering false information in search results

How much more likely were users to encounter false information when using negative search terms than when they used (a) neutral and (b) positive terms?

Discussion

- a. They were $1.93 / 0.53 = 3.64$ times more likely to see false information using negative search terms than neutral terms.
- b. They were $1.93 / 0.40 = 4.83$ times more likely to see false information than if they used positive terms.

The study shows that the choice of search term dramatically alters the likelihood that someone searching for information about vaccination will see falsehoods. A parent who is concerned about the possible risks of vaccination, and who is likely to use a negative term such as *vaccine risk*, is nearly five times more likely to see a myth than one who uses a positive search term such as *vaccine benefits*.

4.3 Countering myths

The number of myths that were countered varied between the three search terms. Websites returned by negative searches countered an average of 3.00 myths apiece. Pages found by neutral searches countered 3.27, but, surprisingly, positive searches were the least effective, countering just 2.87 myths on each page.

The researchers also studied how many myths were countered compared with those that they perpetuated. As you might expect, negative web searches were the least effective at countering myths about vaccination. These sites countered 1.55 myths for each myth they perpetuated. Neutral searches were much better, countering 6.17 myths for every myth

perpetuated. Positive searches were the most effective, countering 7.18 myths for each one perpetuated.

4.4 Explicit recommendations

Of the 84 websites examined, 17.9% recommended vaccination, 4.8% discouraged vaccination, while the majority (77.4%) made no recommendation either way. However, there were significant differences between the number of recommendations made by performing different types of search (Table 1).

Table 1 Percentages of sites making recommendations for or against childhood vaccination with differing search terms

Search term used	Recommend vaccination (%)	Discourage vaccination (%)
Negative	17.9	10.7
Neutral	25.9	3.7
Positive	10.3	0.0

(Ruiz and Bell, 2014)

Activity 11 (exploratory) Analysing search results for vaccination

- Which search terms are *least* likely to (i) recommend vaccination and (ii) discourage vaccination?
- How much *more* likely is a user to encounter a recommendation *against* vaccination when searching with negative search terms than with a neutral search term?

Discussion

- In both cases, positive searches produce the fewest explicit recommendations or discouragements.
- Users were $10.7 / 3.7 = 2.89$ times more likely to see a recommendation against vaccination than if they used a neutral term. Had they used a positive term, users would not have seen any sites discouraging vaccination.

This study demonstrated that the reliability of search engine results can be influenced by the choice of search terms used. In the case of vaccination, the wrong search terms can lead to people following incorrect health advice and being exposed to greater levels of concern and worry about the effects of childhood vaccination.

At first reading, the fact that users entering positive search terms encountered fewer myths appears to be good news. However, where a positive search returned a myth such as *Children's vaccines contain harmful chemicals*, the page was unlikely to counter the myth with factual information. Overall, positive searches were relatively unhelpful to parents who were worried about falsehoods surrounding vaccination. Instead, parents could only research these myths by negative searches, which returned less accurate, potentially harmful, information. The study's authors strongly suggest that websites

offering accurate information about vaccination *must* include information about myths and why they are false.

Note

In the UK, the [NHS has a comprehensive page](#) listing the facts and myths about childhood vaccination.

A second weakness of positive searches was that they were least likely to return pages recommending vaccination. The authors of the study suggest that the creators of positive pages knew the benefits of vaccination and assumed their potential audience were of a similar mind, when in fact this does not seem to be the case. The researchers recommend that site developers test their pages with differing search terms to ensure that factual information *always* appears high in search engine results.

4.5 Summary of Section 4

While we enjoy a high degree of free speech in the west, one aspect of our internet society that went largely unreported until recently is the ability of very large technology companies, such as search engines and social media platforms, to influence what we see – and, in doing so, to influence us.

This section explored the example of medical information available to parents. You saw how our choice of search terms may decide whether we're served with facts or myths. Welcome to a world in which computer algorithms can change our perception of reality. After examining how information is made available to you on the web via search engines, in the next two sections, you will find out more about how information is stored on your computer. You will join Gloria and Rupert and learn about some of the technical details of how data is stored on your computer.

5 Hard disk drives: a conversation

Timing

This section of the course should take you around two hours to complete.

Rupert was bored – he fiddled with his **Rolodex** and sighed deeply. It had been six weeks since he had solved his last case and no business had materialised since then. The mortgage payment was overdue, but at least he had caught up with his Open University module.

A sudden noise caught his attention – the street door of the dingy office complex had banged closed. Probably because of the draughts caused by lots of open windows, reasoned Rupert. No air-conditioning here and the temperature outside was close to 30 degrees.

Listening carefully now, Rupert followed the faint sound of footsteps to the second floor. He held his breath to the count of three after they stopped outside his office door. “Come in,” he said quietly and a bit too quickly, anticipating the knock. A young woman stepped through the door into the gloomy interior. Without invitation, she sank into the classy but tatty green leather chair next to the Art Deco desk. “Gloria,” she said, extending her hand in greeting. A silence followed. Rupert waited, patiently, knowing that saying less often means learning more. Nervously, she scanned the room before spotting the Rolodex (Figure 13). “What is *that*?” she asked.



Figure 13 Rupert’s Rolodex, which has an alphabetical index

“That,” said Rupert, “is my father’s Rolodex. It contains data about all of his clients over his entire career as a private investigator.”

"I have seen the term Rolodex used to mean the whole collection of an individual's business contacts, but I never knew that there was such a thing as a real Rolodex," said Gloria, showing her age.

"Yes, the state-of-the-art data management system for 1958, was the Rolodex," replied Rupert. "They still sell them, you know, but they come with a maximum of 500 cards these days – not like the monsters of the old days that could hold several thousand cards."

There was another short and uneasy silence.

"How do you store your data these days?"

Rupert, sensing that this was some kind of test said, flatly, "I've moved on to punched cards!" (Figure 14).

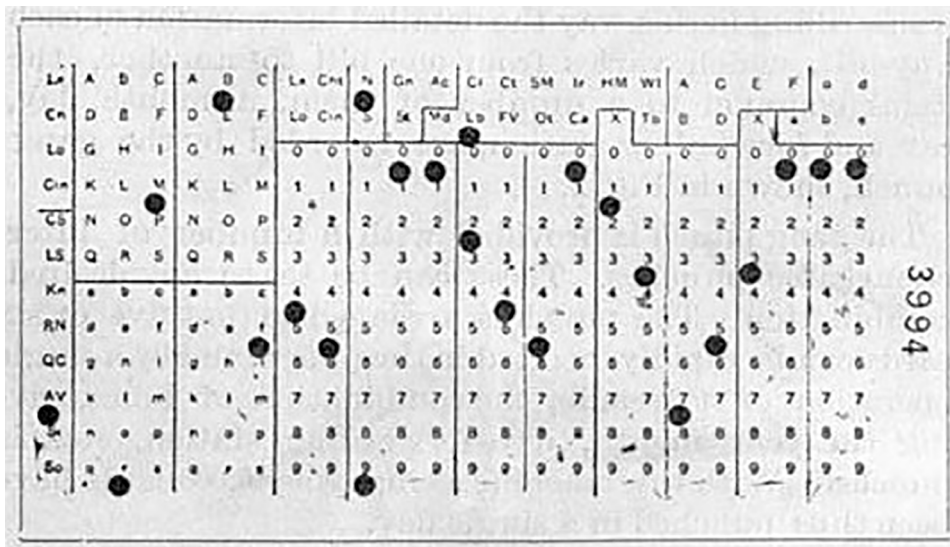


Figure 14 A punched card. Punched cards were widely used during the first half of the twentieth century to store data.

Gloria looked aghast until Rupert relented: "Only joking! Seriously, I know all about data. I am taking a Level 1 Open University computing module. I have learned about databases and the different formats for storing numbers and text."

"That's a good start," she replied. "I have a masters degree in digital data security, and I would like to move into **digital forensics** – you know, the branch of forensic science that is concerned with obtaining legal evidence from computer systems."

"Do you mean **computer forensics**?" said Rupert.

"Yes, same thing. But let's see how much you know already about the secrets that can be held in the data on a computer and how you can unlock them", Gloria continued. "For example,

- Why does solid-state memory create a problem for digital forensics?
- When is deleted data not really deleted?"

Rupert looked down at his shoes, sheepishly. "I haven't got that far in my module yet," he admitted.

"The reason I am here is to make a proposition," said Gloria, leaning towards him in a conspiratorial way. "I know how to find data in the guts of a computer. You are a private investigator who knows how people tick. Why don't we work together?"

5.1 The structure of a hard disk drive

The next morning, Rupert and Gloria were hunched over a disembowelled computer in Rupert's office.

"Let's talk about **memory**," said Gloria. "What do you remember – if you'll excuse the pun?"

"Well, I recall that there are several different types of memory – for example, registers and main memory. But perhaps we should start with the secondary storage because that is persistent memory – its contents are not lost when the power is switched off."

Rupert picked up the component he recognised as the hard disk drive (Figure 15).



Figure 15 The case of a hard disk drive

"So," he continued. "Let's say that I have ... hmm ... 'acquired' a hard drive. How can I find out what interesting data it contains?"

"First, let's look at the anatomy of a **hard disk drive**, or **HDD**, or simply **hard disk**, as we call them in the trade," replied Gloria. Extracting a bunch of small screwdrivers from her handbag, she carefully unscrewed the casing. "I don't advise you ever do this with a hard disk that is still in use," she said. "Not only will you invalidate any warranty it may have, but also the disk was sealed in a clean environment to minimise dust entering the casing. Even a particle of dust could cause it to crash!"

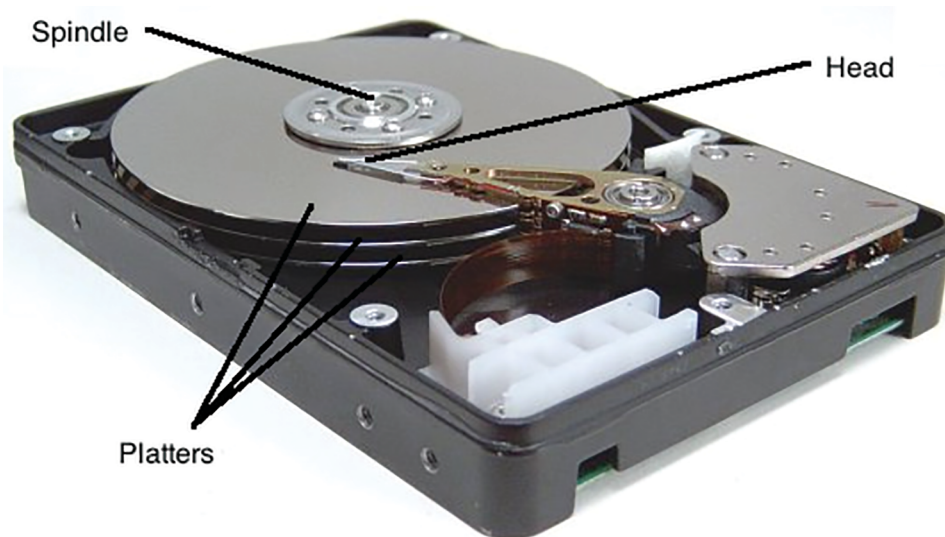


Figure 16 The inside of a hard disk drive

The casing was soon removed and Gloria and Rupert peered inside (Figure 16).

“You see those three circular disks, stacked up in a pile? Those are called platters and they rotate on that central spindle. They are covered on both sides with a metal that can be magnetised in tiny areas to represent zeros and ones. Each platter has a head that can pass over every part of the disk as it is spinning. This head is able to detect and change the magnetic areas, and so can read and write the zeros and ones.”

“It has been a long time since I have studied any physics,” said Rupert, “but won’t a spot on the edge of a platter have a greater speed than one close to the centre?”

“Yes,” replied Gloria. “Most hard disks now rotate at 7200 revolutions per minute and, of course, each point on the disk has this same rotational speed. However, a spot on the outside of the platter will move further in one revolution than one on the inside, as it has to rotate through a circle with a greater radius in the same time – so it moves faster”.

“Does that mean that the data on the outside of a platter can be read faster than data on the inside of a platter?” asked Rupert.

“Indeed,” replied Gloria, looking impressed.

5.2 Storing files on a hard drive

Rupert scratched his head. “I know that the **operating system** manages the data on the hard drive,” he said, “but how does it do it?”

“Well, the first thing to understand is how data is organised. You know that data is stored in files – you can have document files, spreadsheets, database files, music files, and so on? Well, the operating system runs something called a **file system**, which dictates how data is written to – and recovered from – a disk, and also records the location of the file stored on the disk.

“So, just as old-fashioned, paper-based offices might have different filing systems, the same is true of computer file systems. There are file systems that are faster than others, or designed for different storage capacities, or to be more secure or robust. The developers of operating systems have been instrumental in designing file systems, and so most personal computers use one of the file systems that are integrated into Windows, Apple or Linux.”

5.3 Formatting a hard disk

“If there are lots of different operating systems, each with its own kind of file system, do you need a different kind of disk drive for each?” asked Rupert. “Good question,” said Gloria, “but the answer is ‘no’. This is because we can prepare almost any hard disk drive to work with any operating system and its file system by going through a process, called **formatting** it, before we try to save any data to it.

“The most important thing that happens when a disk is formatted is that at least one area of the disk must be loaded with the operating system’s file system in readiness for it to store data. These areas are called **partitions**. You can have a single partition that contains all the storage space on the drive or divide the space into several different partitions. If you want to run more than one operating system on your machine, you can even create partitions that have different file systems. Anyway, you need at least one partition on the drive and, if you have more than one partition, the formatting process will cause them to be displayed as separate drives by your operating system – for example, in Windows Explorer or Finder on a Mac.

“Some formatting procedures may also check the physical structure of the disk for errors, recording their location so that data is not written to these locations.

“Once a disk has been formatted, you can write data to it.”

“I had a ... ‘friend’ once who formatted his disk by accident one day and he lost all of his data,” said Rupert, looking sad.

“Well, if you reformat a disk that has already got data on it, the data does not disappear, but it can’t be retrieved by the operating system in the usual way. So it usually needs special tools to get it back.”

“Ah, it’s a bit like you have destroyed the indexing system in a large library,” said Rupert. “All the books are still there, but they are really difficult to find.”

“Exactly.”

“So should you ever reformat a hard drive?” asked Rupert.

“Well, sometimes you might want to go back to scratch to install a new operating system on the disk or on a partition of the disk. In that case, you would want to reformat the disk or partition for the new operating system. As a last resort, reformatting a disk can also be a way of removing **viruses** from the computer, or fixing other errors.”

“So how is the space organised on the hard drive, once it is formatted?” asked Rupert.

“All hard disks are formed of a series of tracks – sometimes called rings – that can contain data. A disk track is too large to manage the data effectively as a single storage unit. (An individual disk track can store more than a megabyte of data, which would be very inefficient for storing small files.) So, as part of the formatting process, tracks are divided into several numbered, equal divisions known as sectors.

“The sectors are arc-shaped pieces of a track. Almost all file systems create sectors that can hold 512 bytes of data. The sectors are grouped together in clusters. So a cluster is a larger unit of memory whose size depends on the particular file system being used. A cluster always consists of one or more consecutive sectors, but typically there are four or eight (or some other power of two) sectors in a cluster. When a file is written to the hard disk, it always takes up a whole number of clusters.”

“I might be getting lost now,” said Rupert, plaintively.

“No worries. There is a lot of terminology to take in, I know. Here is a summary.

1. Each platter of a hard disk is divided into a number of concentric tracks.

2. Each track is divided into a number of **sectors**, each of which can store the same amount of data. A sector is the smallest physical storage unit on the disk, and on most file systems it is fixed at 512 bytes in size.
3. A **cluster** can consist of one or more consecutive sectors – commonly, a cluster will have four or eight sectors. As a file is written to the disk, the file system allocates the appropriate whole number of clusters to store the file's data."

"Confusingly," continued Gloria, "on diagrams, the sectors on a platter are usually shown as being arranged in segments, like pizza slices. This was the case on older hard disk drives, but it meant that the sectors on the outside of the disk had a larger area than those closer to the centre, which meant that they held fewer bits per unit area, and were less efficient at storing data than the inner sectors. On modern disks, each sector has the same area, so they each store the same number of bits per unit area" (Figure 17).

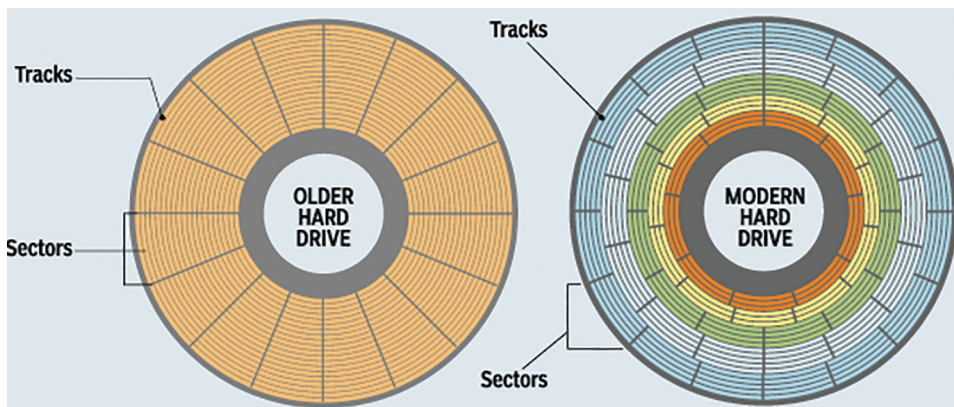


Figure 17 A comparison of sectors on an older and a more modern hard disk drive

Activity 12 (self-assessment) Storing data in bytes, sectors and clusters

Given that a sector is 512 bytes in size, how many bytes of storage are there in a cluster composed of four sectors?

Provide your answer...

Answer

There are $4 \times 512 = 2048$ bytes in a cluster.

"So, once a file has been written to one or more clusters, how does the operating system know where to find the file again?" asked Rupert.

"Well, let's talk fat" was the surprising answer.

"What do you mean, 'fat'?" exclaimed Rupert, sounding surprised at this turn in the conversation.

"No, you idiot," said Gloria affectionately, "not fat, but **FAT**, which stands for **File Allocation Table**. It is the area of the hard disk that is used as an index of every cluster on the disk and records whether a cluster is being used or not. It is what is at the heart of the file system called **FAT32**, which was used by Windows operating systems but is now

mainly used with solid-state memory, such as flash. It can only cope with a maximum file size of 4 GB (gigabytes).

“Windows computers now mainly use a file system called **New Technology File System (NTFS)**, where a table called a **Master File Table (MFT)** does a similar job. Apple has a file system that is unimaginatively called The Apple File System, and the Linux file system is called ext4. They each have similar tables.”

“I suppose that when a new file is to be saved, the operating system can also use this table to determine which clusters are in use and which are free to be allocated?” said Rupert.

“Exactly,” replied Gloria. “The space that is available for files to be written to is referred to as **unallocated space** on the disk and, of course, this is always a whole number of clusters’ worth of bytes.”

5.4 Deleting data from an HDD

“What happens when a file is deleted from a hard disk?” asked Rupert.

“When a file is deleted, the operating system doesn’t erase the file; it simply makes the clusters that the file occupies available for reallocation. So the data is still there until it is overwritten, but there is no reference to it in the file allocation table, and the file is said to be dereferenced. However, the operating system at some point might allocate a new file to one of those clusters, which overwrites the original data.”

Gloria paused while the implications of this dawned on Rupert.

“So even if a file has been deleted, the data might still be right there on the hard disk if it hasn’t been overwritten yet?” he said, barely containing his excitement.

Gloria smiled. “Yes – but it is even better than that,” she said.

“As we have seen, the cluster system means that there is almost always some unused space in the last cluster when a file is stored. The amount of unused space depends on the cluster size and the file size. The **logical size** of a file is a measure of the number of bytes of data a file actually contains. Its **physical size** is almost always bigger than this because it has to be stored in a discrete number of clusters.

“So, for example, take a file that has a logical size of 1280 bytes. In a system where there are four sectors of 512 bytes in a cluster, the file takes up a whole cluster (or 2048 bytes), which means that the physical size of the file is 2048 bytes. The difference between 2048 and 1280 is 768, which means that there is a **slack space** of 768 bytes” (Figure 18).

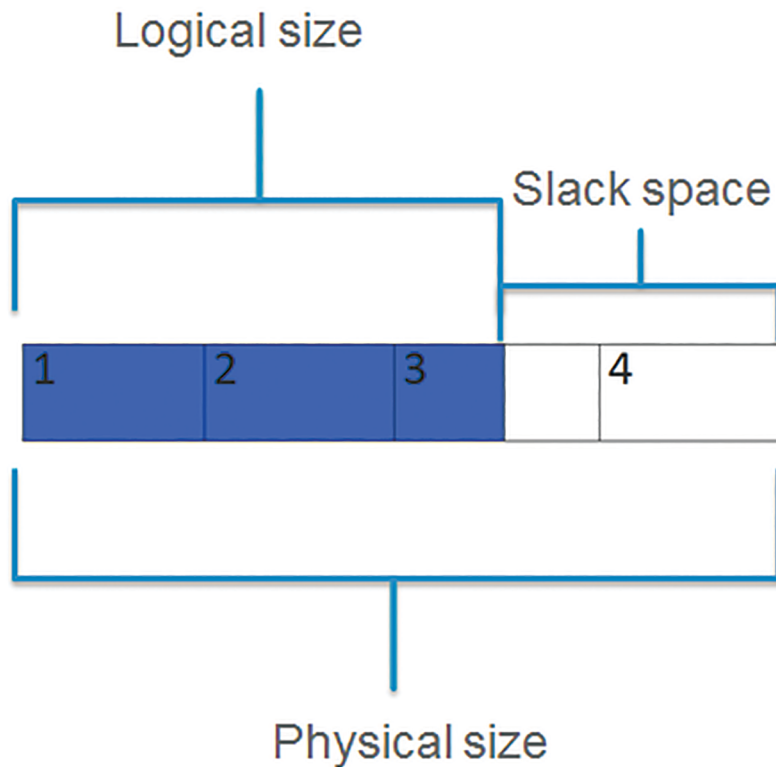


Figure 18 Slack space in a cluster

“After a file is deleted, if the newly allocated file does not occupy the whole of the cluster or clusters, the data in the slack space is not overwritten. And if the deleted file was larger than the newly allocated file, there will be bits of the old file still intact, although not visible to the file system. So data in the slack space might well come from files considerably older than the deleted file.

“This leftover data, which is called **latent data** or **ambient data**, can provide investigators with clues as to what was originally stored in the whole cluster, which may in turn provide leads for other enquiries. In 2016, for example, the Federal Bureau of Investigation (FBI) revealed that it had reviewed millions of email fragments that resided in the slack space of the memory of the personal servers of former Secretary of State Hillary Clinton, in order to determine whether or not she had improperly stored or transmitted classified information.”

Activity 13 (self-assessment) The difference between unallocated space and slack space

Rupert was tired at this point, so he did not quite take in the difference between unallocated space and slack space. Help him out by writing down a couple of sentences to explain it to him.

Provide your answer...

Discussion

Unallocated space is free space on a hard drive that can be used to store data. It is a discrete number of clusters. Slack space is the unused space between the end of the actual file and the end of the cluster.

How to permanently delete data from an HDD

“Just out of interest,” pondered Rupert, “what *can* you do to totally eliminate data on a disk drive?”

“It is surprisingly difficult,” Gloria admitted. “For a hard disk drive, there are only three sure-fire ways – overwriting, degaussing or physical destruction..

- **Overwriting (wiping or shredding).** Some disk management **programs** provide an ‘overwrite’ utility that fills every part of the disk with zeros or ones or a random mix of the two. In theory, this extremely time-consuming process completely obliterates any trace of the previous data. Some operating systems support a ‘secure file delete’ operation, which deletes a file and then performs an overwrite of just the section of the disk formerly occupied by that file.
- **Degaussing.** Data on HDDs (and other magnetic storage media) is stored in patterns of magnetisation. These patterns can be disrupted by a powerful magnetic field, and a sufficiently powerful field can erase an entire disk in a few minutes – a process known as degaussing. An HDD that has been properly degaussed can safely be reused or disposed of. Degaussing requires extremely expensive equipment and expert supervision to be effective (Figure 19). The increased reliability of modern magnetic storage media has been achieved by making them resistant to stray magnetic fields – which requires that even stronger magnetic fields must be used in the degausser. Improper degaussing leaves enough magnetisation on the medium that some or all of the data can be recovered.



Figure 19 A Garner data eliminator which uses degaussing to remove HDD data

- **Physical destruction.** Some people and organisations cannot tolerate the risk that overwriting software might not be completely effective or that their hard disks will not be properly degaussed. Instead, they require that their disks are completely destroyed when no longer needed. Incineration is the most popular method of destroying HDDs (although it must be performed in a specially designed facility, as they may contain toxic substances or produce them when burned), but they can also be crushed or shredded. A web search on how to destroy a hard disk drive comes up with several alternative methods!

5.5 Fragmentation

“I have noticed that a hard disk drive gets slower the fuller it gets. Why is that?” asked Rupert.

“If your operating system tries to save a file that cannot be stored in a single cluster, the file system breaks up the file into cluster-sized chunks and tries to save them in contiguous clusters.”

“What does ‘contiguous’ mean?” Rupert interrupted.

“It just means clusters that are next to each other on the disk. However, if contiguous clusters are not available, the file is fragmented, which means that the remaining clusters are written elsewhere on the same disk. That is called **fragmentation**, and it can slow down system performance because the file system must direct the heads to several different areas of the disk to find all the data in the file you want to read. The extra time for the heads to travel to a number of locations causes a delay before the entire file is retrieved.

“If you want to, you can change the cluster size of your hard disk drive when you format it. But you have to bear in mind that, although a larger cluster size reduces the potential for fragmentation, and reduces the amount of disk space needed to store information about the used and unused areas on the disk, a larger cluster size will also increase the likelihood of unused slack space.”

Activity 14 (self-assessment) Slack space

- a. Given that a cluster size on a particular disk is 2 KB (kilobytes), what is the physical size of a file with a logical size of 69 KB? What is the slack space in this case?

Provide your answer...

- b. Given that a cluster size on a particular disk is 4 KB, what is the physical size of a file with a logical size of 69 KB? What is the slack space in this case?

Provide your answer...

- c. What percentage increase in slack space does doubling the cluster size cause in this case?

Provide your answer...

Answer

- a. To find the physical size, first calculate the number of clusters the file will require, which is given by $69 / 2 = 34.5$. So the file will need 35 whole clusters. The physical size of the file is $35 \times 2 \text{ KB} = 70 \text{ KB}$. Hence the slack space in this case is $70 \text{ KB} - 69 \text{ KB} = 1 \text{ KB}$.
- b. To find the physical size, first calculate the number of clusters the file will require, which is given by $69 / 4 = 17.25$. So the file will need 18 whole clusters. The physical size of the file is $18 \times 4 \text{ KB} = 72 \text{ KB}$. Hence the slack space in this case is $72 \text{ KB} - 69 \text{ KB} = 3 \text{ KB}$.
- c. The percentage increase in slack space is given by $(3-1)/1 \times 100\% = 200\%$

“I have heard of something called ‘defragging’,” said Rupert. “Is that related to what you are talking about?”

“Yes,” said Gloria, “when you **defragment** (or **defrag**) a hard disk, you are using a software utility that moves the chunks of files to try to arrange them in contiguous clusters.” (Figure 20)

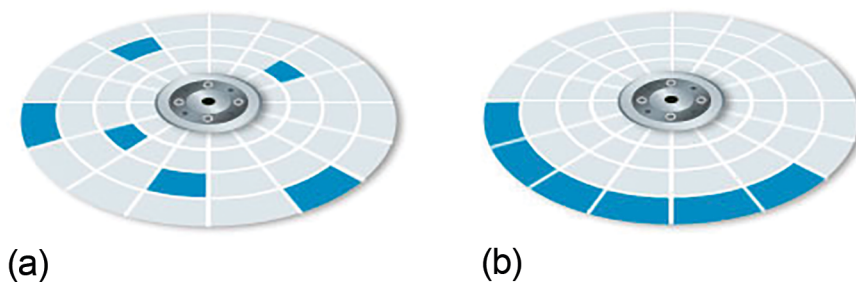


Figure 20 (a) A fragmented file occupying six clusters and (b) the same file after defragmentation

“Windows 10 should automatically schedule a defragging utility for you – once a week, by default,” said Gloria.

“How about on a Mac?” asked Rupert.

“There is no defragging program built into the operating system because OS X defrags files itself using Hotfile Adaptive Clustering (HFAC). It works by identifying the read-only files and files that are accessed frequently, and then allocating them to a space on the outer track of the disk called the ‘hot zone’, which has faster access. In general, then, it is not necessary to defragment a Mac, although it may become necessary to buy and use a utility if the system slows down when working with large files such as multimedia files, video editing tools, and so on, which require frequent read and write operations of large files.”

Activity 15 (self-assessment) Speed of a hard disk drive

Give two factors that determine the speed at which a file can be accessed from a particular hard disk drive.

Provide your answer...

Answer

1. The speed of access depends on where the file is stored on the disk. Clusters on the outer tracks of the disk are accessed more quickly than clusters on the inner tracks of the disk.
2. If a file is highly fragmented, it will take longer to access than a file that is stored in contiguous clusters.

5.6 Summary of Section 5

In this section you followed Gloria, a digital forensics expert, who has teamed up with Rupert, a private investigator. You learned about the structure of hard disk drives and how files are stored and retrieved from HDDs. You also learned that deleting data from an HDD can be undone. But then you were also introduced to various ways to permanently delete data from an HDD. The section ended with an explanation of HDD fragmentation.

In the next section, you will be introduced to solid-state drives, that is SSDs.

6 Solid-state drives: The conversation continued

Timing

This final section of the course should take you around two hours to complete.

“That was very useful information about hard disk drives, but my laptop doesn’t have an HDD – it has a solid-state drive,” said Rupert.

“Ah, you have an SSD,” said Gloria.

“If you say so,” said Rupert, tentatively.

“**SSDs** are **solid-state drives**, which use integrated circuits to store data. They use a technology called **flash memory**, which is a solid-state chip that maintains stored data without any external power source. It is commonly used in portable electronics and removable storage devices such as USB sticks, and as an alternative for laptop and desktop computer hard disk drives.”

“It sounds as if I am going to have to learn a whole new system for storing data,” said Rupert, a bit wearily.

“Well, not completely,” replied Gloria. The good news for you is that the file and operating systems still maintain the same system of dividing the memory into logical sectors and clusters, even though the physical form of a solid-state drive is very different from that of a spinning disk.”

Rupert brightened up a bit at this. “Ah, it sounds as if this is another example of the kind of abstraction my OU module is always banging on about. The operating system doesn’t need to know what physical type of drive it is reading data from, or writing to, as long as it understands the logical file storage structure defined by the file system.”

“You’ve got it,” replied Gloria.

Activity 16 (exploratory) Devices that use flash memory

Many portable devices contain flash memory. Such devices will hold data when they are switched off or unplugged. List any such devices that you use.

Provide your answer...

Discussion

You may have some of the following portable devices in your home: smartphone, GPS, MP3 player, digital camera, laptop, tablet, USB stick, Kindle, iPod, Xbox.

“Unfortunately for those who make their living out of digital forensics,” continued Gloria, “the use of SSDs in computers, rather than HDDs, is increasing. Just look at the stats in this figure.”

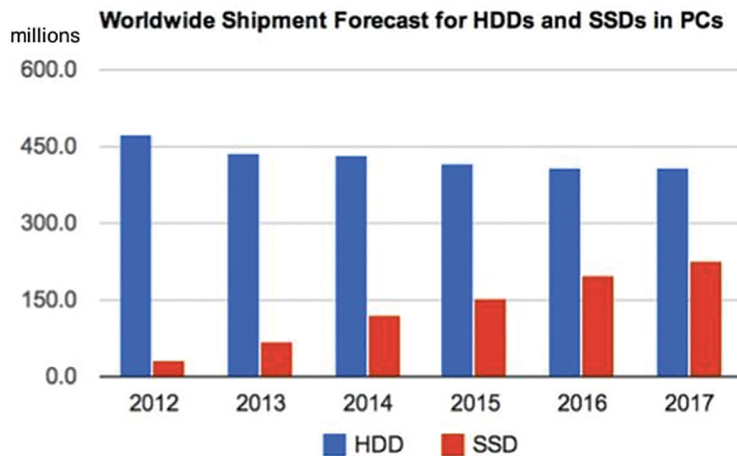


Figure 21 Growth trend of the worldwide use of SSDs in PCs

“But why is it unfortunate for us?” asked Rupert.

“I will tell you why tomorrow. In the meantime, your homework tonight is to look at this table which compares HDDs and SSDs.”

6.1 Comparing SSDs and HDDs

Rupert's homework, a table which compares HDDs and SSDs:

Table 2 A comparison of HDDs and SSDs

Hard disk drives (HDDs)	Solid-state drives (SSDs)
Rotating platters and mechanical arms mean they have a relatively fragile construction, leading to a greater risk of data loss and hard drive failure when transported.	No moving parts means a more rugged construction and little or no risk of data loss when transported.
When the drive starts, it takes some time to get up to the right rotational speed. This, coupled with the mechanical arm movement means comparatively slow response time and performance.	Nothing spins and there is no mechanical arm movement, which means comparatively fast responsiveness and performance.
Reduced battery life in a laptop because of a relatively high power requirement.	Prolonged battery life in a laptop because the lack of moving parts means a relatively low power consumption, which in turn means less power needed for cooling.
Access time depends on where the data is situated on the disk.	Access time is the same regardless of the location of the data.
File fragmentation can lead to decreased performance.	File fragmentation still occurs, but performance remains consistent.
Relatively slow in terms of booting up, launching and running applications and transferring files.	Relatively fast in terms of booting up, launching and running applications and transferring files.
Relatively cheap per byte.	Relatively expensive per byte.

Bright and early the next morning, Rupert was eager for answers.

“I can see lots of advantages for SSDs – in particular, the fact that they don’t have any moving parts seems to save a lot of problems. But why is their increasing popularity bad news for us?”

“One important factor is that fragmentation is not a performance issue on SSDs,” said Gloria.

“How come, if they have similar logical file structures?”

“Well, the key to that also has to do with the fact that an SSD has no moving parts,” replied Gloria. “Although an SSD provides secondary storage, because it is a solid-state device it is truly random access memory (RAM). That is why the access time for an SSD is the same, regardless of the location of the data, and so fragmentation does not lead to the same problems as for an HDD.”

Activity 17 (exploratory) RAM and defragmentation

After this hint from Gloria, Rupert understood why fragmentation does not lead to increased or inconsistent access times when accessing files on an SSD compared with an HDD.

Write a few sentences to explain his reasoning.

Discussion

When the head is trying to find the different clusters on a highly fragmented HDD, it has to make many movements over the disk, and wait for the disk to spin round so the cluster is positioned under the read head. This leads to longer access times.

Furthermore, files of the same size can have inconsistent access times, depending on how many clusters they occupy on the faster outer tracks of the disk or on the slower inner tracks. Both of these effects lead to a degraded performance when the disk becomes fragmented.

On an SSD, each cluster can be accessed in exactly the same amount of time, so the performance is always exactly the same, regardless of whether the files on the SSD are highly fragmented.

6.2 How flash memory works

“As I mentioned earlier, an operating system doesn’t need to know what kind of drive – HDD or SSD – that it is interacting with,” continued Gloria. “This is because there is software and hardware within the SSD device that allows it to emulate the logical structure of a spinning hard drive, including a file allocation table, sectors and clusters.

“However, when it comes down to how data is actually written to an SSD, they are quite different from hard disk drives.

“On a microscopic level, SSDs are made up of semiconducting materials that are configured so that they create a whole series of tiny electrically insulated boxes, which act as memory cells. By applying a small electrical voltage at the top of one of these boxes, additional electrons can be attracted into it, giving it an overall negative charge. Because the box is insulated, the electrons are trapped there, even when the power is switched off. In this state, the insulated box has a bit value of 1. If there are no additional electrons in the box, it has a bit value of 0. A memory cell can be reset to 0 by forcing the additional electrons to flow out of the box by using an electrical voltage in the other direction.

“Unfortunately, though, continual reading and writing of the value of a box causes the insulation enclosing it to degrade over time – and when that happens, the box starts to get ‘leaky’, so it can’t hold the negative charge. This means that it cannot distinguish between a 0 or a 1 any longer, which means it becomes unreliable.”

Rupert thought about this for a moment. “I understand all of that, but don’t spinning disks wear out too?”

“Yes”, said Gloria, “but in a different way. In terms of actual data storage, a hard drive can tolerate millions of read/write cycles, but of course it is prone to mechanical wear. An SSD has no moving parts to break down, but it will endure far fewer cycles – say, between a few thousand to a few hundreds of thousands, depending on the type of flash memory.”

“Oh,” said Rupert, obviously relieved. “I thought that I was in trouble for a moment – I have all of my trainspotting notes on my USB flash drive – I would hate to lose those.”

“I don’t think that you have to worry too much about USB flash drives. Generally speaking, they are used for temporary storage or back-up, and it would take a very long time indeed to wear one out. However, using an SSD in a desktop or laptop computer is a different matter. That is why files are saved in a different way on an SSD drive than on a HDD.”

6.3 Writing data to an SSD

“In order to prevent the ‘leaky’ memory cells from occurring, the software that controls where files are saved tries to prevent the same block in semiconductor memory from being used repeatedly. It does so by ensuring that, every time a file is modified, it is written to different physical locations across the whole of the SSD, selected at random. This technique is called **wear levelling**,” explained Gloria.

“So, SSDs are *designed* to be fragmented?” asked Rupert, slightly incredulously.

“Yes! So the good news for a user is that there is no need to ever defrag an SSD – in fact, even if it could be done, it would be counterproductive because an SSD is designed to take a scattergun approach to storing data across the whole drive. However, the bad news for private detectives is that it is much more likely that a previously ‘deleted’ file will be partially overwritten and that it is much harder to piece together a file if the locations of the constituent clusters are lost.

“And as if this is not bad enough, a memory cell can only be written to when it is empty. So before a file is written to a cluster, the entire contents of that cluster are erased so that the write operation can then be carried out. This slows down write operations considerably. So, to counteract this, many manufacturers provide SSD control software, called **TRIM**, which uses any quiet time to ‘garbage collect’ any unreferenced file fragments caused by wear levelling and proactively initialise these, and other unallocated blocks of the SSD memory, ready for writing. That really is a disaster for digital forensics because, once these areas of memory have been initialised, the data they once contained is lost forever.

“That is terrible!” cried Rupert. “And it gets even worse,” said Gloria. “Unlike a traditional hard drive, just powering up an SSD, even one that has been removed from the computer, can allow the garbage collection routines on the SSD to continue. This means that the data on an SSD drive is continuously altered, making it difficult to use as evidence.”

“Can’t the garbage collector routines be disabled?” asked Rupert.

“Unfortunately not; there is even secrecy around the algorithms for rewriting files and wear levelling, which makes it even more difficult for us to know how to find deleted files. There

are standards emerging to try to ensure that access to the flash memory is tenable but, as yet, there is no requirement for manufacturers to follow them.

“All of this has led to some people saying that SSD is ‘the beginning of the end’ for current practice in digital forensic recovery,” continued Gloria.

“That sounds a bit melodramatic!”

“Yes, but they may be right,” replied Gloria. “We understand a lot about the forensic analysis of hard drives but, more importantly, we also know how to present the evidence so that it is legally acceptable. Now it is like starting again for solid-state drives.”

Activity 18 (self-assessment) Deterioration in the speed of writing to flash memory

The erase operation on an SSD is relatively time-consuming. Use this fact to explain why writing a file to a brand new USB flash drive is very quick but, after a period of use, writing becomes a much slower operation.

Answer

A brand new USB flash drive does not require any blocks to be erased before writing to it for the first time. However, as the SSD fills up, any previously written to blocks need to be erased first before being written to again – this slows down the write operation considerably.

6.4 Deleting data from an SSD

“What about deleting data on an SSD?” asked Rupert.

“Well, of course, you can still physically destroy the drive, but degaussing does not work because SSDs do not rely on magnetism to store zeros and ones. However, most SSD manufacturers have a utility for managing and securely erasing their SSDs using a command called ATA Secure Erase. This works in a similar way to that of the TRIM command that garbage collects memory when a file is deleted, except the ATA Secure Erase command resets the whole of the SSD by applying a spike of voltage to all of the memory cells simultaneously, flushing out all of the stored electrons and forcing the drive to ‘forget’ all of its data.”

6.5 Summary of Section 6

In this section, the story of Gloria and Rupert continued. Their attention shifted to solid-state drives. You learned that the use of SSDs is steadily increasing, in contrast to HDDs, which are in decline. You also learned that SSDs are designed to be fragmented and this is one of the reasons why it is much harder to piece together a previously ‘deleted’ file – a nightmare for digital forensics.

Conclusion

This course introduced several aspects of our increasingly digital lives. You learned about the ever more significant role that digital technologies are playing in our information society, in areas ranging from entertainment to public services. Along with the opportunities available in the online world, you were also introduced to some of the potential problems and how to avoid them.

You looked at a case study about online information on childhood vaccination. Finding information online via web searches turned out to be a truly delicate matter. Depending on your search terms you may be more likely to find myths than facts or vice versa.

Finally, you went on to look at how data is stored on both hard disk drives (HDDs) and solid-state drives (SSDs). You studied the differences from a digital forensics point of view. You examined what it means for data to be written to and deleted from an HDD and if and how 'deleted' information can be retrieved again. You then considered the same questions for SSDs.

This OpenLearn course is an adapted extract from the Open University courses [TM111 Introduction to computing and information technology 1](#) and [TM112 Introduction to computing and information technology 2](#).

Glossary

ambient data

The data in slack space, also known as *latent data*.

automated teller machine (ATM)

A machine that allows bank customers to perform certain transactions, such as withdrawing cash from their bank account. Also known as a *cash machine* or *cashpoint*.

biometric

A term that refers to the identification of people using biological characteristics such as fingerprints, iris recognition and DNA analysis.

cable

A set of wires or optical fibres assembled, with a protective coating, for use as a communication medium.

cluster

An amount of memory on a disk drive that consists of a fixed number of sectors. Commonly a cluster contains 4 or 8 sectors. As a file is written to the disk, the file system allocates the appropriate whole number of clusters to store the file's data.

computer forensics

see digital forensics

copyright

A legal protection that guarantees that the creators of content are rewarded for their work and protects the rights of users.

database

A set of computer-based data that has been organised so that it can be read, written, updated and searched. For example, a library catalogue.

defragment (or defrag)

(Also called *defragmentation*) An operating system process for optimising the physical location of files on a disk. As a disk is used, files are increasingly stored in fragments that can be scattered over the surface of the disk. Such fragmentation increases the length of time taken to read files. Defragmentation joins file fragments together so that a file is stored in continuous memory.

degaussing

The process of applying a strong magnetic field to a magnetic storage device in order to erase any information held on that device.

digital forensics

(Also called *forensic computing*.) A branch of forensic science that is concerned with obtaining legal evidence from computer systems.

Digital rights management (DRM)

A range of technologies used by copyright owners to control how the content they produce is used.

digital technology

Any technology that is based on representing data as sequences of numbers, i.e. as digital data.

emoticons

Text characters or images that indicate someone's mood by representing a simple facial expression. Text-based emoticons such as :-) are usually intended to be 'read' by tilting your head to one side. Also known as *smileys*.

FAT32

A file system developed for Windows that is now mainly used by USB flash drives.

file allocation table (FAT)

A table in a FAT32 file system that holds information about where a file is stored on a hard disk or in an SSD.

file system

The method by which an operating system controls how data is stored on and retrieved from the hardware.

flash memory

A type of non-volatile memory of which its contents can be erased and rewritten.

formatting

The act of preparing a computer storage device for use by a file system. Any data previously held on the device is made unavailable to the operating system and the device appears empty.

forum

An area on the internet designed for discussion, usually on a specific topic. Also known as a *bulletin board system (BBS)*.

fragmentation

An unwanted process in HDDs where computer files are split across a number of physical locations on a disk as the disk becomes full. Fragmentation slows down file access.

global positioning system (GPS)

A set of satellites that continuously transmit their position so that anyone with a suitable receiver can obtain very accurate positioning information.

hard disk

(Also called *HDD* and *hard disk drive*.) A data storage device used for storing and retrieving digital information using one or more rigid, rapidly rotating disks (platters) coated with magnetic material.

hard disk drive (HDD)

(Also called hard disk.) A data storage device used for storing and retrieving digital information using one or more rigid, rapidly rotating disks (platters) coated with magnetic material.

hardware

The physical components of a computer system. These include the large components such as the screen, the small component parts such as circuit boards, and all of the connecting cables.

information society

A term used to describe the social and economic changes related to the development and widespread use of information technologies.

instant messaging (IM)

Text-based communication between people who are online at the same time.

internet

The global internetwork that has grown from a US government-funded project started in the 1960s.

knowledge society

A society based on the acquisition, use and dissemination of information, made possible through advances in computing and IT; one where human knowledge is important for the economy.

latent data

(Also known as *ambient data*.) The data in slack space.

logical size (of a file)

The size of the content of a file. (Compare with *physical size*.)

Master File Table (MFT)

A table in a New Technology File System (NTFS) that holds information about where a file is stored on an HDD or SSD.

memory

The part of a computer system that stores programs and data while they are waiting to be executed by a central processing unit.

moderator

A forum user who has responsibility for managing the forum and 'moderating' discussions to ensure that they follow the forum rules.

netiquette

A set of guidelines for online behaviour.

network

A collection of devices that can communicate with each other. Networks vary in size and complexity, connecting anything from a few devices to many millions.

network society

A term that is sometimes used interchangeably with 'information society', but which emphasises how the flow of information depends on networks.

New Technology File System (NTFS)

A file system commonly used in Windows computers.

online discussion group

A group of people, often with common interests or aims, communicating over the internet.

operating system (OS)

A collection of programs that manages a computer's resources, provides an interface between the user and the computer, and organises the running of other programs. Examples include Windows, Mac OS and Linux.

overwriting (also called **wiping or **shredding**)**

The process of obscuring sensitive information that might remain on a disk before disposal. New data, usually just all zeros or all ones but sometimes a random pattern of both, is written to every part of the disk. This is a time-consuming process, especially if multiple overwrites are used.

partitions

The sections of a hard disk. When you format a hard disk, you can usually choose the number of partitions you want. Partitions are then seen as separate disks by the operating system.

physical size (of a file)

The size of the disk storage that is allocated to store a file, which will always be a whole number of clusters. (Compare with *logical size*.)

plagiarism

Using the work of other people to gain some form of benefit without formally acknowledging that the work came from someone else.

program

A step-by-step set of precise instructions for telling a computer how to carry out a particular task.

public domain

A term referring to any created content that is not subject to copyright. Material that is 'in the public domain' may be used freely by anyone.

Rolodex

A type of physical desktop card index for storing data, invented in the 1950s.

Sector

The smallest physical storage unit on a hard disk. In modern file systems it is almost always defined to be 512 bytes in size.

slack space

The unused space in a disk cluster. The difference between the physical size and the logical size of a file.

smart device

An electronic device that processes information and exchanges it with other devices.

SMS

The 'short message system' that allows text messages to be sent between mobile phones. SMS messages can also be sent between other devices, such as computers and landline phones.

social networking

The activities involved in building and maintaining online relationships and communities.

software

The programs that control the functioning of a computer system.

solid-state drive

(Also called *SSD*.) A storage device that uses integrated circuit assemblies as memory to store data persistently. (Compare with *hard disk drive*.)

SSD

(Also called a *solid-state drive*.) A storage device that uses integrated circuit assemblies as memory to store data persistently. (Compare with *hard disk drive*.)

TRIM

Control software that allows an operating system to inform a solid-state drive (SSD) which blocks of data are no longer considered in use and can be wiped internally.

ubiquitous

A term describing something that seems to be everywhere at the same time.

unallocated space

The space on a hard drive that is available for the operating system to write to.

virus

Malicious software that is designed to attack software on users' computers, spreading quickly and easily from one computer to another.

wear levelling

A technique for prolonging the service life of some kinds of erasable computer storage media, such as flash memory, which is used in solid-state drives (SSDs) and USB flash drives. It does so by ensuring that, every time a file is modified, it is rewritten to a different physical location across the whole of the SSD, selected at random.

world wide web

An internet service that links computer files such as documents, images, audio and video. These files may be located on any computer connected to the internet. Also known as the *web*.

References

Associated Press (2012) 'Mobile carriers face a future without text messaging as data apps take over', *The Guardian*, 1 March [Online]. Available at www.theguardian.com/technology/2012/mar/01/sms-pinger-whatsapp-messaging-mwc . (Accessed 4 June 2020).

Demicheli, V., Rivetti, A., Debalini, M. G. and Di Pietrantonj, C. (2012) 'Vaccines for measles, mumps and rubella in children', *Cochrane Database of Systematic Reviews*, no. 2 [Online]. DOI: 10.1002/14651858.CD004407.pub3 (Accessed 27 July 2017).

Evans, B. (2015) 'WhatsApp sails past SMS, but where does messaging go next?', Benedict Evans, 11 January [Blog]. Available at <http://ben-evans.com/benedictevans/2015/1/11/whatsapp-sails-past-sms-but-where-does-messaging-go-next> (Accessed 4 June 2020).

General Medical Council (2010) *Dr Andrew Jeremy Wakefield: Determination on Serious Professional Misconduct (SPM) and sanction* [Online]. Available at https://web.archive.org/web/20110809092833/http://www.gmc-uk.org/Wakefield_SP-M_and_SANCTION.pdf_32595267.pdf (Accessed 11 October 2019).

International Telecommunication Union (2006) *ITU Internet Report 2006: digital.life* [Online]. Available at www.itu.int/osg/spu/publications/digitalife.

Maglione, M. A., Das, L., Raaen, L., Smith, A., Chari, R., Newberry, S., Shanman, R., Perry, T., Goetz, M. B. and Gidengil, C. (2014) 'Safety of vaccines used for routine immunization of US children: a systematic review', *Pediatrics*, vol. 134, no. 2, pp. 325–37 [Online]. DOI: 10.1542/peds.2014-1079.

Ofcom (2015) *Adults' Media Use and Attitudes: Report 2015* [Online]. Available at https://www.ofcom.org.uk/_data/assets/pdf_file/0021/149124/adults-media-use-and-attitudes-report.pdf.

Ofcom (2016) *Telecommunications Market Data Update Q1 2016* [Online]. Available at <https://www.ofcom.org.uk/research-and-data/telecoms-research/data-updates/q1-2016>.

Office for National Statistics (2019) *Internet access – households and individuals, Great Britain: 2019* [Online]. Available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/home-internetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2019#main-points>

Petrescu, P. (2014) *Google Organic Click-Through Rates in 2014* [Online]. Available at <https://moz.com/blog/google-organic-click-through-rates-in-2014> (Accessed 4 June 2020).

Ruiz, J. B. and Bell, R. A. (2014) 'Understanding vaccination resistance: vaccine search term selection bias and the valence of retrieved information', *Vaccine*, vol. 32, no. 44, pp. 5776–80 [Online]. DOI: 10.1016/j.vaccine.2014.08.042.

Taylor, L. E., Swerdfeger, A. L. and Eslick, G. D. (2014) 'Vaccines are not associated with autism: an evidence-based meta-analysis of case-control and cohort studies', *Vaccine*, vol. 32, no. 29, pp. 3623–29 [Online]. DOI: 10.1016/j.vaccine.2014.04.085.

Acknowledgements

This free course is based on extracts from the Open University courses [TM111 Introduction to computing and information technology 1](#) and [TM112 Introduction to computing and information technology 2](#). It was edited by Elaine Thomas and Paul Piwek. It was first published in June 2020.

Sections 1.1, 1.2 and 1.3 of TM111 Block 1 Part 1 (Author: Elaine Thomas).

Introduction, Section 1.1 and 1.2 of TM112 Block 3 Part 1 (Author: Lindsey Court)

Section 6.3 (up to and including 6.3.1) of Block TM112 Block 3 Part 6 (Author: Mike Richards).

Except for third party materials and otherwise stated (see [terms and conditions](#)), this content is made available under a

[Creative Commons Attribution-NonCommercial-ShareAlike 4.0 Licence](#).

The material acknowledged below is Proprietary and used under licence (not subject to Creative Commons Licence). Grateful acknowledgement is made to the following sources for permission to reproduce material in this free course.

Images

Figure 2: Grace Hopper beside a 1950s mainframe computer: Taken from <http://media-cache-ec0.pinimg.com/736x/3f/6c/35/3f6c35695014f2741197107b7e536fa2.jpg>

Figure 3: © Jim Zuckerman/Alamy Images

Figure 4: Taken from https://www.ofcom.org.uk/__data/assets/pdf_file/0024/83166/adults_media_use_and_attitudes_charts_-_section_3.pdf

Figure 5: © www. xkcd.com / Creative Commons Attribution-Non-commercial-ShareAlike Licence <http://creativecommons.org/licenses/by-nc-sa/2.5/>

Figure 6: © www. xkcd.com /This file is licensed under the Creative Commons Attribution-Noncommercial-ShareAlike Licence <http://creativecommons.org/licenses/by-nc-sa/3.0/>

Figure 7: Taken from <http://www.clipartsuggest.com/images/562/clip-art-gap-generation-gap-5LCMhE-clipart.jpg>

Figure 10: © www.xkcd.com / This file is licensed under the Creative Commons Attribution-Noncommercial-ShareAlike Licence <http://creativecommons.org/licenses/by-nc-sa/3.0/>

Figure 13: © Image Source / Getty Images

Figure 15: © Sashkin / www.shutterstock.com

Figure 16: Taken from <https://www.pctechguide.com/hard-disks/hard-disk-hard-drive-construction>

Figure 17: Taken from http://sherylcanter.com/articles/pcmag_20031001_DataRecovery.php

Figure 19: Taken from <http://www.datadev.com/degausser-hd2.html?viewfullsite=1>

Figure 21: Taken from <http://www.zdnet.com/article/ssds-set-to-grab-over-one-third-of-pc-storage-solutions-market-by-2017-ihs/>

Every effort has been made to contact copyright owners. If any have been inadvertently overlooked, the publishers will be pleased to make the necessary arrangements at the first opportunity.

Don't miss out

If reading this text has inspired you to learn more, you may be interested in joining the millions of people who discover our free learning resources and qualifications by visiting The Open University – www.open.edu/openlearn/free-courses.