UNIT 1

INTRODUCTION

VCE Unit 1 of *Computing* looks at how individuals and organisations use, and can be affected by, information and networked digital systems in their daily lives.

Throughout the unit, students will apply the design and development stages of the problem-solving methodology. They will acquire and apply the knowledge and skills to work with different data types to create solutions that can be used to persuade, educate, inform and entertain.

This unit also examines the role of networked information systems in the communication of data within a global environment and an exploration of mobile devices.

Several issues relating to the effect of information systems on students themselves are also examined and students are required to work collaboratively to examine these issues.

There are three outcomes to be completed in Unit 1.

AREA OF STUDY 1: DATA AND GRAPHIC SOLUTIONS

Outcome 1

You will collect your own data and information, and design and develop solutions that meet specific purposes using software to create a graphic solution. The information you produce should be in graphic form. You are only required to apply the design and development stages of the problem-solving methodology, but you should undertake your own investigation.

AREA OF STUDY 2: NETWORKS

Outcome 2

You must propose a networked information system with wireless capability for a specific purpose, and explain the security threats that exist within the networked information system. You must also explain the configuration of the network and suggest the risks and benefits of its use for intended users, including potential legal requirements and ethical responsibilities. Throughout this Outcome, you will explore the exchange of data and information within a networked information system. You will also learn the use of mobile devices within networks, and how the security of data and information exchanged within a network can be threatened.

AREA OF STUDY 3: COLLABORATION AND COMMUNICATION

Outcome 3

You will work in teams to design and develop a website that analyses a chosen contemporary issue and supports your team's point of view. In analysing an information systems issue, you will consider the tensions and conflicts between different stakeholders and then, using visualising thinking tools, explore your own opinions. You will manipulate acquired primary and secondary data and, optionally, develop graphical representations that form part of your website. You will also engage in project management and use digital systems to form and monitor plans.

CHAPTER 1

DATA ANALYSIS

Key knowledge

After completing this chapter, you will be able to demonstrate knowledge of:

Data and information

- types and purposes of qualitative and quantitative data
- sources of, and methods and techniques for, acquiring and referencing primary data and information
- factors affecting the quality of data and information such as relevance, accuracy, bias and reliability
- techniques for authorising the collection and use of data and information such as using consent forms
- techniques for protecting the privacy of the providers of data and information such as de-identifying personal data

Digital systems

- physical and software controls used to protect the security of stored data such as backing up, usernames and passwords, systems protection software and encryption

Interactions and impact

- Australian Privacy Principles relating to the acquisition, management and communication of data and information, including non-identification of individuals (Principle 2), information only being held for its primary purpose (Principle 6)
- ethical dilemmas arising from data acquisition strategies.

For the student

Students will conduct an investigation into an issue, practice or event and collect primary data, interpret and manipulate this data into a graphical solution to represent their findings.

For the teacher

This chapter is based on Unit 1, Area of Study 1 and, together with Chapter 2, provides the key knowledge required to complete Unit 1, Outcome 1. At the end of Chapters 1 and 2, students should be able to acquire, secure and interpret data and design and develop a graphical solution that communicates their findings of an investigation.

Shutterstock.com/Rawpixel

Understanding research

Some people consume research, and others produce research. Consumers of research spend a lot of time reading other people's research rather than conducting their own. On the other hand, producers of research investigate or explore an area that has relevance to them, interpret their data and then communicate their findings.

Producers of research often start with a theory or a research question, from which they generate a **hypothesis**, test the hypothesis and then interpret the results, as can be seen in Figure 1.1.

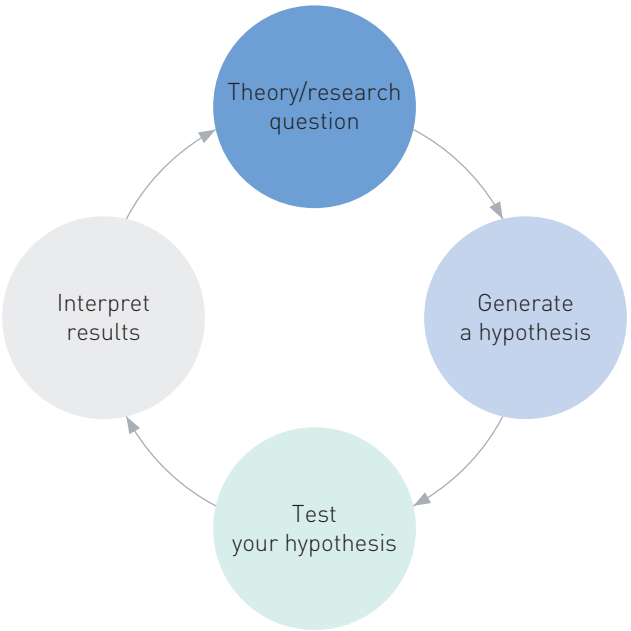


FIGURE 1.1 Producers of research often start with a theory or a research question; then they generate a hypothesis, test the hypothesis and interpret the results

Theories are usually general statements that describe something, provide an explanation of why something happens and can be applied to predict what will happen in the future. Theories are in principle falsifiable or disprovable; that is, they contain information about the sorts of events that, if they were to happen, would show the theory to be false. Some research questions are tied closely with theories. Research questions assist researchers to narrow the focus of the topic of investigation. For example, 'Do science and innovation boost our standard of living and contribute to economic growth?'

A scientific theory summarises a hypothesis that has been supported with repeated testing. A theory is valid as long as there is no evidence to dispute it.

Hypotheses, on the other hand, are based on probabilities about what will happen according to the applied theory. Theories are tested by using data collection tools such as questionnaires and/or interviews, and then the results of the study will either confirm or disprove the hypothesis.

Types of research

An approach to investigate the topic of interest is through quantitative or qualitative research.

Quantitative data is measurable and specific and therefore easier to chart or graph. At a simplistic level, quantitative data gathering is based on verifying theory through the use of statistics and largely numerical data, while qualitative data provides a more in-depth understanding. An example of quantitative data is:

> Of the teachers who teach mathematics to secondary school students (Years 7 to 10), 61 per cent have studied mathematics at university to at least second-year level.
>
> Benchmarking Australian Science, Technology, Engineering and Mathematics, Office of the Chief Scientist, November 2014, p.98

SPSS and MiniTab are statistical software packages used to analyse quantitative data.

When data has been gathered using surveys, focus groups, observation or other methods, quantitative data can be analysed by using software such as Excel, the Statistical Package for the Social Sciences (SPSS) and Minitab. This takes time and often involves hours of data entry, depending on the complexity of the data gathering instrument. For simple data gathering, online surveys such as SurveyMonkey allow users to create surveys and manage the collection and analysis of quantitative data. Most online survey software also permits **qualitative data** to be entered.

Qualitative data is harder to measure than quantitative data. You can gather qualitative data using instruments such as interviews, focus groups, video footage and observation. Generally, qualitative data needs to be recorded accurately and transcribed at a later stage. The analysis of the qualitative data is also quite different from that of quantitative data. With quantitative data, the researcher looks for themes or patterns through the use of numbers, while with qualitative data, the researcher establishes rich descriptions and finds themes through reading the text and classifying these themes. An example of qualitative data gathering is more descriptive:

NVivo is a qualitative data analysis computer software package. It has been designed for qualitative researchers working with very rich text-based and/or multimedia information, where deep levels of analysis of data are required.

> From 2003 to 2012, the mean PISA scores in mathematics declined, while the number of countries performing better than Australia increased. PISA also shows that Australian students' proficiency in mathematics is declining, with the proportion of low performers rising and the proportion of top performers falling.
>
> Benchmarking Australian Science, Technology, Engineering and Mathematics, Office of the Chief Scientist, November 2014, p.100

Advantages and disadvantages of quantitative and qualitative data

Participants are more willing to be part of a quantitative study as it is less demanding of them. Often, quantitative studies use questionnaires, which can capture a large sample size. Having a large sample size provides statistical validity, and helps to accurately reflect the population. Data is then interpreted, relationships identified and findings are then communicated. Conversely, because questionnaires do not have the provision to probe the participants further, the answers provided do not have as much depth and are at times superficial. If too much information were provided, researchers would be overwhelmed by the amount of data collected and would not be in a position to analyse it.

Qualitative research provides for rich, in-depth study of participants. Researchers can ask further questions, especially if something of interest arises. Generally, qualitative studies are small, and provide a narrative description of a sample group. Data gathering tools can include interviews and focus groups. However, because the sample size is small and the sample is not very random, conclusions may not generalise to a larger sample size; therefore, findings may be peculiar to a particular sample.

Data and information

Technically speaking, a datum is a single item of data; however, the term 'data' is commonly used and accepted as both the singular and plural forms of the word.

Information systems focus on both the transformation of **data** into **information** and the management of that information. Critical to using digital systems to solve information problems is an understanding of how data can be input to a computer and then manipulated to create meaningful information. The term 'data' refers to the raw, unorganised facts, figures and symbols fed to a computer during the input process. Data can also mean ideas or concepts before they have been refined. In addition to text and numbers, data also includes sounds, and images (still and moving).

Information is produced when data is manipulated by the computer's processor into a meaningful and useful form, thus becoming information. This can be achieved by organising the data and presenting it in a way that suits the needs of the intended audience. The information produced can be used to inform, entertain or persuade an audience.

Gathering data: Primary data and information

Sources

Primary sources are usually the **stakeholders** in a particular issue – the topic that you are investigating as part of your hypothesis. To question them or survey their opinions can provide different insights and often more in-depth data than information from **secondary sources**. The data will often be more up-to-date and can provide more unusual and important insights into issues, especially at the immediate local level, than secondary sources can, which often present overall conclusions and general summaries. When data is collected, often by non-stakeholders, they frequently use observation and measurement.

Techniques and methods

Collecting data from the stakeholders directly is usually conducted through methods such as surveys using questionnaires, and interviews. While the results of questionnaires are easy to present graphically, interview results often can only be presented as written summaries and conclusions. However, both require analytical discussions to interpret their meaning.

A **questionnaire** is usually a set of questions that ask for a response to be selected from a list of alternatives, such as A, B, C, D; or a range, either 1–5 or very low to very high. Such questionnaires can easily be given to many people, and are easily processed and analysed using computer-based methods since the answers are able to be recorded as numbers.

Another type of questionnaire provides space for short, focused free-form answers similar to those obtained from an interview about particular aspects of an issue; for example, 'Describe the feelings you have when you are playing your favourite computer game' (in an interview, this question would be followed up with 'Why?' or 'What part of the game causes these feelings?'). However, because of their free form, which does not lend itself to being recorded numerically, these answers tend to be more difficult to analyse.

Interviews are usually conducted face to face (technique), sometimes in groups, and can take a substantial amount of time. A major feature of an interview is the opportunity for in-depth follow-up and clarification questions that cannot be done with questionnaires, which are often answered in private. Interviews are very useful for eliciting the feelings, attitudes and opinions of people that are too complex to easily record in a questionnaire.

Other ways to collect data electronically include sensors such as traffic cameras, satellites and online sources, such as websites or data logs. The data collected can also be used for a variety of purposes, including describing, predicting and improving processes within an organisation, or for research.

ISSUE

MH370 search: How new satellite data confirmed Malaysia Airlines plane was lost

[25 March 2014, Nick Miller, Europe correspondent] Sydney Morning Herald

British satellite company Inmarsat analyses seven, hourly pings sent by the missing Malaysian Airlines flight to determine its final resting place.

London: A new satellite tracking technique is what gave Malaysian Prime Minister Najib Razak enough confidence to announce that Malaysia Airlines flight MH370 went down in the remote south of the Indian Ocean.

British firm Inmarsat was behind an earlier analysis that indicated the plane had been flying in one of two big 'corridors', one in the northern hemisphere and one in the southern.

However last week it went back to its data and tried a new mathematical analysis, which concluded on Sunday.

The new analysis allowed them to discard the northern corridor, and focus more precisely on the southern route.

Based on this new information, Mr Najib announced on Monday that MH370's last known position was in the middle of the Indian Ocean, west of Perth.

UK firm Inmarsat was behind an earlier analysis of the path of missing flight MH370.

The nature of the pings indicated that the plane was still moving during that time.

'This is a remote location far from any possible landing sites,' he said. 'It is therefore with deep sadness and regret that I must inform you that … flight MH370 ended in the southern Indian Ocean.'

The aeroplane had Inmarsat's 'Classic Aero' satellite system, which collects information such as location, altitude, body heading and speed, and sends it through Inmarsat's satellites into their network.

This 'ACARS' (aircraft communications addressing and reporting system) was switched off or interrupted early in the flight, meaning no such information was available to track the plane.

However the Classic Aero system still sent hourly 'pings' back to Inmarsat's satellite for at least five hours after the aircraft left Malaysian airspace, the company discovered.

These pings contained no data – they were just a simple 'hello' to keep the link open – however, their timing and frequency contained hidden mathematical clues.

The company looked at the 'Doppler effect' – tiny changes in the frequency of the ping signal, caused by the relative movement of the satellite and the plane (the Doppler effect is the reason why, for example, police sirens are a different pitch or frequency depending on whether they are travelling toward you or away from you).

This analysis allowed Inmarsat to map two huge 'corridors' for the plane's possible location, in big arcs stretching thousands of kilometres north and south of the point where the last radar contact with MH370 was made.

Australian and US experts took this information, added some assumptions about the plane's speed, and narrowed the southern option into an area of ocean that could be realistically searched.

Meanwhile, Inmarsat went back to its satellite data. Its new analysis found that the northern route did not quite correlate with the frequency of the pings from the plane – meaning the plane must have been heading south.

It also suggested that the plane had been travelling at a steady cruising altitude above 30,000 feet.

They compared satellite data from MH370 with that from previous Malaysian Airlines Boeing 777 flights, going back a few weeks, in order to better model the movement of the plane.

'This really was a shot in the dark,' Chris McLaughlin, senior vice president of external affairs at Inmarsat told the BBC. 'It's a credit to the scientific team that they managed to model this.

'Just a single "ping" can be used to say the plane was both powered up and travelling. And then by a process of elimination comparing it to other known flights and established that it went south.'

The UK's Air Accidents Investigation Branch also contributed to the analysis.

THINK ABOUT COMPUTING 1.1

1 List the various data sources used to search for the missing plane.

2 Why is there a need to use a variety of data sources to assist with locating the missing plane?

3 How has the data assisted with estimating the location of the plane?

Data collection methods

Before we can produce information, we first must start with data. Data collection methods such as surveys/questionnaires, interviews or observation provide a means of capturing data.

Surveys and questionnaires

Surveys and questionnaires are common methods used to collect data. They can provide data about what the respondents think is true, or their preferences for consumer goods and political parties. A questionnaire can be a quick way of gathering large amounts of data. Questionnaires and surveys need to be carefully designed, otherwise the participants' responses may not provide suitable data to analyse, rendering them useless. Questions used in a survey must be carefully worded so that the response will provide meaningful and useful data without the need for further clarification.

Focus groups

A **focus group** is the meeting of a small group of individuals who are guided through a discussion by a researcher. The focus group is carefully selected, so it fits a particular demographic and the researcher can obtain the necessary data through a guided discussion that probes the participant's attitudes about the topic. Focus groups often comprise five to 12 people and the discussion is loosely structured to encourage ideas to flow.

Interviews

Interviews are used to elicit the opinions and beliefs of people. They can be used to gather data for research projects. Interviews are usually conducted one-to-one in a quiet, relaxed atmosphere. They should be recorded, with permission of the interviewee, with easily used and unobtrusive audio equipment or video. Writing down the responses during the interview is not helpful to the interviewer or the interviewee. Collating and analysing information can be difficult and time-consuming, and may require the use of someone with expertise.

Open-ended and closed questions

Questions used on a survey and during an interview can be open-ended or closed. **Closed questions** limit the responses available to the respondent (Figure 1.2). They include 'Yes/No' boxes, multiple-choice questions, and scales on which the attitudes and beliefs are measured, such as 'strongly agree', 'agree', 'disagree' or 'strongly disagree'.

Closed (or closed-ended) questions are generally considered to be quantitative in nature. They are called 'closed' because the range of answers the participant can choose is limited. Closed questions are also known as quantitative, as the response options can be converted to numbers. For example:

How often do you feel that you are overworked with homework?
5 I always feel overworked
4 I sometimes feel overworked
3 I occasionally feel overworked
2 I feel overworked once in a while
1 I never feel overworked

Each of these options can have a value placed next to them. However, we do not talk in numbers and we shouldn't create surveys that only have numbers. Questionnaires and surveys should be thought of as a conversation between the person asking the questions and the person answering them.

Open-ended questions do not limit the answers that can be given by the respondent (see Figure 1.3). They should be worded so that the responses received are capable of correct interpretation. For instance, if you asked the question, 'How do you feel about the widespread use of computer games?', the responses would probably be too broad to be usefully categorised and analysed. The wording must therefore limit the scope of the possible responses to specific areas of interest: 'How has the playing of computer games affected the school results for your children?' Open-ended questions also allow for follow-up questions, which are called probing questions, such as 'Why?' or 'Please give an example'. Such questions tend to elicit more detail.

Open-ended questions allow people to answer the question as they want to. They are called 'open-ended' because participants are free to answer in any manner they choose. Unlike closed questions, there are no response options specified. They are qualitative because responses are considered and measured by feel rather than by numbers.

Closed questions are easier to develop, quicker to administer and answer, easier to collate and analyse, and can provide a large and balanced sample; however, they may not be useful for complex issues. In this case, open-ended questions may be needed as they elicit greater detail in the responses, can bring forth unusual ideas and can show links between various aspects of the issues.

Closed questions
1 How long have you shopped at this store?
2 How many times per week do you go shopping?
3 How much do you spend per week?
4 Which of the following sources of information most influences your purchasing habits?
• Advertising pamphlets delivered to the home
• Newspaper advertisements
• Television promotions
• Recommendations from friends
5 Do you use a computer? • Yes • No
6 What time of day do you normally go shopping?

FIGURE 1.2 Closed questions should be designed to elicit short, straightforward answers

Open-ended questions
1 What is your opinion of the games available from this store?
2 How influential do you think the advertising campaign has been?
3 What are some of the errors in data entry that you have observed?
4 Describe the most frustrating experience you have had when using the computer system.
5 What are some of the problems you experience in receiving information on time?
6 What changes would you recommend to improve the billing system?

FIGURE 1.3 Open-ended questions try not to limit the answers the respondent can give

Observation

Observation is a way of finding out about the world around us. Using our senses, such as sight, smell and hearing, we are able to pick up detailed information about our environment. However, as a method of data collection, observation is more than just looking or listening, as we can be selective about what we perceive to be most useful to us. Researchers engaged in observation attempt to learn what life is like for someone in a particular setting, while they remain an outsider. While observing, they make careful notes of what they see, record all accounts including conversations and interactions. Observation generally takes place in community settings such as classrooms or locations believed to have some relevance to the research questions. Observation is unlike other forms of data collection tools, as the researcher approaches participants in their own environment rather than having the participants come to the researcher.

Quality of data and information

One aspect of identifying relevant data from a given data set is ensuring that the data is usable. To be usable, data must be relevant, accurate, free from bias and reliable.

Relevance

To produce usable information, data must be relevant. For example, if a computing department in an organisation is evaluating PC-only software, then surveying people who only use a Mac is irrelevant. The data collected from Mac users would not be relevant to the overall data collection.

Data also needs to be processed while it is current, because decision making should not be based on outdated data.

Accuracy

Data that is entered into a computer must be accurate. Transcription is often a cause of error. Transcription errors occur when the person entering the data misreads the information through, for example, a lapse in concentration, being interrupted or pressing the wrong key. It is easy to make a mistake when entering a large amount of data, particularly numbers with many digits that may not contain spaces or punctuation to signify thousands. Clearly, if the data collected is incorrect, the information produced will be incorrect. If data has been gathered from a primary source, it is a good idea to check it against this. If data has been gathered from a secondary source and is suspect, it is worthwhile verifying the data using other secondary sources.

Freedom from bias

Bias can easily creep into data and make the information processed from it unreliable. Several influences can result in the introduction of bias into data: namely, vested interest, timing, small sample size, bias through sorting and bias through graphic representations.

Vested interest

Bias can enter data if the respondent to a survey or interview has a **vested interest** in the outcome of the research. A common example is celebrities who are paid to promote particular products in commercials or social media. It would be unreasonable to trust their statements that one product is better than others purely based on the fact that they are celebrities; they are only saying what they have been paid to say and may not necessarily be providing an independent judgement that has been derived from research or experience.

Timing

The timing of the data collection may also introduce bias. For example, you plan to survey a sample of the population for their views about Australia becoming a republic. The data you gather may be biased if, just prior to the survey being conducted, a royal tour takes place and there is extensive media coverage about the Royal Family. The timing of the data collection would introduce bias because it coincides with a significant event that could influence the responses.

Note too that bias is not restricted to data gathered from surveys or during interviews. For example, suppose that Qantas needed to decide whether to schedule two new weekly flights to New York. The decision could depend on the demand for existing flights. If the airline collected data from bookings made over a four-week period just before or during a significant event, the data gathered would be biased. Such data should not be relied on for making this decision because the influence of this event on customer demand is irregular and unlikely to occur again.

Small sample size

Choosing a sample size that is too small may also create bias. The sample size and composition must be suitable for the purpose of the data collection and, usually, a larger sample size leads to greater precision, provided the sample composition is suitably representative of the target population. The sample size must be big enough to make any conclusions drawn and information produced credible. For example, if you wanted to determine whether or not the school uniform should be changed, it would be remiss to only survey students in your class. Not only would this sample not be representative of the student body, but it would also not include other stakeholders, such as parents and school administrators. Similarly, if you wanted to gather sales data over a four-day period to predict monthly sales at a fish and chip shop, this time scale would not be sufficient to make a prediction. For instance, by choosing the four Mondays in the month, you may be selecting the quietest trading days in the week. If you pick the four Fridays in the month, you may be picking the busiest trading days. When selecting a sample size, you need to ensure that it is representative of the whole population.

Bias through sorting

The way in which you sort lists can introduce bias, although frequently this is unavoidable. A classroom teacher often consults a class list that is sorted alphabetically; for example, to select students for special tasks. The list is biased towards students whose surnames appear early in the alphabet and thus at the top of the class list. If you need to hire an electrician and consult a paper-based telephone book or an online directory, it is more likely that you will pick an early entry than one from the second page of listings. Bias of this type is difficult to avoid, so it is preferable to educate the user to recognise that the output has built-in bias and to encourage strategies to overcome that bias.

Bias through graphic representations

Bias can occur through your choice of graphic type, scale used and size chosen. Graphic representations should be sized proportionally to avoid overstating or trivialising the importance of one of the **variables** involved. For example, in Figure 1.4, a teacher has created a graph to show the heights of her students. The graph does not really give a clear picture of how the heights vary. The bars look as if they are all of similar height, so it is difficult to see the differences between them. In contrast, the different heights in Figure 1.5 are more distinct. This has been achieved by decreasing the scale of the vertical axis. The variation in the student heights looks much greater, even though the data has not changed. This graph makes it easier to see the small differences.

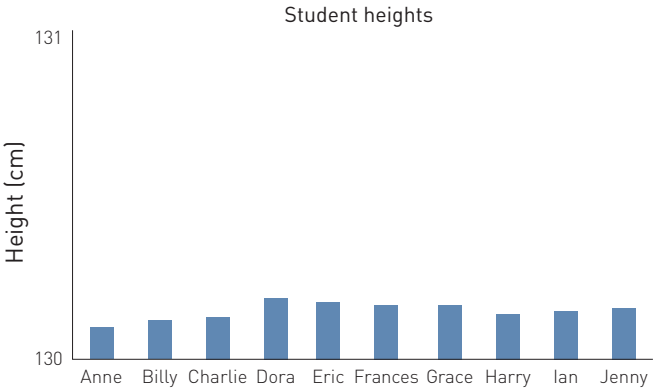


FIGURE 1.4 Column graph with little variation

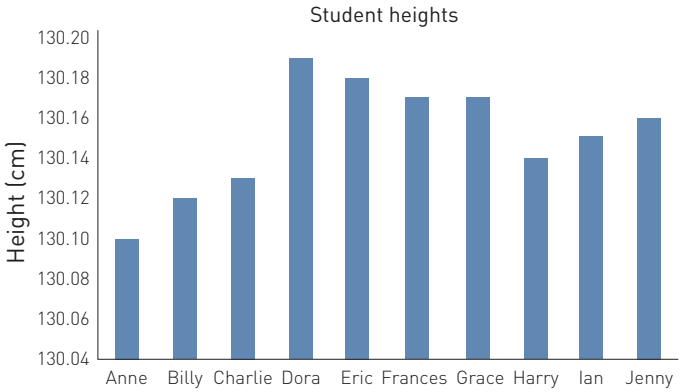


FIGURE 1.5 Column graph with greater variation

Reliability

The internet has made it easier for people to communicate their views and present information in a format that is easily accessible to others. There are many personal websites, homemade videos, wikis, podcasts, vodcasts and a plethora of unchecked information on the internet, and some of the views presented may not be widely accepted or proven to be accurate. Alternatively, there are many sources on the internet that are reliable such as the World Health Organization (WHO), Smithsonian, Australian Bureau of Statistics and universities.

THINK ABOUT COMPUTING 1.2
Find three reliable sources and describe in one paragraph the type of data they produce.

Referencing primary sources

Once primary data has been gathered, details need to be carefully recorded to enable appropriate **referencing**.

Interviews

For an interview, the following details need to be documented so that an interested person can go back to the source for checking, clarification and further information.

- Name of interviewee
- Date of interview
- Place of interview
- Qualification to be an interviewee – that is, whether the interviewee is a stakeholder in the issue and/or an expert about it

- Organisation to which the interviewee belongs (if relevant)
- Contact information for interviewee – phone number, address, email address, online chat handle
- How the interview was conducted; for example, in person, by phone, email or online chat
- Name and contact details of interviewer

Questionnaires

If you want to cite an individual response to a questionnaire or a survey, you need to record these details.

- Name of respondent
- When the questionnaire was completed
- Title of questionnaire
- Organisation to which the questionnaire belongs (if relevant)
- How the questionnaire was conducted – paper/online

Observation

For observation, the following details need to be recorded.

- The name of the person observed
- When the observation was conducted (date/time)
- Where the observation was conducted

Examples of referencing

Citations in a document help readers to find the source of the information and also assist students to avoid plagiarism. There are many ways to cite sources, such as providing footnotes, in-text citations or listing sources at the end of the document through a bibliography or references list.

Footnotes

Footnotes are listed at the bottom of the page on which a citation is made. Some academic disciplines prefer to use footnotes (notes at the foot of the page) to reference their writing. Although this method differs in style from the 'author, date' system, its purpose – to acknowledge the source of ideas, data or quotations without undue interruption to the flow of the writing – is the same. Footnotes are usually sequenced: series of numbers above the text (superscript) are placed in the appropriate part of the text to indicate the cited work and are matched at the bottom of the page after the footnote. A footnote lists the author, title and details of publication, in that order. Here is an example of a footnote:

> For example, in a 2009 article in the *Australian Financial Review*, journalist Jacqueline Maley wrote about the changing pattern of consumers' expenditure on leisure goods and services following the GFEC.[1]
>
> [1] Jacqueline Maley, 'Tough times bring home life's simple pleasures', *Australian Financial Review*, 7 September 2009, p. 3.

FIGURE 1.6 An example of a footnote: a superscript number is inserted in the body of the text; however, the full reference is provided at the bottom of the same page

APA

The American Psychological Association (APA) created a style guide to assist with academic writing such as publications, essays and books. The APA style is widely used and is one of the most common reference styles that students are expected to use. Citations within the text and their corresponding source details in a references list at the end of the work are necessary elements of the APA style. These show the reader where ideas and research have come from. Typically, when referencing using APA, the author's surname and the date of the publication are featured in the text. If quoting directly from the source, then the page number is also included. For example:

In some states, regimes of control on missions and reserves (a core element of the 'protectionist' system that dominated most of the twentieth century) continued into the 1970s or even 1980s. However, from the 1960s the dominant policy approach towards Indigenous Australians began to change, recognising, to some extent, Indigenous peoples' rights to preserve their cultures (Parbury 1999). Key events in this period included amendments to the *Commonwealth Electoral Act* in 1962 to give Indigenous people the vote in federal elections and the 1967 constitutional referendum, which meant Indigenous people would be counted in the national census (Attwood & Markus 2007) and the Commonwealth could legislate for them. By 1973, the new policy approach of Indigenous self-determination was introduced by the Federal Government (Sanders 2002). Along with the introduction of laws for equal wages, land rights and antidiscrimination in the 1960s and 1970s, these changes have sought to ensure fairer treatment of Indigenous Australians and better respect for their human rights.

Attwood, B. & Markus, A. 2007, *The 1967 Referendum: Race, power and the Australian Constitution*, Aboriginal Studies Press, Canberra.

Parbury, N. 1999, 'Aboriginal education: A history' in R. Craven (ed.), *Teaching Aboriginal Studies*, Allen & Unwin, Sydney, pp. 63–86.

FIGURE 1.7 An example of APA-style citation: the author's name and date of publication are inserted in the body of the text; the full reference is provided in the references list at the end of the document

Within the main body of the document, any reference to the publications will need to be cited. As illustrated in Figure 1.7, when using the APA style, the format usually follows author-date-page where the author's last name, the year of the publication, and the page number of the quote are referenced. These are all separated by commas, and are placed within parentheses following the text. The page number is preceded by a lower case 'p' with a period (full stop) after it; for example, (Parbury, 1999, p. 65).

EndNote is a commercial reference management software, used to manage bibliographies and references. It is similar to a database and is used to keep all references in one place. EndNote also integrates with MS Word, so that references can be easily inserted into the text with minimal effort, as shown in Figure 1.8. Alternatively, MS Word has its own built-in referencing capabilities that will allow you to create a bibliography and manage your sources without the need of additional software.

For more detailed information, visit the APA website.

For more information, visit the EndNote website.

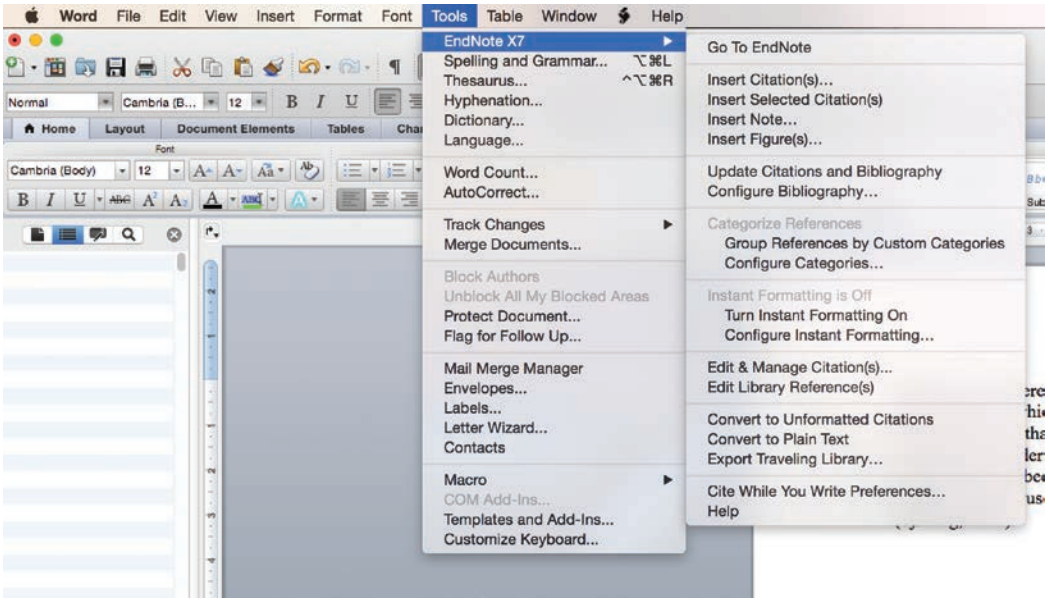
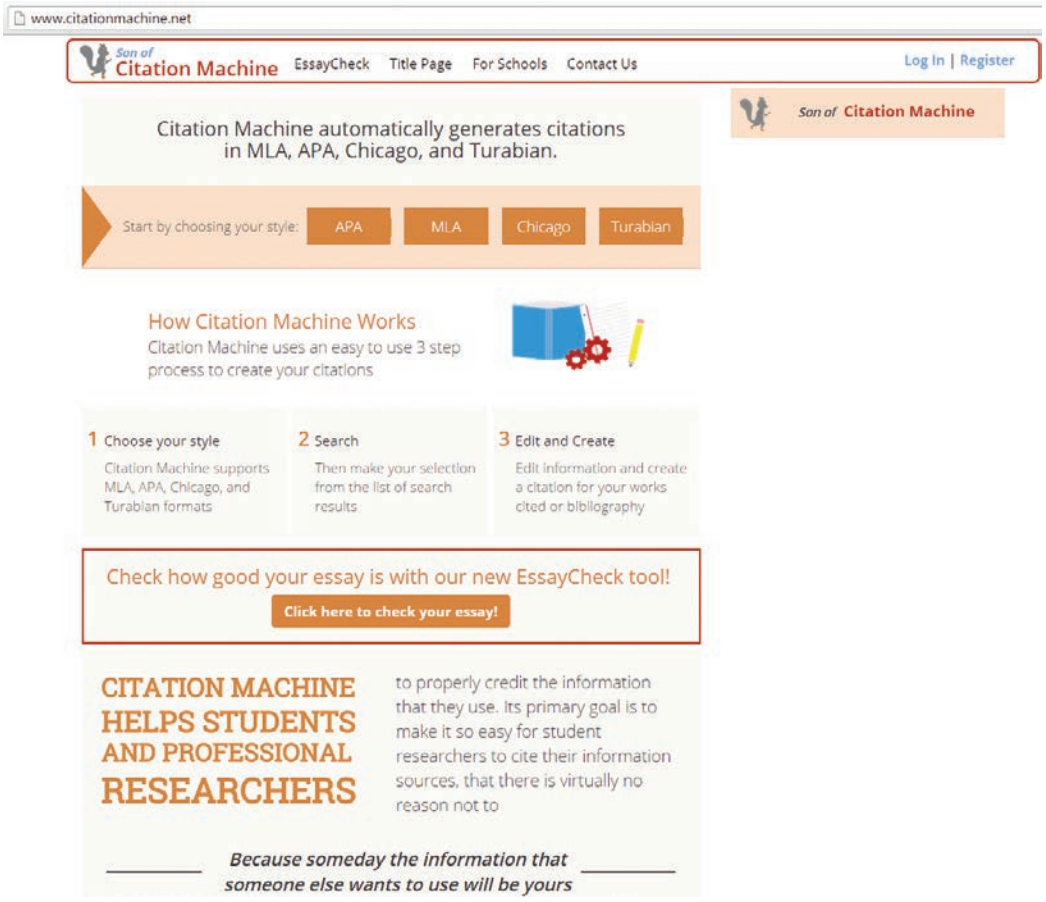


FIGURE 1.8 EndNote integrates easily into Microsoft Word



http://www.citationmachine.net

The free Citation Machine website, shown in Figure 1.9, generates citations in APA style. It can assist students to generate references in the appropriate format.

FIGURE 1.9 The Citation Machine can be used by students to assist with referencing

Seeking permission

Permission must be sought for collecting any data or information that involves people, because of privacy laws such as the Australian Privacy Principles (APPs). For example, to photograph or film individuals or groups, you should obtain permission. The organisation or individual who wants permission needs to let the people photographed or filmed know the purpose of the photographs or film and what it may be used for. Permission is usually provided in written form and is often known as consent. Permission needs to be sought because a photograph or video image in which an individual can be identified is considered to be personal information. Pictures of people can be used in advertisements, or for marketing purposes, and sometimes pictures can be used thoughtlessly and depict people in a false light.

Participants in any research need to be informed about what the research entails. They need to know what they are required to do and how much of their time it will take, and how often they will be required. All details of the research need to be given so that participants can make an informed decision to participate. For example, when researchers want to conduct a questionnaire, they need to specify how much time it will take (for example, 20 minutes) and how many times the questionnaire will need to be completed. Participants need to know whether the questionnaire will be paper-based or electronic, and how they will get access to it. All these details need to be explicitly stated so that each participant has a clear understanding before agreeing to take part. Participation needs to be voluntary. The researchers cannot put pressure on the participants or use coercion, or provide financial or other incentives for them to participate. Participation in research needs to be voluntary and informed.

In universities, research that involves people or animals cannot begin until researchers obtain ethics clearance. Each university has an ethics committee established and they follow the guidelines set out by the National Health and Medical Research Council (NHMRC). Before obtaining ethics clearance, the researcher/s must demonstrate that they have followed correct procedures and processes before collecting data to ensure that all risks have been addressed, and that it is established that participation is informed and voluntary.

Consent forms

One method of obtaining permission for research purposes is to use consent forms. **Informed consent** by all research participants is necessary. This means that before agreeing to participate in research, they are aware of what the research involves, the time required from them and the possible risks that may arise. Participation in research not only has to be informed, but also voluntary; that is, participants are not pressured to be involved and have the capacity to make their own decisions based on their understanding of the research.

You must obtain consent when interviewing or observing participants, creating questionnaires or surveys or collecting any type of data.

Consent forms should have the following information clearly listed.

- The title of the project
- The name of the researcher
- What the project is about and why it is being undertaken
- What is required from the participants in terms of time, effort, resources and costs
- The rights and interests of the participants – that they freely consent to be involved in the research and can withdraw at any time without having to provide a reason
- A statement of whether the participant's identity will be preserved

- Statements of confirmation, such as:
  - 'I have been informed of and understand the purposes of the study.'
  - 'I have been given an opportunity to ask questions.'
  - 'I understand I can withdraw at any time without prejudice.'
  - 'Any information that might potentially identify me will not be used in published material.'
  - 'I agree to participate in the study as outlined to me.'
- Name of participant, signature and date

FIGURE 1.10 A sample consent form

Name of school:

Project title:

Investigator(s):

1 I consent to participate in the project named above. I have been provided a copy of the project consent information statement to which this consent form relates and any questions I have asked have been answered to my satisfaction.

2 In relation to this project, please circle your response to the following:

• I agree to be interviewed by the researcher	**Yes**	**No**
• I agree to allow the interview to be recorded by electronic device	**Yes**	**No**
• I agree to make myself available for further information if required	**Yes**	**No**
• I agree to complete questionnaires asking me about <Insert topic>	**Yes**	**No**

3 I acknowledge that:

a my participation is voluntary and that I am free to withdraw from the project at any time without explanation;

b the project is for the purpose of research and not for profit;

c any identifiable information about me which is gathered in the course of and as the result of my participating in this project will be **(i)** collected and retained for the purpose of this project and **(ii)** accessed and analysed by the researcher(s) for the purpose of conducting this project;

d my anonymity is preserved and I will not be identified in publications or otherwise without my express written consent.

By signing this document I agree to participate in this project.

Participant name: .....

Signature and date: .....

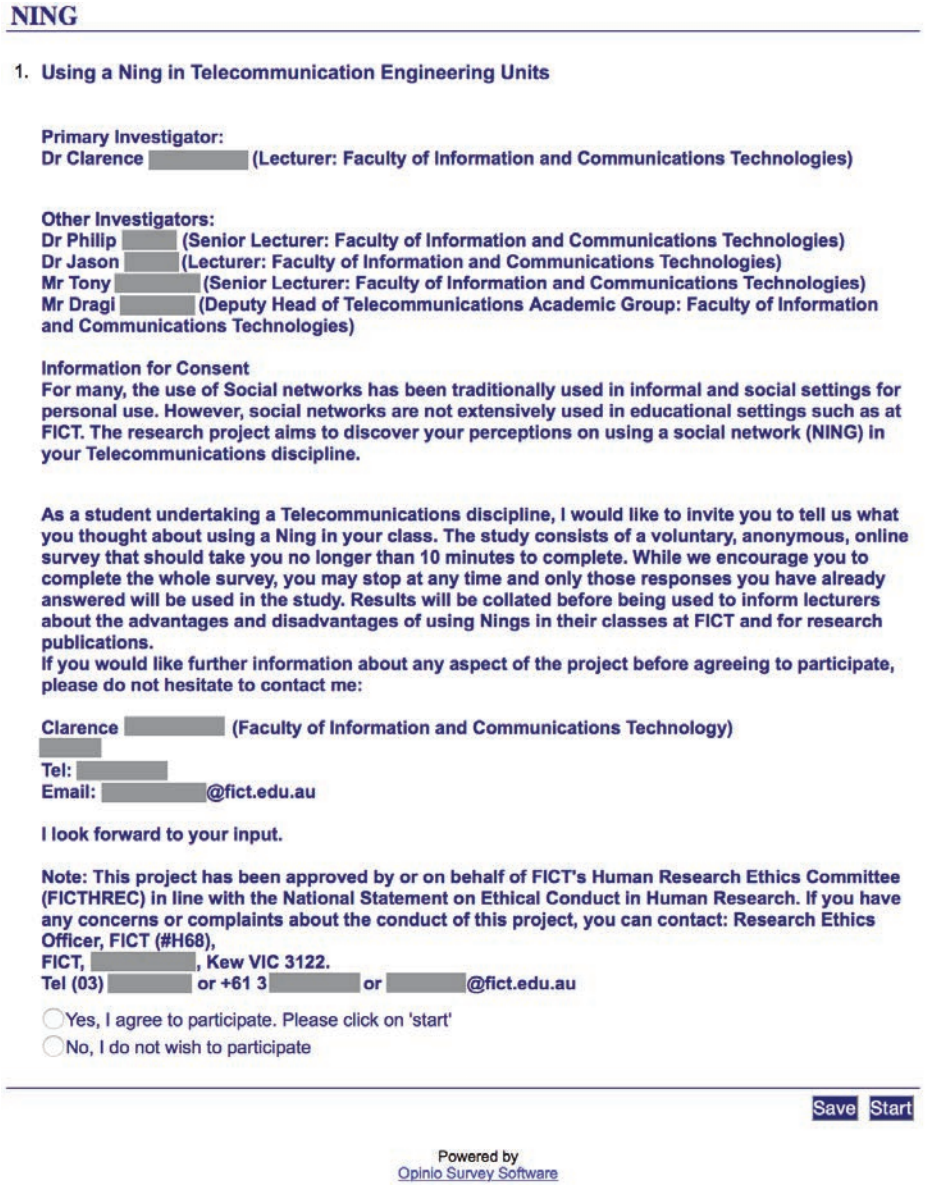
Participation information statements

The most common way to obtain informed consent is to provide a **participant information statement** written in plain English along with a consent form. This provides participants with information about the research and also provides the scope to answer questions that participants may have. Once they have read and understood the information statement, they are asked to sign the consent form to indicate their agreement to participate. When developing a participation information statement, consider the following.

- Use language your audience can easily understand.
- If technical terms are required, make sure they are clearly explained the first time they are used.
- The statement should be suitable for a 12-year-old reading age.
- Write in a conversational style, as if you were speaking to the participant.
- Language used should be clear, concise, invitational, culturally appropriate and logically set out.

- Use pronouns, such as 'I', 'we' and 'you'. This encourages the use of active voice and will be clearer to the reader. For example: 'You will be asked to participate'.
- Use reader-friendly formatting so that your document is easy to read.
- Ask others to read and edit your document.
- Include a statement on how the data will be stored.
- State where and how the research findings will be published.
- Give further information about the project, such as whom to contact.

Even if the data collection technique is an online survey and the researchers do not meet the participants, respondents must still be informed of the research, and asked to consent online prior to completing the questionnaire. They may be asked to consent by accepting and selecting the 'I agree to participate'. If they choose the 'I do not wish to participate' option, then they do not get access to the questionnaire. See Figure 1.11 for an example of an integrated online information statement and consent form.



NING

1. **Using a Ning in Telecommunication Engineering Units**

Primary Investigator:
Dr Clarence [redacted] (Lecturer: Faculty of Information and Communications Technologies)

Other Investigators:
Dr Philip [redacted] (Senior Lecturer: Faculty of Information and Communications Technologies)
Dr Jason [redacted] (Lecturer: Faculty of Information and Communications Technologies)
Mr Tony [redacted] (Senior Lecturer: Faculty of Information and Communications Technologies)
Mr Dragi [redacted] (Deputy Head of Telecommunications Academic Group: Faculty of Information and Communications Technologies)

Information for Consent
For many, the use of Social networks has been traditionally used in informal and social settings for personal use. However, social networks are not extensively used in educational settings such as at FICT. The research project aims to discover your perceptions on using a social network (NING) in your Telecommunications discipline.

As a student undertaking a Telecommunications discipline, I would like to invite you to tell us what you thought about using a Ning in your class. The study consists of a voluntary, anonymous, online survey that should take you no longer than 10 minutes to complete. While we encourage you to complete the whole survey, you may stop at any time and only those responses you have already answered will be used in the study. Results will be collated before being used to inform lecturers about the advantages and disadvantages of using Nings in their classes at FICT and for research publications.
If you would like further information about any aspect of the project before agreeing to participate, please do not hesitate to contact me:

Clarence [redacted] (Faculty of Information and Communications Technology)
[redacted]
Tel: [redacted]
Email: [redacted]@fict.edu.au

I look forward to your input.

Note: This project has been approved by or on behalf of FICT's Human Research Ethics Committee (FICTHREC) in line with the National Statement on Ethical Conduct in Human Research. If you have any concerns or complaints about the conduct of this project, you can contact: Research Ethics Officer, FICT (#H68),
FICT, [redacted], Kew VIC 3122.
Tel (03) [redacted] or +61 3 [redacted] or [redacted]@fict.edu.au

Yes, I agree to participate. Please click on 'start'
No, I do not wish to participate

Save Start

Powered by
Opinio Survey Software

FIGURE 1.11 Example of an integrated online information statement and consent form. Note that personal information has been redacted for privacy reasons.

The consent information statement should be clearly labelled as such (or quite similar) and contain key pieces of information as follows:

Project title
Give the project a full title

Investigators and other project personnel
List clearly all researchers.

Introduction to project and invitation to participate
Both introduce the project and invite participation.

What this project is about and why it is being undertaken
Sufficiently explain what the project is about, its aims, why it is being conducted.

Project and researcher interests
Researcher and project interests should be sufficiently disclosed as applicable; for example, project is partly, mainly or wholly to satisfy the requirements for a student's VCE subject.

What participation will involve – time, effort, resources, costs
Give sufficient clear detail as to what is being asked of participants – voluntary consent to their time, effort, supply of information/body tissue/records/personal effects, etc. Avoid language that can be read as orders or directives (i.e., not 'You will do this or that'; but better as 'We will ask you', etc.) and presumptuous language (e.g., 'Dear Participant').

Participant rights and interests – risks and benefits/contingencies/back-up support
Outline realistically any potential risks (minimal or otherwise) and what preventative, minimisation or redress arrangements are in place. If some research questions or issues can be considered particularly sensitive, give sample questions or topics as an indication of the information that will be discussed or requested. Describe any benefits pertaining to individual participation or more generally. Avoid grandiose claims.

Participant rights and interests – free consent/withdrawal from participation
Participation should be voluntary, free from any coercion or perceived coercion. Detail on this matter should be clear – that an individual is free to participate or not and the circumstances. If, for example, the participants are students, patients or employees, it will help to clarify that their decision to whether or not to participate will have no bearing on their results, treatment or employment (in some cases this may need further explanation, such as details about recruitment).

A statement about the participant's right to withdraw participation, data or material contributed, ordinarily without question or explanation, needs to be included.

Don't forget to outline how valid consent is to be obtained (by signed consent form, completion and return of an anonymous questionnaire, any witnessing procedure if applicable, etc.). Often it helps to highlight or bold this part.

Participant rights and interests – privacy and confidentiality
Give clear information about secure arrangements for data access, collection, use, retention and/or disposal. This needs to comply with mandatory Australian Privacy Principles.

If signed consent forms are required, state whether they will be stored separately from any data collected and who will have access to them.

Remember, people are increasingly concerned about data access and data matching. Clear information will help allay any concerns.

Research output
Outline intended or anticipated publication or reporting of research findings. If need be, reiterate or refer to privacy arrangements for confidentiality/anonymity. Offer to make available any report or article or summary, where appropriate, and indicate how this will occur.

Further information about the project – whom to contact
You will need to nominate at least one person to contact regarding further information about the research activity or participation in the project.

FIGURE 1.12 Information that should be included in a participation information statement

Privacy

When undertaking research, it is very important to uphold the privacy of the participants. Privacy is a fine balance between the interests of researchers and participants. Privacy laws attempt to stop inappropriate intrusion into the lives of individuals. However, the collection of data is often not the problem, but how the data is used or misused by people entrusted with it. To maintain privacy, and to de-identify the data, personal identifiers such as names and birthdates that are associated with individuals need to be removed so that information cannot be traced or identified.

TABLE 1.1 Data that has been collected but not de-identified

Given name	Surname	Height (cm)	Weight (kg)	Sex	Birthday	Postcode	Age	Active
Charlotte	Maine	145	48	F	22/10/2003	3103	11	No
Claudette	Shine	153	40	F	10/10/2003	3105	11	Yes
Abdul	Mensur	160	37	M	19/10/2002	3122	12	Yes
Eric	Anton	142	38	M	12/12/2003	3040	11	Yes
Jaspreet	Singh	148	41	M	31/12/2002	3041	12	No
Sienna	Megane	152	42	F	2/12/2003	3103	11	Yes
Elijah	Wu	144	43	M	6/01/2003	3122	11	No
Josephine	Wu	149	47	F	6/01/2003	3122	11	Yes
Hamish	Green	154	37	M	2/02/2003	3044	11	No
Jacinda	Black	143	46	F	4/04/2003	3111	11	No

TABLE 1.2 Data that has been de-identified: personal information such as birthdates, first names and surnames has been removed so that data cannot be identified or traced

Participant	Height (cm)	Weight (kg)	Sex	Postcode	Age	Active
ID1001	145	48	F	3103	11	No
ID1002	153	40	F	3105	11	Yes
ID1003	160	37	M	3122	12	Yes
ID1004	142	38	M	3040	11	Yes
ID1005	148	41	M	3041	12	No
ID1006	152	42	F	3103	11	Yes
ID1007	144	43	M	3122	11	No
ID1008	149	47	F	3122	11	Yes
ID1009	154	37	M	3044	11	No
ID1010	143	46	F	3111	11	No

Care must be taken to de-identify data. The example we have provided of de-identifying data only removes the names and dates of birth of survey participants. Depending on the remaining data and the information generated, this may be insufficient to protect the privacy of individual research participants from being identified (even accidentally) by at least some users of the information. This is particularly likely if the users know who the members of the sample group are, or if they are familiar with the wider population that is being sampled. Users of the information may be able to identify a survey participant even after his or her name and date of birth have been removed. This could be caused by a characteristic known to be rare in the sample group or wider population, because this helps to narrow possibilities. Combinations of rare characteristics would help to narrow them even further, so that users of the information could make an intelligent guess as to a person's identity.

Physical security controls

Researchers must ensure that data and materials generated and collected as part of their research, regardless of the format, are stored securely in a durable and accessible form. Data must be stored in a way that meets all legal and confidentiality requirements. Stored data can be protected with both physical- and software-based controls such as backing up data or shredding confidential documents.

Data needs to be securely retained and then securely disposed of or destroyed when no longer needed. How long the data or material need to be retained depends on the type of research or data. The minimum standard is set out in Section 2 of the *Australian Code for the Responsible Conduct of Research 2007*. The minimum timeframe for data to be retained is five years. For some specific types of research, other timeframes may be required. For data such as identifiable health research data, the minimum period may need to be seven years. For identifiable health research data involving children, the data may need to be kept until the individuals involved turn 25, or five years after any research outcome based on the data is published or made available for release, whichever is the longer.

Encryption

Encryption is the process of translating data into a secret code that can only be read by authorised users. To read an encrypted file, you must have access to a secret key that you use to **decrypt** the data. **Unencrypted data** is also known as 'plaintext'. Data that has been collected for research purposes needs to be secured so that only authorised people have access. Encryption is one way to ensure that the data is secured.

Backups

Backups form an essential step in data management. Regular backups protect against a number of risks including human error, computer crashes and software faults. Critical data files or data that are used regularly should be backed up frequently.

It is not unusual for someone to accidentally delete an important file or edit a document and later realise that some important information was removed. One strategy to minimise loss of data by an authorised user is to maintain a backup system. Important files inadvertently lost can be retrieved from the backup media.

A **full backup** copies all of the files from a device to a storage medium. It can take considerable time and is usually performed once over a time period (such as a week, fortnight or month). A **differential backup** copies only those files that have been changed since the last full backup. Restoration of data would involve restoring files from the full backup and then from the differential backup. An **incremental backup** is similar to a differential backup, the difference being that it uses more than two backup media, while a differential backup uses only two media. An incremental backup only copies files that have been changed since the last incremental backup. It is the most complicated strategy from which to restore files since it requires restoration from a full backup and then from a series of incremental backups. It is a good practice to clearly label all backup media so that you know when the backup was made and what is on it.

Location of backup files

Once you have created backups, where do you put them? Ideally, your backups should be stored in a location that is safe from theft and damage caused by extremes of temperature or disasters. Most small businesses have a fireproof and waterproof safe in which valuable company

documents are stored. This might also be used to store backups. It is preferable, however, to store backups at a remote location, perhaps even in the cloud. This means that if there is a large natural disaster, such as a huge flood or an earthquake, the backups will be safe.

One last point to remember is to ensure that backups actually work when you want to restore the data. It is important to test the effectiveness of your backup files by running a disaster recovery simulation. If files cannot be restored from the backup or the system refuses to recognise them, it is better to discover this before a real emergency.

Cloud-computing companies provide offsite storage, processing and computer resources to individuals and organisations. These companies are typically third party and they store data to a remote database in real-time. The internet provides the connection between this database and the user's computer. One of the advantages of cloud storage include the ability to access data from any location that has internet access, eliminating the need to carry a USB, or a hard-drive to retrieve and store data. The ability to share files with other people and collaborate simultaneously, such as by using Google Docs, is also an advantage. Finally, if something were to happen to the computer, such as a fire or natural disaster, and the data on it were to be destroyed, having the data saved offsite in the cloud would prevent the data from being lost.

Google Docs allows users to upload documents, spreadsheets and presentations to Google's servers (Figure 1.13). Users can edit files using a Google application and work on them at the same time as others, so they can read or make edits simultaneously.

The word 'cloud' is used because the internet was originally, and still is, represented by a cloud in network diagrams. An example of cloud computing is Google's Gmail, which uses cloud-computing processing powers and storage facilities.

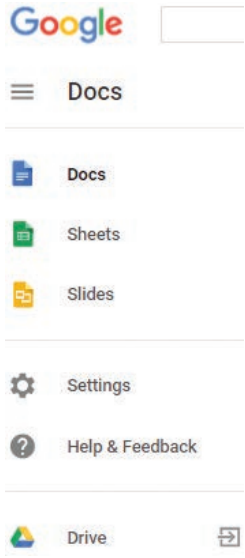


FIGURE 1.13 Google Docs

Usernames and passwords

A username usually is identifiable as belonging to a particular person and can be easily remembered. Usernames are uniquely assigned to users. Passwords are set by the user and should be known only to that user. To maintain high levels of security, user passwords should:

- be at least eight digits long
- include non-alphabetical characters
- not be easily guessed (for example, a favourite pet's name is not suitable)
- be changed every month.

Some network policies force passwords to be changed on a regular basis, and do not allow passwords to be repeated.

In terms of securing individual files, a password can be placed by an individual to prevent unauthorised people accessing the data. Password protection can be placed on files to prevent them from being viewed by unauthorised people, as shown in Figure 1.14.

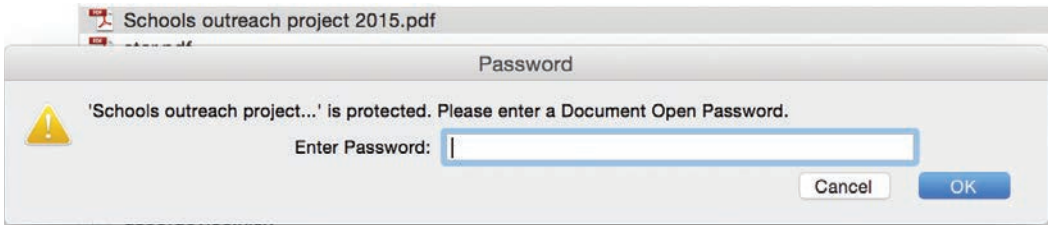


FIGURE 1.14 Password protection has been placed on the file, preventing access to it unless a password is entered

Firewall

A **firewall** is a server and software combination that filters the information coming through an internet connection into an organisation's internal network. Any packet of data that is flagged by the filters as unwanted is not allowed through.

The filters used by a firewall include examining the IP address of computers that request information from an internal server, blocking all access to certain domain names, banning certain protocols (e.g. file transfer protocol, mail protocol or Telnet protocol) from accessing particular servers, and searching for certain words and phrases included in packets of information, and blocking transfers that contain them.

A firewall can also be used to restrict employees' access to sensitive information. For example, a firewall can be used to stop some personnel from accessing the payroll database.

Most firewalls use two separate network interface controllers (NICs); one is connected to the internal network and the other to the outside world. Material can only move from one card to the other through the CPU of the server computer that is acting as the firewall. While the data or information is being checked for authenticity, it is also examined for viruses and other malicious codes. Everything that comes in from outside is examined for danger.

Firewalls are explained in more detail in Chapter 3.

Malware protection

Malware is malicious software that includes spyware, adware, Trojan horses, worms and viruses. Spyware and adware use cookies to track the internet sites that a user might visit. Trojan horses can leave your computer open to others to read your personal information by creating backdoor access to your system. Viruses and worms can hijack your system to send multiple emails to others or perform other acts of mischief. Both can use up essential system resources, which may result in the computer freezing.

Network administrators usually require workstations to run virus protection software. The antivirus software is often updated automatically via the network. A firewall is also useful to block malware from sending personal information over the internet. Anti-adware programs should also be run on workstations.

Australian Privacy Principles

Originally, the *Privacy Act 1988* only dealt with the handling of data by Federal government agencies. Many people criticised this limitation because it seemed that private organisations were not required to apply even the most basic of safeguards on data they collected. Even worse, there were no regulations preventing non-government organisations from collecting data by any method and using it for any purpose without the consent of the people whose private details were concerned. In particular, the rapid growth of electronic transactions, especially over the internet, led many people to demand some sort of legal protection from those who might gather data about internet browsing habits. The government was keen to encourage the development of electronic commerce, while protecting the confidentiality of consumers and increasing public confidence in electronic transactions. These amendments have now been incorporated into the *Privacy Act 1988* (Cth) and are the most significant changes to privacy laws since the inception of the legislation.

There have been several additional powers included within this Act since 1988, but its main purpose has remained unchanged. The *Privacy Act 1988* (Cth) was amended by the Privacy Amendment (Enhancing Privacy Protection) Bill in 2012. This came into effect in 2014. As part of this Act, the Australian Privacy Principles replaced the National Privacy Principles and the Information Privacy Principles so that Australia now has one set of privacy principles. The Australian Privacy Principles (APPs) generally apply to Australian Government agencies. They do not apply to Local Councils or State or Territory Governments. Some States have their own privacy laws. For example Victoria has The *Privacy and Data Protection Act 2014* (PDPA).

The changes to the *Privacy Act* include 13 new privacy policy principles known as the Australian Privacy Principles (APPs). The APPs were devised to set out the standards, rights and obligations for collecting, handling, holding, accessing, using, disclosing and correcting personal information. The APPs oversee the handling of personal information by:

- Australian and Norfolk Island Government agencies
- all private health service providers
- businesses that have an annual turnover of $3 million or those that trade personal information.

Of particular interest are Principle 2, *Anonymity and pseudonymity*, and Principle 6, *Use or disclosure of personal information*.

TABLE 1.3 The Australian Privacy Principles

APP 1 Open and transparent management of personal information
Ensures that APP entities manage personal information in an open and transparent way. This includes having a clearly expressed and up-to-date APP privacy policy.

APP 2 Anonymity and pseudonymity
Requires APP entities to give individuals the option of not identifying themselves, or of using a pseudonym. Limited exceptions apply.

APP 3 Collection of solicited personal information
Outlines when an APP entity can collect personal information that is solicited. It applies higher standards to the collection of 'sensitive' information.

APP 4 Dealing with unsolicited personal information
Outlines how APP entities must deal with unsolicited personal information.

APP 5 Notification of the collection of personal information
Outlines when and in what circumstances an APP entity that collects personal information must notify an individual of certain matters.

APP 6 Use or disclosure of personal information
Outlines the circumstances in which an APP entity may use or disclose personal information that it holds.

APP 7 Direct marketing
An organisation may only use or disclose personal information for direct marketing purposes if certain conditions are met.

APP 8 Cross-border disclosure of personal information
Outlines the steps an APP entity must take to protect personal information before it is disclosed overseas.

APP 9 Adoption, use or disclosure of government related identifiers
Outlines the limited circumstances when an organisation may adopt a government related identifier of an individual as its own identifier, or use or disclose a government related identifier of an individual.

APP 10 Quality of personal information
An APP entity must take reasonable steps to ensure the personal information it collects is accurate, up to date and complete. An entity must also take reasonable steps to ensure the personal information it uses or discloses is accurate, up to date, complete and relevant, having regard to the purpose of the use or disclosure.

APP 11 Security of personal information
An APP entity must take reasonable steps to protect personal information it holds from misuse, interference and loss, and from unauthorised access, modification or disclosure. An entity has obligations to destroy or de-identify personal information in certain circumstances.

APP 12 Access to personal information
Outlines an APP entity's obligations when an individual requests to be given access to personal information held about them by the entity. This includes a requirement to provide access unless a specific exception applies.

APP 13 Correction of personal information
Outlines an APP entity's obligations in relation to correcting the personal information it holds about individuals.

APP 2 Anonymity and pseudonymity

Australian Privacy Principle 2 provides individuals dealing with organisations the option of using a different name or a **pseudonym** in relation to a particular matter. This measure is in place so that individuals cannot be identified. In addition, individuals can also remain anonymous. For example, when an individual calls an organisation, often a message states that the call will be recorded for training purposes. If the individual objects, the call is not recorded. At a later date, when staff from the organisation receive training, the names of the individuals whose voices have been recorded must be changed in order to protect their identity when these real examples are used.

APP 6 Use and disclosure of personal information

Australian Privacy Principle 6 states that the information that is being held is in line with the primary purpose it was intended for. Information cannot be used for a secondary purpose unless the holders of the information have received consent from the individuals concerned. For example, if a sporting organisation collected information about their players for the purpose of organising competitions and making the details available to the coaches and captains, then that would be the primary purpose of the data. However, if the sporting organisation wanted to provide the details of their members to a shop that specialised in sporting merchandise, then it could not do so, as this would be using the information for a purpose for which it was not intended.

Ethical dilemmas

Everyone wants the benefits of digital systems; however, intended and unintended negative effects can impinge upon people's rights. As a result, those who design, control and use digital systems have a responsibility to consider the real and potential negative effects and to eliminate or lessen them as much as possible. Sometimes even this may not be enough to justify the proposed collection or creation of data. It is important to take into account legal objections and ethical considerations when creating or acquiring data. The purpose for collection needs to be clear. This also needs to be articulated in the participant information statements and the consent forms provided to the people from whom information will be sought.

Ethics refers to behaving in ways that are based on our morals and accepted standards. These standards may be common in a particular society or specific to a single organisation. They apply to questionable activities over and above any legal requirements. Ethics often provide us with a set of guidelines of appropriate behaviour. If we choose to ignore these guidelines, we may not be committing a crime, but we may be sacked by an employer or shunned by society. For example, the impact of violent video games on children has long been debated. Some people have voiced their concerns that video game writers should not include animated violence in their games because it has a negative impact on children. A system of classification exists for games, similar to television and film classifications, but there is no legal restriction preventing these games from being created. These examples demonstrate how ethics hinge on society's values and standards. In this example, there are two competing principles. On the one hand, some proponents would argue that protecting children from possibly harmful video games is the right thing to do. On the other hand, others would argue that it is more important to maintain freedom of expression. Often, then, questions of ethics become debates over which

of two principles is more important. Such conflict may be said to constitute a **dilemma**. This is especially the case when the consequences of action are open to debate or interpretation.

The standards or guidelines that determine whether an action is good or bad are known as ethics. Ethics are the moral guidelines that govern, among other things, the use of data collection. Often ethical principles/guidelines have an accompanying law, but the ethical principle is usually broader and the law applies only to certain circumstances or applications of the principle.

For example, it is ethical to obtain permission to publish photos of people on websites or in promotional material. Sometimes people may object to their images being used for these purposes. The purpose for taking the photo and how it is intended to be used need to be made clear. Ethically it is wrong to use a photo for a different purpose from that for which it was originally collected. Similarly, when using data-collection tools such as surveys, interviews and questionnaires, it is important to reassure participants that the data provided, within the limits of the law, will remain anonymous and that their individual comments will not be able to be identified by others. It is not simply that it is important to put participants' minds at rest regarding their concerns about protecting privacy; it is also important to ensure that their privacy is in fact protected (and also to ensure that non-participants in the larger group of which the sample is supposed to be representative, are not put at unacceptable risk of suffering as a result of mistaken identification).

This may be of particular concern when deciding what must be removed to de-identify data sufficiently to protect all its potential users. It is also important when reporting personal information anonymously or using pseudonyms in a newspaper report, for example.

Ethical frameworks assist people to work through ethical dilemmas. A six-step framework as described below can be used for handling ethical dilemmas to solve a legal, ethical or social tension. This framework will provide support and guidance in making a decision when presented with an ethical dilemma.

1 Identify the problem: What decision has to be made and what facts are required?
2 Identify the stakeholders: Who are they? What interests do they have? What power do they have? Who is vulnerable? How are the vulnerable to be protected?
3 Identify possible alternatives: What options are available? What are the likely consequences?
4 Identify ethical standards: Are there any applicable laws? Are there any morals or standards that could be applied? Is there a precedent?
5 Evaluate options: Identify strengths and weaknesses. Identify the option that causes least harm. Can the decision be reversed?
6 Make a decision: Select the preferred option. Justify the option.

CHAPTER SUMMARY

ESSENTIAL TERMS

backup copying files from an information system to some type of storage device to guard against possible data loss

bias a prejudicial or unreasoned judgement

closed (or closed-ended) questions queries that are restricted in the range of options provided so that only specific answers are elicited

data unprocessed, unorganised and distinct facts or ideas; in addition to text and numbers, data also includes sounds, images and video

decrypt to translate encrypted data back into ordinary text that can be read by anyone

differential backup used in conjunction with a full backup, only files that have been altered since the last full backup are copied; restoration requires the full backup to be restored first, followed by files from the differential backup

dilemma when people must choose between two (or more) equally desirable (or undesirable) options; for example, between allowing the sale of violent video games to preserve freedom of expression and banning their sale in order to protect children from possible harm

encryption the process of translating data into a secret code that can only be read by authorised users; to read an encrypted file, you must have access to a secret key that you use to decrypt the data; encrypted data is also known as 'ciphertext'

ethics guidelines based on our morals and accepted standards; these standards may be common in a particular society or specific to a single organisation, and apply to questionable activities over and above any legal requirements – ethics often provide us with a set of guidelines of appropriate behaviour

firewall hardware and software that restrict access to data and information on a network

focus group the meeting of a small group of individuals who are guided through a discussion by a researcher; focus group members are carefully selected, so that each one fits a particular demographic and the researcher can obtain the necessary data through a guided discussion that probes the participant's attitudes about the topic

footnotes a type of reference that is listed at the bottom of the page on which the citation is made

full backup copying all chosen files to a backup device; it can be slow to perform, but is the easiest and quickest way to restore data form

hypothesis a statement based on probabilities proposing what will happen according to the applied theory; some hypotheses are tested using data collection tools such as questionnaires and/or interviews, and the results of the study will either support or disprove the hypothesis

incremental backup similar to a differential backup in that it works in conjunction with a full backup, but only backs up files that have been altered since the last incremental backup; it is the most complicated strategy from which to restore files

information processed, organised and value-added data, which can be paper-based (hard copy) or digital (soft copy)

informed consent a necessity for all participants before agreeing to take part in research, participants must be informed of what the research involves, the time commitment expected and the possible risks that may arise

malware programs designed to infiltrate and cause harm to a computer or network without the owner's knowledge or consent, including viruses, worms, Trojans, adware, spyware, logic bombs and keyloggers

observation a way of finding out about the world around us; using our five senses, we pick up detailed information about our environment

open-ended questions queries that allow people to answer the question in the manner they wish; they are called open ended because the person responding is free to answer in any manner he or she chooses

participant information statement a document that provides participants with information about the research in an unbiased way, and also provides the scope to answer questions that the participant may have

primary sources original, uninterpreted data and resources; that is, information that has not been processed, analysed or interpreted in any way, such as interviews, speeches, emails, debates and meetings; primary data usually come from stakeholders

pseudonym a fictitious name that is given to a person, or that is chosen by a person, to hide or protect their identity

qualitative data collected data that is harder to measure because it is based on subjective data collection techniques such as interviews, focus groups, video footage and observation

quantitative data collected data that is measurable and specific; quantitative data gathering is based on verifying theory through the use of statistics and largely numerical data

questionnaire or **survey** usually a set of questions that ask for a response to be selected from a list of alternatives, such as A, B, C, D; or from a range, 1–5 or very low to very high; questionnaires can easily be given to many people, and are easily processed and analysed using computer-based methods because the answers can be recorded as numbers

referencing citations in a document that assist readers to know where to locate the source of an original idea or quote in a piece of work, and assist students to avoid plagiarism

secondary sources information that has been processed, interpreted or analysed in some way by others, such as textbooks, websites, magazines, newspapers and TV programs

stakeholder an individual or group who either has an interest in or is affected by the decisions and actions of an organisation

survey see **questionnaire**

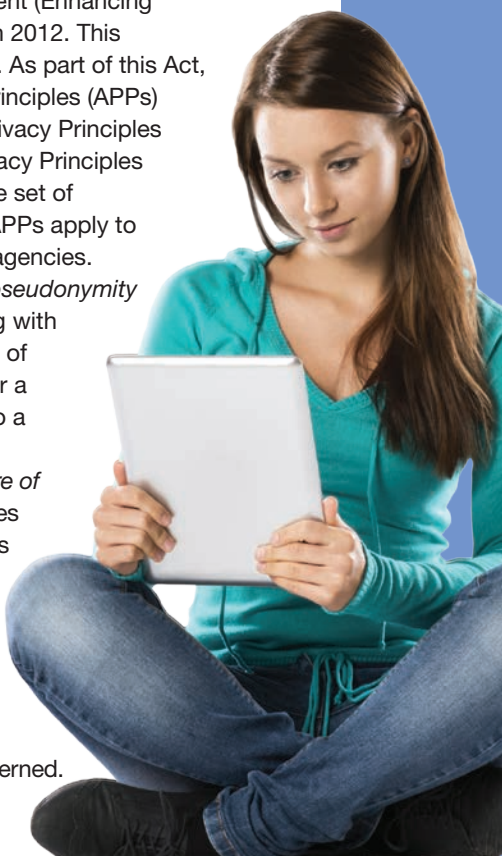
unencrypted data also known as 'plaintext'

variable in programming, a key word, phrase or symbol that represents a value that may change

vested interest arises when an individual, group or organisation has a strong personal interest because there is an advantage to be gained

IMPORTANT FACTS

1 Data must be relevant to produce usable information. Data needs to be processed while it is current because decision-making should not be based on outdated data.
2 Data that is entered into a computer must be accurate. Transcription is often a cause of error. Transcription errors occur when the person entering the data misreads the information through, for example, a lapse in concentration, being interrupted or pressing the wrong key.
3 Interviews are usually done one-to-one, but can sometimes be done in groups, and can take a substantial amount of time. A major feature of an interview is the opportunity for in-depth follow-up and clarification questions that cannot be done with questionnaires, which are often answered in private. Interviews are very useful for eliciting the feelings, attitudes and opinions of people that are too complex to easily record in a questionnaire.
4 Bias can infiltrate data if the respondent to a survey or interview has a vested interest in the outcome of the research, if the timing of the data gathering is inappropriate, or if the sample size chosen is too small.
5 Timing of events needs consideration when collecting data as it can cause skewed results, which can lead to inaccurate or misleading conclusions.
6 Sample size must relate to the purpose of the data collection and, generally, a larger sample size leads to greater precision.
7 There is a plethora of unchecked information on the internet; however, some of the views presented may not be widely accepted or proven. Sources cited should be reliable.
8 The American Psychological Association (APA) created a style guide to assist with academic writing such as essays, books and other publications. The APA style is widely used and is one of the most common styles that students are expected to reference.
9 Privacy is a fine balance between the interest of researchers and the participants. Privacy laws attempt to stop inappropriate intrusion into the lives of individuals. However, the collection of data is often not the problem, but how the data is used or misused by people entrusted with it.
10 Researchers must ensure that data and materials generated and collected as part of their research, regardless of the format, are stored securely in a durable and accessible form. Stored data can be protected with both physical and software-based controls such as backing up and shredding confidential documents.
11 Cloud-computing companies provide offsite storage, processing and computer resources to individuals and organisations. These companies are typically third party and they store data to a remote database in real-time.
12 To verify users' rights to access a network, security features are required. A system of establishing usernames (or user IDs) and passwords allows for the identification and authentication of each user.
13 The *Privacy Act 1988* (Cth) was amended by the Privacy Amendment (Enhancing Privacy Protection) Bill in 2012. This came into effect in 2014. As part of this Act, the Australian Privacy Principles (APPs) replaced the National Privacy Principles and the Information Privacy Principles so Australia now has one set of privacy principles. The APPs apply to Australian Government agencies.
14 APP 2, *Anonymity and pseudonymity* offers individuals dealing with organisations the option of using a different name or a pseudonym in relation to a particular matter.
15 APP 6, *Use and disclosure of personal information* states that the information that is being held is in line with the primary purpose it was intended for. Information cannot be used for a secondary purpose without consent from the individuals concerned.



TEST YOUR KNOWLEDGE

Review quiz

DATA AND INFORMATION

1 What can occur if information is produced from incorrect or incomplete data?
2 Why is it important to ensure that data is accurate?
3 What are the properties of usable data?
4 What is a common cause of inaccurate data?
5 How can the accuracy of a primary source be determined?
6 With an example, explain the importance of timeliness in ensuring the quality and usability of data.
7 What influences the introduction of bias into data?
8 What is the difference between quantitative and qualitative data? Provide two examples.
9 What strategies could be used when gathering quantitative data?
10 Provide an example of referencing based on APA.
11 Why is it important to obtain permission when collecting data?
12 What is the purpose of consent forms?

DIGITAL SYSTEMS

13 Why is encryption important in data security?
14 How do usernames and passwords protect data?
15 What makes a strong password? Provide an example of a very strong password.
16 Describe a strategy on backing up data?
17 What is the difference between a full backup and an incremental backup?
18 Why is it important to secure data when conducting research?

INTERACTIONS AND IMPACT

19 How do the Australian Privacy Principles affect the individual?
20 Under which legislation does the Australian Privacy Principles fall?
21 Why is it important to de-identify personal data?
22 What is an ethical dilemma in the context of data collection for research purposes?

APPLY YOUR KNOWLEDGE

STREET TRAFFIC

1 Arthur believes that there is too much traffic in his street, particularly on weekdays – in the morning and evening. Arthur lives in a small residential street, but cars use this street as a shortcut to avoid traffic lights and the main roads. Arthur wants to do some research to support his theory about the traffic on his street, and present a case to the local council.

- **a** Clearly state the topic Arthur will investigate.
- **b** What type of data will Arthur need to assist in his investigation?
- **c** Identify an appropriate data gathering technique Arthur could use.
- **d** Justify the selected data gathering technique.
- **e** How will Arthur keep the data safe?
- **f** Does Arthur need to get permission to conduct his research and if so, from whom?
- **g** What tools will Arthur use to interpret the results?
- **h** What types of relationships and patterns is Arthur looking for?
- **i** How will Arthur present the data to the local council?

INTERNET USAGE

2 Go to the UN Data website and find the Information and Communications Technology database. Select to view the data on the percentage of individuals using the internet.

UN Data

- **a** Filter the data to show only the percentage of individuals using the internet.
- **b** Copy and paste this data into a spreadsheet program.
- **c** Create a column chart, a scatter diagram and a bubble chart that depict these statistics.
- **d** Discuss which graphic representation best conveys the data and why.