In this chapter you will learn to:

- for a given scenario, identify alternatives for data collection and choose the most appropriate one
- use a range of hardware collection devices to collect different data types
- describe the operation of a range of hardware collection devices
- make predictions about new and emerging trends in data collection based on past practices
- choose the most appropriate combination of hardware, software and/or non-computer tools to collect data from a given source
- use the Internet to locate data for a given scenario
- design forms that allow data to be accurately recorded and easily input into software applications
- select and use appropriate communication skills to conduct interviews and surveys so that data can be accurately collected
- identify existing data that can be collected for an information system for a given scenario
- recognise personal bias and explain its impact on data collection
- identify the privacy implications of particular situations and propose strategies to ensure they are respected
- predict errors that might flow from data inaccurately collected
- predict issues when collecting data that might arise when it is subsequently analysed and processed

Which will make you more able to:

- describe the nature of information processes and information technology
- classify the functions and operations of information processes and information technology
- identify and describe the information processes within an information system
- recognise and explain the interdependence between each of the information processes
- identify and describe social and ethical issues
- describe the historical developments of information systems and relate these to current and emerging technologies
- select and ethically use computer based and non-computer based resources and tools to process information
- analyse and describe an identified need
- generate ideas, consider alternatives and develop solutions for a defined need
- recognise, apply and explain management and communication techniques used in individual and team-based project work
- use and justify technology to support individuals and teams

In this chapter you will learn about:

Collecting – the process by which data is captured or entered into a computer system, including:

- deciding what data is required
- how it is sourced
- how it is encoded for entry into the system

Hardware used for collection

- scanners and/or digital cameras to collect images
- microphones and/or recording from peripheral devices to collect audio
- video cameras and/or peripheral devices with appropriate interfaces to capture video
- keyboards and/or optical character readers to collect numbers and text
- data capture devices such as counters for counting cars on a road
- historical and emerging trends in hardware collection devices

Software used for collection

- device drivers that allow hardware to interface with the operating system
- software that allows participants to enter or import data
- software that allows participants to move data between applications

Non-computer procedures in collecting

- literature searches
- surveys and interviews
- form design for data collection
- manual recording of events
- existing non-computer data

Social and ethical issues in collecting

- bias in the choice of what and where to collect data
- accuracy of the collected data
- copyright and acknowledgment of source data when collecting
- the rights to privacy of individuals on whom data is collected
- ergonomic issues for participants entering large volumes of data into an information system

3

TOOLS FOR INFORMATION PROCESSES: COLLECTING

Collecting is essentially an information input process; it gathers data from the environment for use by the system. In Chapter 2 we discussed aspects of the collecting process that need to be understood prior to the actual collection of data commencing; this includes deciding on what to collect, from where it will be collected as well as how it will be encoded during collection. In this chapter we focus on the tools available for use during the collection process. Each of these tools is suited to the collection of particular media types from particular sources using particular collection techniques. For example a microphone is used to collect audio data from sound waves. It does this by sensing the compression waves and converting them into changes in voltage; this data is then suitable for conversion into digital sound samples. Conversion from analog into digital, is strictly speaking, an organising process, however as it is integral to the operation of many collection tools it makes sense to describe this operation as we discuss each tool.

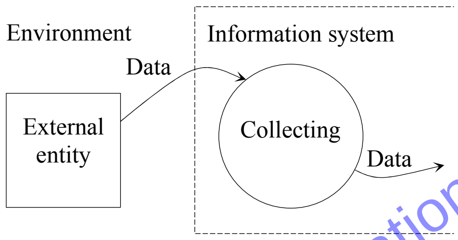


Fig 3.1
Collecting gathers data from the environment for use by the system.

Many of the tools considered in this, and subsequent chapters perform actions from more than one of the seven information processes; these tools, as is the case with the information processes themselves, do not operate in isolation. For example the CPU is clearly involved in all seven of the information processes yet we examine its operation in the processing chapter. Each tool has thus been categorised according to the particular information process that most closely aligns with the actions it performs.

In this chapter we first consider the operation of various hardware devices used to collect each of the different media types. In terms of collecting, hardware includes such common input devices as keyboards, scanners, microphones and video cameras together with various specialised data capture devices. We then examine the software used when collecting data; the software used to interface or communicate with input devices, software that forms the user interface for data entry and software that imports data so it can be moved between applications. We examine a number of non-computer procedures for collecting and finally discuss some social and ethical issues relevant to the collection process.

FORUM

GROUP TASK Activity
Brainstorm a list of input devices. Categorise this list according to the different types of media each device is designed to collect

HARDWARE USED FOR COLLECTION

It is not possible to examine all types of input device, so in this section we restrict our discussion to include at least one type of device for each of the five media types.

- Keyboard for collecting of text.
- Mouse for collecting various media types.
- Scanner for collecting images.
- Digital camera for collecting images.
- Microphone and sound card for collecting audio.
- Video camera for collecting video.
- Vehicle counting and monitoring for collecting various media types.

Be aware that often the data collected by these devices is organised into a different type after collection; from the users' perspective this is often not obvious. For example when using a spreadsheet, we think we are entering numeric data, in actuality the keyboard collects the data as text; it is the spreadsheet application that converts this text data into numbers and displays it on the screen. Similarly a barcode scanner is really collecting image data, organising it into text (usually digits) and sending it to the computer for further processing.

KEYBOARD

In essence a keyboard is a collection or matrix of switches; each switch completes a circuit to indicate a particular key, or combination of keys, has been pressed. A digital code representing the key is then sent as an electrical signal to the computer. Sounds relatively simple, however in reality the keyboard is an amazing mix of ergonomic and technological design.

Fig 3.2
Section of a QWERTY keyboard. Note the staggered rows and standard size and shape of each key.

To structure our discussion let us work through the operations occurring as a single character is collected. That is, from the time the user presses a key until the computer receives the information.

First the user decides which key to press and locates that key. This may seem obvious but there are many aspects of keyboard design that facilitate this process. Consider the standard design of the keys; in most cases a QWERTY layout is used, the layout of the keys needs to be familiar if the user is to efficiently locate the correct key. Consider the physical size and shape of each key and the way each row is staggered relative to other rows (see *Fig 3.2*); these attributes are common to almost all keyboards, they allow users to transfer their keyboard skills from one keyboard to another. At first glance most keys appear to be cubes; actually they are tapered, with the top surface of each pad slightly concave; these design elements assist the fingers to positively locate the required keypad without touching adjoining keys.

So the user now presses the key; during this process the keyboard provides feedback to the user. Finger pressure moves the key down then upon release the key springs back to its original position, at the same time an audible 'click' is often produced. This feedback is a major factor in determining the general feel of a keyboard and is perhaps the most significant reason conventional keyboards are considered superior to most notebook keyboards; notebooks minimise the downward throw of each key to reduce their thickness.

Contained under each keypad is a switch which completes a circuit indicating precisely which key has been pressed. There are various designs of key switch used for this process; older designs use a matrix of mechanical switches, each switch being similar to those used for other applications such as door bells. At the time of writing, the most common designs utilise flexible rubber or silicone domes. Some use separate domes containing a carbon button for each key (see *Fig 3.3*). When a key is pressed the dome flexes, causing the carbon button to complete a circuit on the underlying circuit board, when released the dome springs back to its original shape. Other designs utilise two printed circuit boards separated by a thin film containing a hole for each key (see *Fig 3.4*); the domes are contained within a single silicone membrane. When a dome is depressed the contacts touch through the hole in the film to complete the circuit. All these designs are far simpler, and cheaper to produce, than traditional switches; furthermore the domes protect the actual contacts from dust and liquid spills.

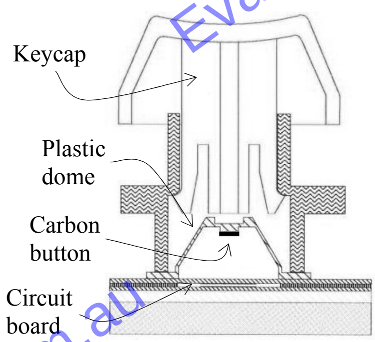


Fig 3.3
Detail of a typical flexible dome key switch similar to those used on many keyboards.

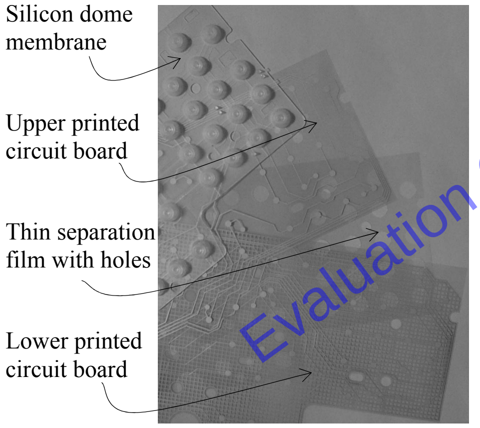


Fig 3.4
Detail of a keyboard design utilising a silicon membrane of domes and two circuit boards separated by a thin film.

The circuit board is really a matrix of wires; the intersection of a row and a column identifying a specific key. Each row and column is connected to the keyboard's internal controller which is a microchip contained within the keyboard case. The controller's job is to make sense of these signals and convert them into binary data for transmission to the computer. In actuality the controller detects changes in voltage; as a key is pressed the voltage in that circuit goes from low to high, and similarly when a key is released the voltage returns from high back to low. Every key is associated with a pair of scan codes; the 'make code' is generated as the key is pressed and the second, known as the 'break code', is generated when the key is released. The controller produces these scan codes, stores them in its own internal memory and sends them to the computer via an interface cable.

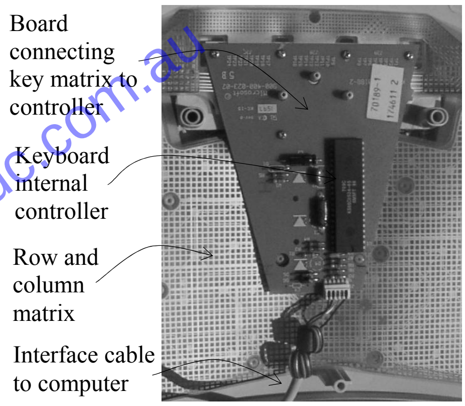


Fig 3.5
Detail of the keyboard controller within Microsoft's 'Natural' Keyboard.

The interface cable contains four active wires; two are used to power the keyboard, one is a clock signal and the last is used for transmission of the scan codes and other data. Commonly the cable connects to the motherboard using a USB port; USB ports use synchronous serial communication; synchronous means the data arrives and departs at a steady rate controlled by the clock signal and serial means a single wire is used to transmit data one bit after the other.

When a series of scan codes arrive at the motherboard they are stored in memory and the operating system is notified using an interrupt request. The operating system, with assistance from the keyboard driver software, then examines the scan codes and responds accordingly. In most cases the scan codes are converted into a representation that includes the key's ASCII code together with information in regard to any modifier keys that may have been used. This data is passed to the currently active application. In other words the operating system transforms the scan code data into information that is meaningful to the application. This means different keyboard layouts are specified at the operating system level rather than at the keyboard itself; *Fig 3.6* shows a screen used to implement this facility within Microsoft's Windows XP; obviously the labels on each key would require alteration to reflect the changes made to such settings. The operating system also intercepts keystrokes that are intended for system level tasks, such as switching between applications, starting new applications or even rebooting the system.

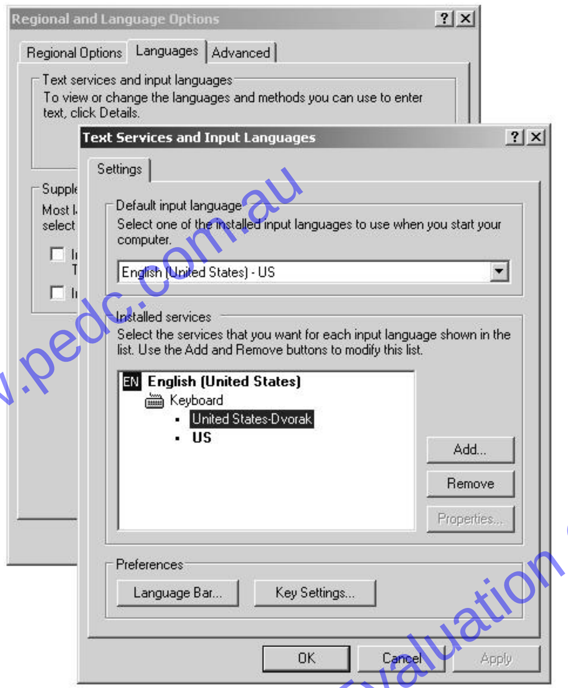
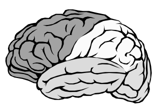


Fig 3.6
Changing the keyboard layout to Dvorak using the control panel in Microsoft's Windows XP.

So far we have only considered the transfer of data from the keyboard to the computer, however data also travels in the other direction. For example when the caps lock is pressed the operating system responds to these scan codes by sending the keyboard a message to turn on or off the caps lock light. There is also data returned to the keyboard each time an error occurs in the transmission of a scan code; each error message signals the keyboard's internal controller to resend the last scan code.



Consider the following:

All keyboards contain groups of keys that perform related actions. Consider the following groupings:

- Alphanumeric and punctuation keys (e.g. A-Z)
- Modifier keys (e.g. Shift)
- Numeric keypad (e.g. 0-9)
- Function keys (e.g. F1)
- Cursor control and navigation keys (e.g. Page Up)
- Other specialised keys (e.g. keys for Internet access)

GROUP TASK Activity
Most standard keyboards contain at least 104 keys. Examine the keyboard you use and classify each of the keys using the above bulleted list.

Consider the following:

The QWERTY layout was allegedly designed to slow typists; old typewriters used hammers that would get caught on each other so all the common letters where moved away from the centre or home row. Other sources indicate the real reason for the design of the QWERTY layout was somewhat less technical; all the letters in the word typewriter were deliberately located in the top row to allow typewriter salesmen to type the word typewriter at incredible speeds! Regardless of the reasons, the QWERTY layout is an inefficient design, but through consistent usage it has and is likely to remain the most widely used. *Fig 3.7* shows a standard Dvorak layout; notice that this layout has the most commonly used letters located in the home row.

Fig 3.7
Dvorak keyboard layout.

Some claims made in favour of the Dvorak layout compared to QWERTY include:

- Finger travel distance is up to 20 times less.
- 70% of letters occur in the home row in Dvorak, compared to 31% in QWERTY.
- The error rate for QWERTY typists is about twice that of Dvorak typists.
- The costs of retraining typists to use Dvorak layouts could be recovered in 10 days, due to increased productivity.
- Dvorak typists experience lower instances of repetitive strain injury (RSI).

GROUP TASK Discussion
Each of the above claims can only be substantiated using solid reliable evidence. Suggest techniques that could be used to collect such evidence.

MOUSE

The basic design of the mouse were first conceived by Douglas Englebart in 1964; it was some 20 years later, when Apple released the Macintosh, that the mouse became the input device of choice. Today it is hard to imagine using a computer without a mouse.

The mouse is primarily used to collect movement data in two dimensions; usually this data is used by the computer to control the position of the cursor on the monitor. In addition mice include a number of buttons and many also include a scroll wheel that doubles as an extra button.

GROUP TASK Activity
There are many other input devices that collect movement data similar to that collected by a mouse. Create a list of such devices.

So what happens when we move a mouse; that is, how does the mouse detect this movement and transform it into digital data for use by the computer? Currently there are two common designs; one based on a rolling ball and another using a purely optical design.

Let us first consider the rolling ball design (see *Fig 3.8*). A ball, inside the mouse case rolls along the desktop as the mouse is moved. The case contains two rolling shafts, one for the X direction and another for the Y direction; these shafts are in contact with the ball hence they revolve as the ball moves. A disk with many small slits around its circumference is attached to the end of each shaft; as each shaft spins then so too does the attached disk. A light emitting diode (LED) is mounted on one side of each disk and an LED sensor on the other side; as the disk revolves the slits allow pulses of light from the LED to reach the LED sensor. The LED sensor, in simple terms, opens a circuit each time it encounters a pulse of light. This pulsating signal is connected to the mouse's internal controller whose job is to count the number of pulses generated by each LED sensor. The controller sends this data to the computer approximately forty times every second; after each send the counters are reset back to zero. So if the mouse is moved fast to the left then maybe 200 pulses will be detected in the X direction and no pulses recorded in the Y direction; the controller sends the binary equivalent of these numbers to the computer.

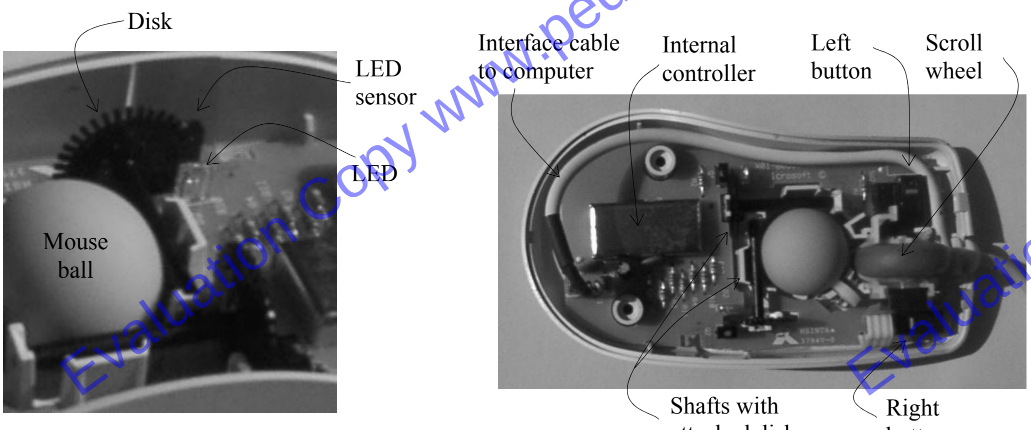


Fig 3.8
Inside a typical rolling ball mouse.

You may have noticed that the above description does not explain how the mouse differentiates between left and right or forward and backward motion. In actuality the LED sensor is really two LED sensors built into a single unit (many mouse designs actually use two pairs of LED and LED sensor for each direction). This twin sensor is positioned such that when one sensor sees the light clearly the other is in a state of change. Without explaining the mathematical details, this provides sufficient information for the controller to determine the direction of movement; which is subsequently sent to the computer along with the number of light pulses detected.

The rolling ball mouse has a major drawback; the rolling action soon picks up any dust and debris present on the desktop. This rubbish rapidly accumulates on all the internal parts; in particular around each of the shafts. Once this occurs it is only a matter of time before the mouse operation is degraded and eventually stops. The obvious solution is to clean the mouse regularly, however perhaps a better solution is a mouse sealed from the outside world that has no moving parts; the purely optical mouse is one such design.

A purely optical mouse replaces all the moving and optical parts with just three components; a red LED, an image sensor and a digital signal processor (DSP). The red light from the LED is reflected off the surface of the desktop and into the lens of the image sensor (see *Fig 3.9*). The image sensor is essentially a mini digital camera; it takes a picture of the desktop some 1500 times per second. Each of these images is sent to the DSP whose primary task is to detect the direction and size of any movement by comparing features in successive images. The precision and speed of the DSP provides far more detailed information in regard to mouse movement than previous technologies; hence an optical mouse provides much smoother response and control for users.

Fig 3.9
Underside of an optical mouse.

Virtually all mouse designs include three buttons; a left and right button together with one activated by pressing down on the scroll wheel. What about the scroll wheel itself; scroll wheels do not rotate smoothly, rather they rotate in a series of clicks, each click is either in the forward direction or in the backwards direction. Consequently the data generated by the scroll wheel is represented identically to that generated by two of the other buttons; either the wheel was clicked forward or it was not, similarly it was either clicked backward or it was not. The data sent to the computer includes information in regard to the state of each of these buttons; each button is either clicked (1) or it is not (0).

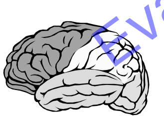
Fig 3.10
A typical mouse containing three buttons together with a scroll wheel.

Let us summarise the data collected by a typical mouse:

- Numbers representing the distance moved in both X and Y dimensions.
- The direction of the movement. Either left or right and either backwards or forward.
- The state of each button; either on or off.
- Scroll wheel events, either forward click or not, and either backward click or not.

This data, in binary form, is generated and sent approximately 40 times every second. Older mice use a PS2 port, whilst most now use a USB port for connection to the computer, hence the method of data transmission is essentially the same as that used for keyboard data.

Fig 3.11
PS2 plug (left) and USB plug (right) used to connect mice.



Consider the following:

The mouse, together with most other types of pointing device, can be used to collect a variety of different media types. This is not usually their primary task, rather they are used to collect information used to initiate or facilitate the actions occurring in other information processes.

GROUP TASK Discussion
Suggest seven scenarios, one for each information process, where a mouse is used to initiate or facilitate the actions of that process.

SCANNER

There are various different types of image scanner; all collect light as their raw data and transform it into binary digital data. This digital data may then be analysed, organised and processed into numbers or text, or it may remain as image data in the form of bitmaps. Perhaps the most familiar forms of scanner are barcode scanners, used in most retail stores and flatbed scanners used to collect images in bitmap form. Let us consider the operation of common examples of each.

Barcode scanners operate by reflecting light off the barcode image; light reflects well off white and not very well off black. This is the basic principle underlying the operation of all types of scanner. A sensor is used to detect the amount of reflected light; so to read a barcode we can either progressively move the light beam from left to right across the barcode or use a strip of light in conjunction with a row of light sensors. Each of these techniques are used for different designs of barcode scanner; those based on LED, laser and CCD technologies dominate the market, *Fig 3.12* shows an example of each. Most barcode scanners incorporate a decoder to organise the data into a character representation that mimics that produced by the keyboard. This means most barcode readers can be installed between the keyboard and the computer without the need for dedicated interface software.

Fig 3.12
Clockwise from top-left: LED wand, multi-directional laser and CCD based barcode scanners.

Barcode wands use a single light emitting diode (LED) to illuminate a small spot on the barcode. The reflected light from the LED is measured using a single photocell. As the wand is steadily moved across the barcode, areas of high and low reflection change the state of the photocell. The photocell absorbs photons (a component of light). As the intensity of photons absorbed increases so to does the current flowing through the photocell; large currents indicating white and smaller currents indicating black. This electrical current is transformed by an analog to digital converter (ADC) to produce a series of digital ones and zeros. The same LED technology is used for slot readers, where the barcode on a card is read by swiping the card through the reader.

GROUP TASK Activity
A barcode is scanned using an LED barcode scanner and the following stream of bits is produced: 000000110011000000111111001100111111
Draw the most likely original barcode.

Lasers are high intensity beams of light and as such they can be directed very precisely. Laser barcode readers can therefore operate at greater distances from the barcode than other technologies, commonly up to about 30cm away. The reflected light from the laser is detected by the photocell using the same technique as LED scanners. There is no need to manually sweep across the barcode as the laser beam is moved using an electronically controlled mirror. Basic models continually sweep back and forth across a single path, whilst more advanced models perform multiple rotating sweeps that trace out a 'star like' pattern. These advanced models are much more effective as the user need not hold the scanner parallel to the barcode; rather the scanner rotates the scan line until a positive read is collected. Supermarkets often use this type of barcode scanner mounted within the counter top.

Charge coupled devices (CCDs) contain one or more rows of photocells built into a single microchip. CCD technology is used by many image collection devices including CCD barcode scanners, digital still and video cameras, handheld image scanners, and also flatbed scanners. For both barcode and image scanners a single row CCD is used. The light source for these scanners is typically a single row of LEDs with the light being reflected off the image back to a mirror. The mirror reflects the light onto a lens that focuses the image at the CCD. Each photocell in the CCD transforms the light into different levels of electrical current. These levels are converted into bits using a similar technique to that used in LED and laser barcode scanners. CCDs in flatbed scanners differ slightly; they convert the electrical current from each photocell into a binary number, normally between 0 and 255, using a more complex analog to digital converter (ADC).

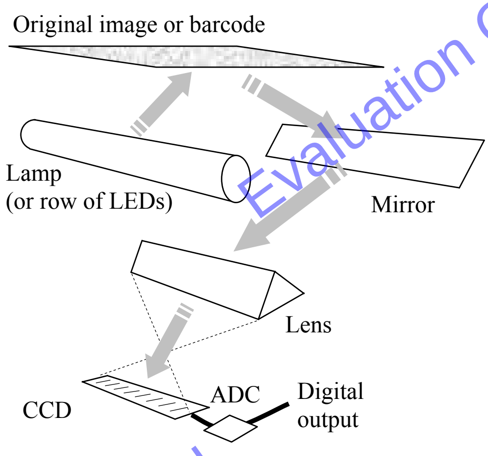


Fig 3.13
The components and light path typical of most CCD scanner designs.

GROUP TASK Investigation
Barcode scanners are used in most retail stores and libraries. Over the next 24 hours observe closely each barcode scanner you encounter. Classify each as using either LED, laser or CCD technology. Justify your choices.

Let us now consider flatbed scanners based on CCDs in more detail. This type of flatbed scanner is by the far the most common; scanners based on other technologies are available, but currently they fall into the higher quality and price ranges. We mentioned above that the binary numbers returned from a flatbed scanner's ADC range from 0 to 255; this is the range of different numbers that can be represented using 8 bits (1 byte). If white light is used then these numbers will represent shades of grey, ranging from black (0) to white (255). So how do flatbed scanners collect colour images? Quite simply, they reflect red light off the original image to collect the red

component, green to collect the green component and blue for the blue component. Some early scanners performed this action by doing three passes over the entire image using a different coloured filter for each pass; this technique is seldom used today. Today most scanners use an LED light source that cycles through each of the colours red, green, blue; hence only a single pass is needed.

The LED lamp, mirror, lens and CCD are all mounted on a single carriage; these components are collectively known as the scan head. All the components on the scan head are the same width as the glass window onto which the original image is placed. This means a complete row of the image is scanned all at once. The number of pixels in each row of the final image is determined by the number of photosensors contained within the CCD; typical CCDs contain some 600 sensors per inch, predictably this results in images with horizontal resolutions of up to 600 dpi (dots per inch).

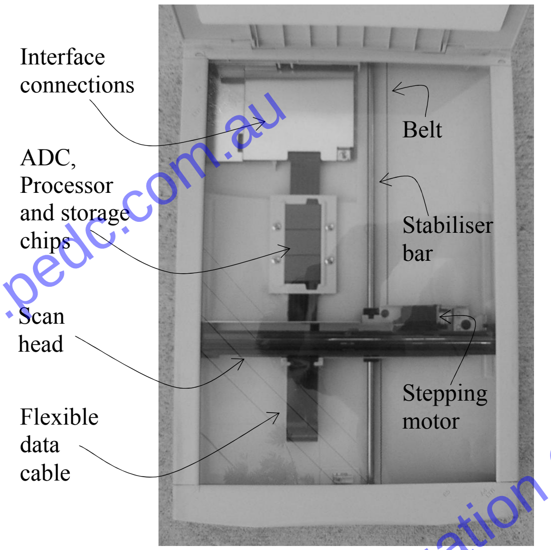


Fig 3.14
Components of a flatbed scanner.

So what operations occur to collect a colour image?

- The current row of the image is scanned by flashing red, then green, then blue light at the image. If you open the lid of a scanner you'll predominantly see white light, this is due to the colours alternating so rapidly that your eye merges the three colours into white. After each coloured flash the contents of the CCD is passed to the ADC and onto the scanner's main processor and storage chips.
- The scan head is attached to a stabilising bar, and is moved using a stepping motor attached to a belt and pulley system. The stepping motor rotates a precise amount each time power is applied; consequently the scan head moves step by step over the image; pausing after each step to scan a fresh row of the image. The number of times the stepping motor moves determines the vertical resolution of the final image.
- As scanning progresses the image is sent to the computer via an interface cable. The large volume of image data means faster interfaces are preferred; commonly SCSI, USB or even firewire interfaces are used to connect scanners. Once the scan is complete the scan head returns back to its starting position in preparation for the next scan.

GROUP TASK Discussion
Some scanners use 36 or even 42 bits internally to represent each pixel, yet they only output 24 bit per pixel images. Why would this be? Discuss

GROUP TASK Discussion
The packaging of a scanner implies it is able to scan at 2400dpi, you know the CCD contains just 600 sensors per inch. What is going on? Discuss

Pages 91 to 100

Not included in this Evaluation Copy

HSC style question:

(a) Outline the technology and processes occurring as a digital camera collects and digitises image data.

(b) Outline the technology and processes occurring as a flatbed scanner collects an image.

(c) Outline the technology and processes occurring as sound is recorded using a microphone, sound card and computer to produce a sampled digital file.

Suggested Solution

(a) At the back of a digital camera is a CCD (or a CMOS chip) which contain a photocell for each pixel. The photocells respond to the amount of light falling on them through the lens of the camera. As the photocells in a CCD are hit by incoming light, they emit a current relative to the brightness of the light. Different colours are detected through the use of a red, green and blue bayer filter covering the photocells. The electrical current is converted using an analog to digital convertor into an equivalent bit pattern for each pixel. The digital data is then processed to generate a red, green and blue colour value based on the values of the adjoining pixels. Finally the digital image data is compressed and stored on the camera's flash card.

(b) A light source and row of sensors are mounted on a head assembly that moves in precise steps over the entire image being scanned. The more light that is reflected, then the lighter that part of the image must be. A series of focusing mirrors and lenses focus the reflected light onto the row of light sensors, which is typically a CCD. As the reflected light hits the sensors a current is produced – the more light, the stronger the current. An ADC (analog to digital converter) then converts this electrical signal into an equivalent series of bits for that section of the image. For colour scanning, there are 3 coloured lights, one for each of red, green and blue light. Each coloured light flashes in sequence over each line of the image, hence the CCD collects the intensity of red, green and blue light. The equivalent bit patterns representing the colour of each pixel of the image can then be determined.

(c) Sound waves are detected by the microphone as a sequence of compressions and decompressions in the air. These movements in the air move a diaphragm inside the microphone which in turn generates an equivalent current. This analogue current is sent along the connecting cable through to the sound card. At the sound card, the current is sampled many thousands of times a second and put through a ADC. Each sample represents the height of the original sound and is converted by the ADC into a binary number. This effectively converts the sound wave into a long sequence of binary numbers. Most sound card also contains a digital signal processor (DSP) which smooths the transitions and filters out noise. The resulting sound samples are then transferred via the system bus to the CPU where they are directed to a storage device.

SET 3B

1. The main difference between CCDs used in digital cameras and those used in flatbed scanners is:
   (A) digital camera CCDs have a single row of photosites, flatbed scanners have a two dimensional grid of photosites.
   (B) flatbed scanner CCDs have a single row of photosites, digital cameras have a two dimensional grid of photosites.
   (C) flatbed scanners generate their own light, digital cameras use natural light.
   (D) digital cameras generate their own light, flatbed scanners use natural light.

2. A bayer filter:
   (A) is used to remove unwanted detail during image collection.
   (B) converts analog data into digital data.
   (C) alternates red and green rows with blue and green rows.
   (D) is used to compress image data.

3. Microphones are used to:
   (A) convert sound waves into bits.
   (B) convert sound waves into electrical energy.
   (C) convert electrical energy into bits.
   (D) All of the above.

4. The component on a sound card that filters and compresses the audio data is known as a:
   (A) ADC
   (B) CCD
   (C) SAR
   (D) DSP

5. CCDs that contain two layers of sensors are commonly used in:
   (A) analog video cameras.
   (B) digital video cameras.
   (C) digital cameras.
   (D) Both (A) and (B).

6. The main components of a dynamic microphone include:
   (A) two plates and a power source.
   (B) a diaphragm, wire coil and magnet.
   (C) a capacitor, comparator, DAC and SAR.
   (D) a wire coil, ADC and DSP.

7. Which of the following is true for progressive scan CCDs?
   (A) Every second line of each image is retained.
   (B) The data is collected to suit normal analog television.
   (C) The entire contents of the CCD is read for each image collected.
   (D) They use interlacing to reduce the amount of data.

8. Mechanisms on traditional cameras to control the amount of light entering the camera include:
   (A) altering the size of the aperture and the time shutter is open.
   (B) altering the size of the shutter and the time the aperture is open.
   (C) the use of different types of flim that have varying sensitivities to light.
   (D) moving the lens in and out to focus the light more accurately.

9. One issue to consider when using vehicle counters based on a single air pressure switch is:
   (A) the road temperature and air temperature commonly cause false readings.
   (B) they are unable to produce digital data.
   (C) they cannot differentiate between vehicles with two axles and larger vehicles with more than two axles.
   (D) the vehicle must be stationary if the magnetic field is too influence the induction loop.

10. The Safe-T-Cam system uses a video camera and a still camera. Why is this?
   (A) Still images are used to isolate individual vehicles from the background, video is used to determine the registration.
   (B) Video is used to isolate individual vehicles from the background, still images are used to determine the registration.
   (C) The still images are used as a backup should the video camera fail.
   (D) This is not true, Safe-T-Cam only uses a video camera..

11. Describe the operation of a digital camera.
12. Describe the operation of a condenser microphone.
13. Explain how an ADC performs its conversion using the services of a DAC.
14. Explain the differences between CCDs used in digital still cameras and those used in video cameras.
15. Research and describe how speed cameras operate.

SOFTWARE USED FOR COLLECTION

> **Software**
> The instructions that control the hardware and direct its operation.

Software is comprised of instructions that control the hardware and direct its operation. There are two general types of software present in all computer systems; system software and application software. System software includes the operating system and device drivers for each hardware device. In regard to collection, these software components communicate with each hardware collection device and with the current software application. That is, they provide the interface between application software and hardware collection devices. Application software, such as word processors, spreadsheets and databases, receive the collected data and process it during other information processes.

In many cases the source of the data is the user, and in virtually all software applications the user controls or initiates the collection process; the application software must display data entry forms to enable the efficient collection of data. Therefore software performs two essential tasks during the collection process; it provides an interface with the collection hardware and it provides a mechanism for data entry.

Existing data is often moved from other systems rather than being collected directly. For instance, a large variety of different data is available from the Australian Bureau of Statistics. This data can be converted into a form suitable for use by a variety of different applications. We consider a variety of examples where participants can move data between different applications.

In this section we consider:

- Device drivers that allow hardware to interface with the operating system.
- Software that allows participants to enter data.
- Software that allows participants to import data.

DEVICE DRIVERS THAT ALLOW HARDWARE TO INTERFACE WITH THE OPERATING SYSTEM

> **Device driver**
> A program that provides the interface between the operating system and a peripheral device.

Most hardware collection devices, and most other devices connected to computers, are used by a variety of different software applications. It would be most inefficient for each software application to contain its own set of instructions for communicating with each device. It makes more sense to store all the instructions required to control and communicate with a particular device separately. These software programs are called 'device drivers'. Most hardware manufacturers develop and distribute device drivers specific to each of their hardware devices. For example, if you buy a new printer, the packaged software accompanying the device will contain a device driver from the manufacturer. Most operating systems also include various different device drivers that are capable of communicating and controlling common hardware devices, however advanced features of specific devices may not be supported. For example the generic scanner device drivers supplied with an operating system are unlikely to support advanced features such as document feeders.

When a hardware collection device, and most other peripheral hardware devices, wish to send data to the system they communicate via their device driver; the device driver provides the software interface between the hardware device and the operating system and application software. Messages to control the process are sent by the hardware to the device driver which in turn communicates with the operating system. The operating system determines which software application requires the data and informs the software application. The software application then receives the data from the hardware collection device via the device driver. The operating system controls the whole process by communicating with the device driver and the software application; *Fig 3.29* describes this process.

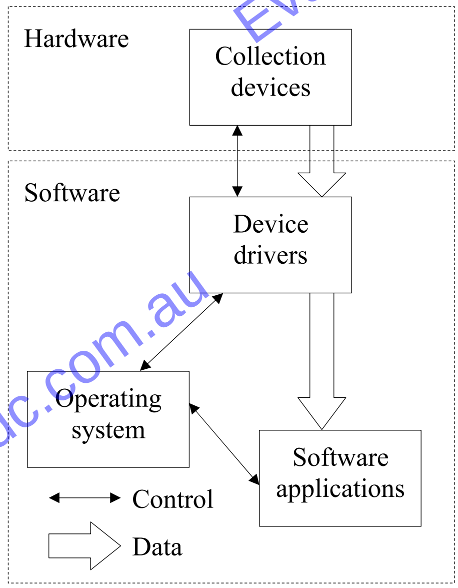
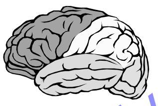


Fig 3.29
The software interface between collection devices and software applications.



Consider the following:

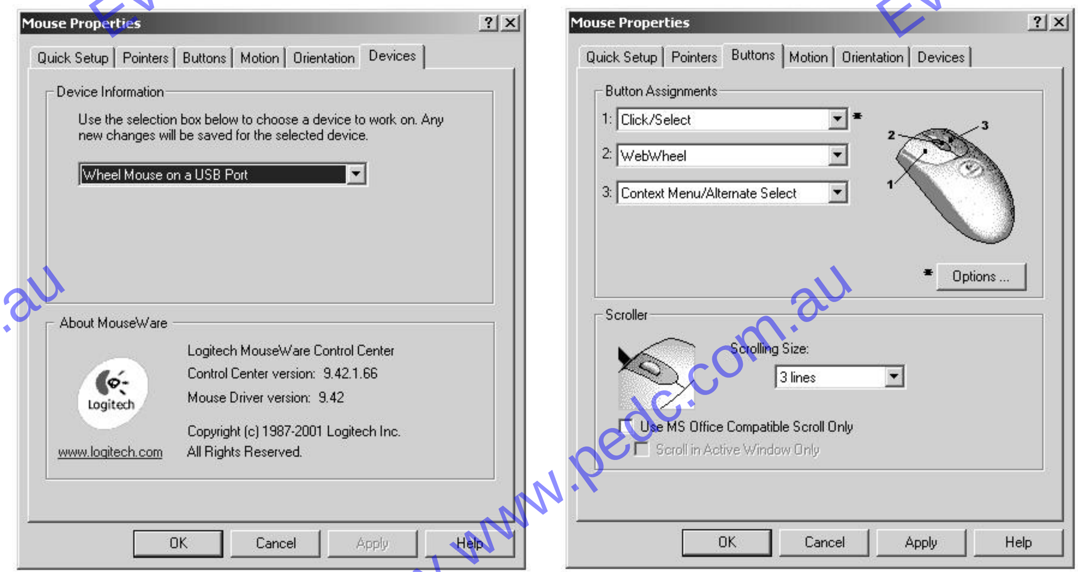
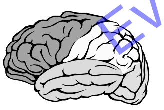


Fig 3.30
Screen shots from the user interface of a Logitech mouse driver used in Windows XP.

The screen shots in *Fig 3.30* above form part of the user interface of a Logitech mouse driver designed for use with the Windows XP operating system. These screens allow the user to alter characteristics of the mouse's device driver which in turn affects the operation of the mouse within all software applications.

GROUP TASK Discussion
Discuss how changes made to the above screens cause the mouse to operate in the same way within all software applications installed on the computer. Consider the flow of data from the mouse until it reaches the application as part of your discussion.



Consider the following:

The screen shots in *Fig 3.31* are for a keyboard installed in Windows XP via a USB port. This keyboard is described as an 'HID Keyboard Device'; HID is an acronym for 'Human Interface Device'. HID devices include most common hardware devices for collecting data from the user; this includes barcode scanners, mice, joysticks and keyboards. HID is a standard that forms part of the USB standard. HID compliant devices do not require their own device drivers; rather they use the HID device driver included with the operating system.

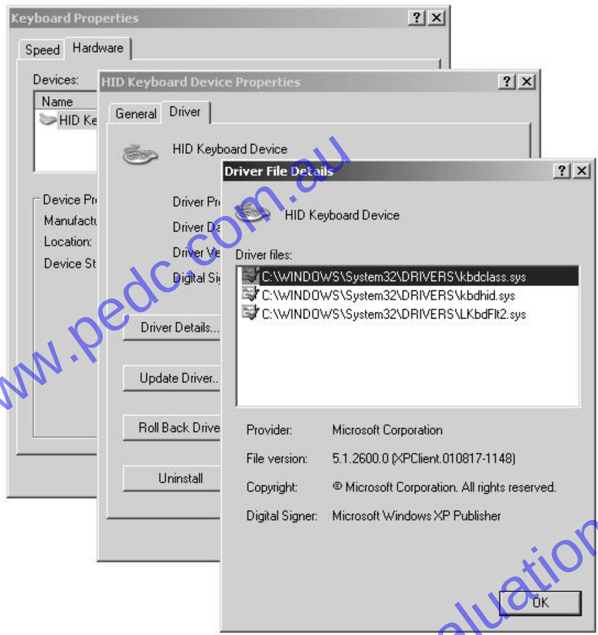


Fig 3.31
Driver details for a standard HID keyboard in Windows XP.

When an HID compliant device is first plugged in, the HID device driver accesses data from the device in regard to its operation. This data provides information specific to the particular device in a similar way that a dedicated device driver provides specific information about a device. As a consequence specific functionality available in the new device can be used without the need to install a dedicated device driver.

GROUP TASK Investigation
Examine the device drivers for each hardware collection device installed on either your school or home computer. Do any of these devices utilise the HID standard?

GROUP TASK Discussion
List and describe advantages and disadvantages of standards, such as the HID standard, for both users and hardware manufacturers.

SOFTWARE THAT ALLOWS PARTICIPANTS TO ENTER DATA

Application software
Software that performs a specific set of tasks to solve specific types of problems.

Virtually all software applications collect data directly; hence a means for data entry must be provided. Most applications use the keyboard and mouse in conjunction with data entry forms displayed on the screen to enter data; this is particularly the case for text and numbers, together with computer generated image, audio and video data. Even image, audio and video data collected using various other hardware collection devices use the keyboard and mouse to collect data to control the operation of the device.

The aim of all data entry screens is to collect data in the most accurate and efficient manner. The accuracy of data within a system is known as 'data integrity'. Data integrity is the aim of all information systems; the data needs to be correct and accurately reflect its source at all times. High quality data has high levels of data integrity. Achieving high levels of data integrity is a time consuming and difficult task for most information systems. It is an ongoing process whereby new and existing data is repeatedly checked for accuracy, not just at the time of collection but throughout the life of the data.

Data Integrity
Occurs when data is correct and accurately reflects its source. The quality of the data.

Ensuring the integrity of data during collection is accomplished using data validation checks as well as data verification checks. The efficiency of data entry processes is largely determined by the design and behaviour of the data entry screens; the user interface. Let us consider data validation, data verification and some basic principles in regard to user interface design.

Data Validation

The computer performs data validation as each data item is entered. Data validation ensures each data item is reasonable. For example, when entering the cost of a product, data validation criteria would likely include checks for a number and ensuring that the number is positive. Data entry screens often use self-validating components that ensure only valid data can be entered. For example sets, of radio buttons restrict the range of data that can be entered to one of the available choices, hence radio buttons are said to be self-validating.

Data Validation
A check, at the time of data collection, to ensure the data is reasonable and meets certain criteria.

GROUP TASK Activity
Examine data entry screens from software applications installed on your school or home computer. Describe the differents types of data validation used for each component on these screens.

Data Verification

Data verification is a more difficult task than validating the reasonableness of the data. For example, the computer can quite easily check that a phone number contains the correct number of digits however verifying that these digits are indeed the persons phone number is a more difficult task. Furthermore, people often change their phone numbers therefore data verification must be ongoing. Data verification includes all the procedures that are used to verify the correctness of the data within an information system. In regard to data entry into application software, data verification is often implemented as a procedure whereby the user compares the source data to the data just entered. For example, when taking a credit card order over the phone, the operator verifies the credit card number entered by reading it back to the customer.

Data Verification
A check to ensure the data collected matches the source of the data.

User Interface Design

The aim of the user interface is to guide the user through the collection process in such a way that the data is collected accurately and efficiently. The user interface is more than just the placement of components on the screens; rather it provides the total interaction between the user and the software application. In regard to collecting data the user interface displays information to the user to guide them through the collecting process.

There are numerous design factors that influence the efficiency and accuracy of user interfaces. The study of user interface design is itself a complete discipline; nevertheless let us consider some basic principles that could be used when assessing the quality of user interfaces.

> **User Interface**
> Part of a software application that displays information for the user. The user interface provides the means by which users interact with software.

- Know who the users are. What are their goals, skills, experience and needs? Answers to these questions are required before an accurate assessment of the user interface can be made. For example, a data entry screen that will be used every day by data entry operators will be quite different to one used infrequently by unskilled users, such as members of the public.
- Consistency with known software and also consistency within the application. Users expect certain components to operate in similar ways and to be located in similar locations. For example, the file menu is located in the top left hand corner of the screen and placing it elsewhere would be inconsistent and confusing. Consistency allows users to utilise there existing skills when learning new software applications.
- Components on data entry screens should be readable. This includes the words used as well as the logical placement and grouping of components. The interface should include blank areas (white space) to visually imply grouping and to rest the eye. Colour and graphics should be used with caution and only when they convey information more efficiently than other means.
- Clearly show what functions are available. Users like to explore the user interface; this is how most people learn new applications, therefore functions should not be hidden too deeply. If a particular function is not relevant then it is better for it to be dulled or greyed out than for it to be hidden, this allows users to absorb all possibilities. At the same time the user interface should not be overly complex.
- Every action by a user should cause a reaction in the user interface. This is called feedback; without feedback that something is occurring, or has occurred, users will either feel insecure or will reinitiate the task in the belief that nothing has happened. Feedback can be provided in simple ways; such as the cursor moving to the next field, a command button depressing or the mouse pointer changing. Tasks that take some time to complete should provide more obvious feedback indicating the likely time for the task to complete.
- User actions that perform potentially dangerous changes should provide a way out. Many modern software applications include an 'undo' feature, whilst others provide warning messages prior to such dangerous tasks commencing. In either case the user is given a method to reverse their action.

Consider the following data entry screen:

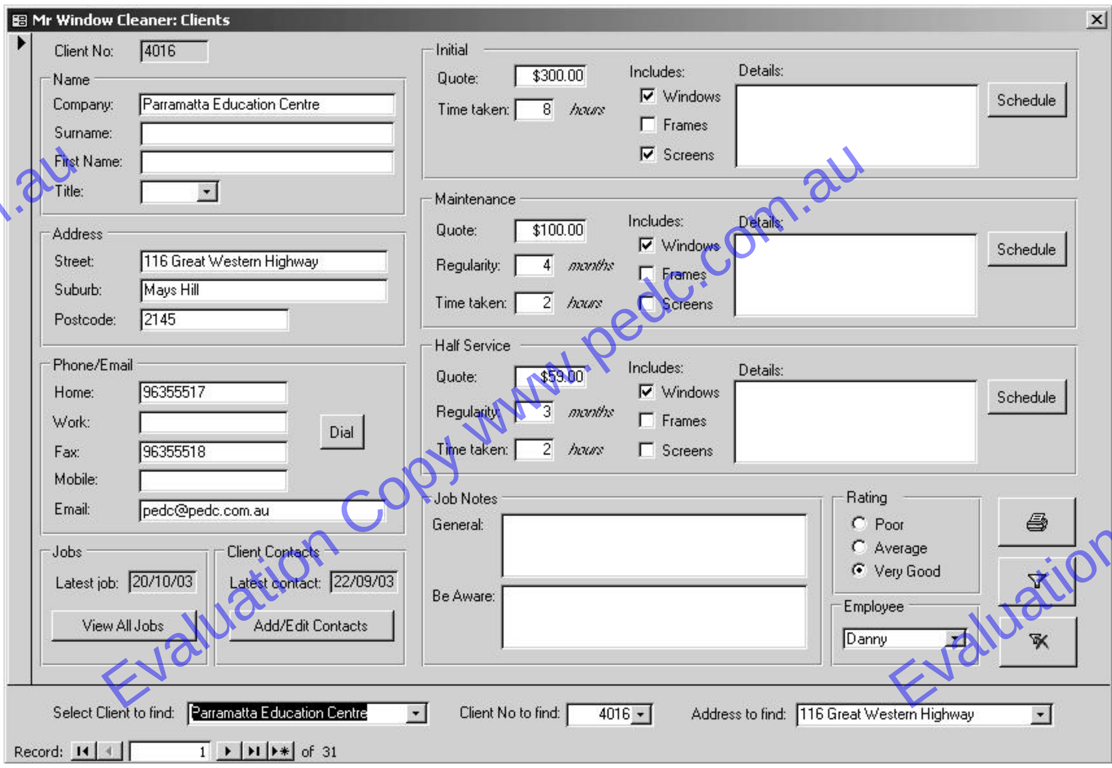


Fig 3.32
A data entry screen used to collect and display client information.

The above data entry screen is used to collect and display client information for a window cleaning company. The screen is linked to the company's phone system; as the phone rings the information system retrieves the 'Caller ID' and this data is then compared to the phone number details held in the database. If a match is found then the appropriate client details are displayed. At some stage during the call the operator confirms that the data held in the first column on the screen is correct. If no match to the Caller ID is found, then a blank client screen is displayed and in this case the operator must first perform a search operation using one or more of the find components at the bottom of the screen. If the client cannot be located then a new record is created.

Only the left hand column of data is entered during a phone call for a new client, the remaining data is collected via a paper form by the window cleaner assigned to complete the first job or quote. The paper form from the window cleaner provides the source of the data for the remaining data items on the screen.

GROUP TASK Discussion
List and describe techniques used on the above screen and in the above scenario to improve the data integrity of the client data.

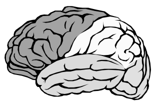
GROUP TASK Discussion
Critically evaluate the above data entry screen based on the user interface design principles outlined on the previous page.

Collection via data entry web pages hosted on a web server.

> **Browser**
> A software application that interprets HTML code into the words, graphics and other elements seen when viewing a web page from a web server.

Data entry screens in the form of web pages can be created to collect data from users visiting the web site. These data entry screens perform in much the same way as other data entry screens. One major restriction is the speed with which the data can be validated. Simple data validation can be carried out by the browser within the downloaded web page (often using JavaScript) whereas any validation that requires examining the data source held on the web server will take time to occur and is often implemented as a batch process after a complete page of data has been entered.

There are various server-side technologies available for generating data entry web pages. At the time of writing many data entry web pages are implemented using server-side scripting languages such as PHP (Hypertext Pre-processor), PERL (Practical Extraction and Report Language) or ASP (Active Server Pages). All these server-side technologies use programming code executing on the web server to generate HTML files for transmission to the user's web browser. Server-side scripts can contain instructions that retrieve data from a database and place it into the page prior to its delivery; similarly the data entered by the user is returned to the web server where it is stored in the database. These server-side technologies mean that specific data can be displayed for the specific user and collected from the user interactively, yet no additional software is required on the user's machine.



Consider the following:

Google is a popular search engine used to search the World Wide Web. The screen shot in *Fig 3.33* shows Google's advanced search data entry screen displayed within Microsoft's Internet Explorer browser.

Self-validating screen components, namely drop down boxes and radio buttons, are used for data validation. When a user clicks the 'Google Search' command button the data entered by the user is sent to Google for processing and the results of the search are then returned for display in the user's browser.

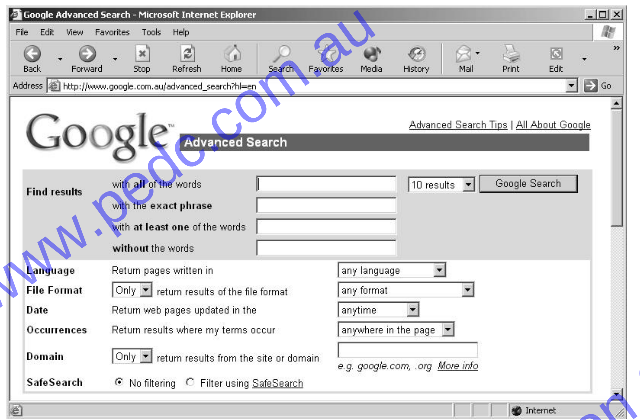


Fig 3.33
Google's advanced search data screen displayed in Microsoft's Internet Explorer browser.

> **GROUP TASK Discussion**
> Identify the source of the data from the perspective of Google and then from the perspective of a user collecting data from a website found via Google's search engine. Discuss your responses.

SOFTWARE THAT ALLOWS PARTICIPANTS TO IMPORT DATA

Many software applications use data created or originally collected by other applications. The data is moved across for further processing in some other software application. For example, images are collected using a digital camera and then moved to a computer where they are edited using a graphics application. These images are then imported into a broad range of software applications, such as word processors, desktop publishing applications and HTML editors. Data held in databases is often moved to or imported into other applications for specialised processing. For example, word processors use databases when mail merging to create personalised letters. Spreadsheets import data from databases so they can perform statistical analysis or create charts. Web browsers in combination with search engines are used to locate and download data over the Internet. This data can then be used by other software applications. In essence, the destination software application is collecting existing data created by another application.

Importing data often, but not always, involves altering the organisation of the data to suit the needs of the destination software application. Therefore, although importing is essentially a collecting process, it often includes organising and perhaps other information processes as sub-processes. For example, if a photograph is to be used on a web page then the image would likely be resized, compressed and saved using the JPEG format. In this case the conversion changes the data and its file format so both processing and organising processes are performed by the graphics application. The HTML editor imports the resulting JPEG image file without altering its content or organisation. In other cases the source application exports the data in a format understood by both applications. In this case both the source and destination applications reorganise the data. For instance, many database management systems (DBMSs) are able to output delimited text files. These text files can be imported into spreadsheets. In this case, the DBMS converts or reorganises the data from its native format to create the delimited text file. The spreadsheet performs a conversion from the text format to reorganise the data into its own format. There are also scenarios where just the destination or importing application performs the conversion. For example, most word processors include the ability to import files that are in the native format of other word processors – for instance, Microsoft Word can import files in WordPerfect format. In this case Microsoft Word performs the reorganisation as it imports or collects the WordPerfect file.

Web browser and search engines used to locate and download Internet data.

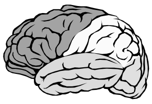
Web browsers are software applications used to collect data and information from web servers. The problem is not so much actually getting the information from the web server to your browser, rather the problem is to locate the web server that holds the information you require. Search engines are websites dedicated to assisting in the task of locating information on the World Wide Web. When using a search engine you are not directly searching the web, rather the search engine queries its own database for possible URLs of relevant web sites. When suitable data or information is located on a website it can be collected by copying and pasting, downloading files or by saving the HTML code. When collecting data in this way it is difficult to control the format of the data and even more difficult to reliably assess its integrity.

GROUP TASK Discussion
Why is it difficult to control the format and assess the integrity of data collected directly from the web using a web browser? Discuss.

Pages 111 to 120

Not included in this Evaluation Copy

- Individuals should be informed about the purpose of data collection. This includes any legal obligations, such as taxation office requirements, as well as any possible consequences of not providing personal details. For example, "if an ABN is not provided then you will not be considered for this contract."
- Whenever possible, personal details should be collected directly from the individual; the aim is for individuals to have a clear understanding of who holds their personal details. If secondary sources are used to collect personal information then individuals should be informed of any organisations that will later use the information.
- Sensitive personal information should not be collected without individuals giving their specific consent. Convincing reasons for collecting such information must also exist and be made clear to the individual. For example, a blood bank is required by law to collect data about each donor's sexual preferences.



Consider the following:

1. You have just purchased a new pair of jeans and the shop assistant asks for your name, address and phone number.
2. On a job application one of the questions asks if you have a criminal record.
3. You subscribe to an Internet newsgroup, which involves entering your email address. Subsequently you begin receiving various marketing emails from other businesses.

GROUP TASK Discussion
Identify privacy issues present in each of the above scenarios. Discuss suitable techniques that should be used during the collecting process to help resolve each of the issues you have identified.

ERGONOMICS FOR DATA ENTRY PARTICIPANTS

Extended periods of time entering data magnifies the possible effects of any ergonomic inadequacies. In Chapter 1, we listed a number of broad ergonomic issues; it may be worth reviewing page 27 to familiarise yourself with these issues.

It is generally accepted that participants who spend more than two hours of their day at a computer workstation are susceptible to health problems including repetitive strain injury (RSI), vision problems and general muscle strain. Most data entry operators spend far in excess of two hours per day at their workstations, so the risk of such health problems occurring is significant.

Vision problems and muscle strains experienced by data entry participants are commonly caused by muscles being held in a single strained position for an extended length of time. In relation to vision problems, the muscles controlling the eyes are focussed at a set distance on the screen. Similarly muscles in the back, shoulders, neck and arms are held still during data entry. Muscles within the body are designed to expand and contract; this movement causes blood to flow freely to each muscle. When a muscle is held in a static position the blood does not flow freely, hence oxygen supply to the muscle is reduced and waste products are not efficiently removed; the result being the pain experienced by the operator. Such problems are rarely long term and can be corrected by improved ergonomics.

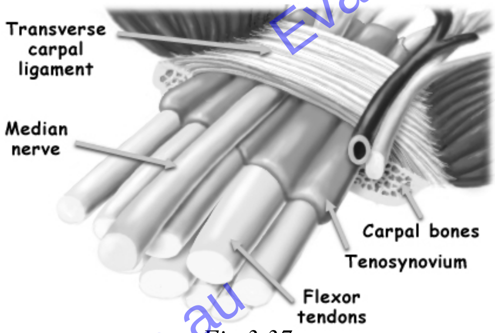


Fig 3.37
Detail of the carpal tunnel showing tendons, median nerve and tenosynovium.

Repetitive strain injury, which is also known as occupational overuse syndrome (OOS) is a much more serious problem. RSI is caused by continually performing the same task; muscles and tendons are not designed for such repetitive tasks. Almost any part of the body is susceptible to RSI, however in relation to computer users; arms, wrists and fingers are the most likely victims. The most common type of RSI caused by repetitive keyboard use is called "Carpal Tunnel Syndrome". The carpal tunnel (see *Fig 3.37*) is within the wrist and is surrounded by the transverse carpal ligament; this ligament surrounds most of the tendons that operate the fingers. When these tendons are overused the lubricating sheath (tenosynovium) around each tendon swells causing restrictions within the carpal tunnel. Such restrictions result in pressure being applied to the median nerve within the carpal tunnel; the result being the characteristic numbness of the fingers experienced by sufferers. RSI, and in particular carpal tunnel syndrome, can result in long-term damage, therefore it is vital to prevent such problems occurring using sound ergonomics within the workplace.

GROUP TASK Discussion
Sufferers of carpal tunnel syndrome are often prescribed wrist braces or anti-inflammatory drugs. In severe cases surgery is used to cut the transverse carpal ligament. How do you think each of these techniques could help relieve the symptoms of carpal tunnel syndrome? Discuss.

The two most significant ergonomic considerations for preventing vision, muscle strain and RSI problems are the design of the work routine and the design and adjustment of equipment. Let us consider both of these in more detail.

Design of the work routine

The aim is to design a work routine that allows data entry operators the ability to change their physical position regularly and to design tasks that do not require repetitive actions for extended periods of time. To accomplish this aim each data entry operator should be assigned a variety of different tasks and they should have control over the order in which they complete these tasks. Tasks that do require significant time at the keyboard should be interspersed with other tasks or with rest breaks.

Everyone has different needs in regard to their most suitable work routine; therefore it is not appropriate to insist that all operators complete tasks in the same order and for the same length of time, rather each operator should have the freedom to complete their tasks in the order that best suits their needs. Structuring the work routine to suit each individual not only assists in directly preventing injuries but it also increases job satisfaction for each operator. Improved levels of job satisfaction lead to increased productivity, resulting in increased profits for the business.

GROUP TASK Discussion
Studies have shown that employees with higher levels of job satisfaction experience a far lower number of workplace injuries. Suggest likely reasons why this is the case.

Design and adjustment of equipment

The importance of ergonomically designed and adjusted equipment increases with the amount of time spent at the computer. As data entry operators spend more than two hours at a computer workstation they should take particular care in regard to the design, placement and adjustment of the equipment they use.

Chairs, desks (or keyboard) and monitors should all be height adjustable. Often desks do not allow their height to be adjusted, if this is the case then a footrest of variable height may be required. Many desks designed for computer usage incorporate a height adjustable keyboard panel.

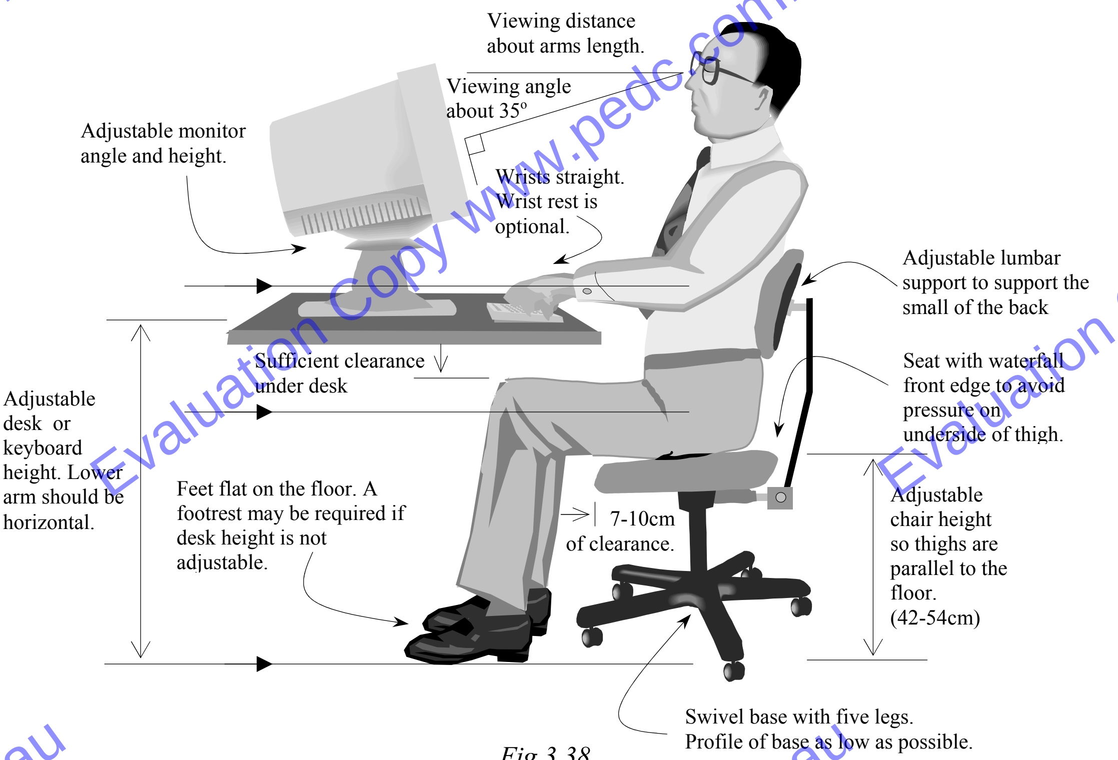


Fig 3.38
Features of an ergonomically sound workstation.

Fig 3.38 shows many of the recommended adjustments and features that should be present in the design of an ergonomically sound workstation. The chair height should first be adjusted so that the thighs are parallel to the floor when the feet are flat on the floor. Next the height of the desk (or keyboard) is adjusted so the forearms are parallel with the floor when using the keyboard. Finally the monitor height is adjusted so the centre of the screen is viewed approximately 35 degrees below the horizontal. Various minor adjustments can then be made to each individual piece of equipment to ensure all muscles are maintained in a relaxed state.

FORUM

GROUP TASK Activity
Try to adjust the equipment for either your school or home computer workstation to comply with *Fig 3.38*. List any items that are not present or adjustments that are not possible.

FORUM

GROUP TASK Discussion
In Chapter 1 on page 27 we listed other ergonomic considerations apart from work routine and equipment design and adjustment. Discuss the significance of these other considerations for data entry participants.

HSC style question:

A market research company has developed a paper-based survey asking people to rate specific products they use on a ten-point scale from poor to excellent. Surveys will be answered anonymously. However, each survey includes a unique Survey ID number. An extract from an example survey is reproduced below – actual surveys will typically contain hundreds of products. It is anticipated that hundreds or even thousands of people will complete surveys.

Survey ID: 345289509		I use this product	Rating Poor ....Excellent									
Code	**Product**		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1436	Heinz baked beans											
2845	Black and Gold baked beans											
1865	Homebrand baked beans											

(a) The intention of the survey was for people to rate only the products they use. During the collection process it is found that some people have rated products they do not use and others have not rated products that they do use. Explain how such issues could have been avoided using an online survey form.

(b) During data collection using the paper forms it is noticed that many individuals do not use the full range of possible ratings. For example, some people rate all products as either poor (1) or excellent (10), others use a small range of ratings such as from 3 to 7 and many use just three ratings, often 2, 5 and 9.

Describe likely effects on the survey results and describe possible techniques for minimising these effects.

Suggested Solution

(a) Validation during online entry would avoid these issues. The software could dull the rating field so that users cannot proceed to enter the rating until a product has been selected as one they use - perhaps from a list box. Equally, once a product has been selected, the user cannot proceed to select another product until a rating has been clicked on or entered for the current product.

The problem with the anonymous paper-based form is that once the form is sent in, it is impossible to check back with the person what their intentions were, if data is not entered correctly.

(b) The effect on the results of the survey is that any findings will not necessarily be valid. The problem is likely due to the fact that ‘poor’ and ‘excellent’ are qualitative measures that do not have the same meaning for all people. This is made more complex by providing a scale of 1 to 10, when perhaps a smaller range, say from 1 to 5, along with descriptions for each rating might be easier for people to intuitively use. Currently it seems people are using different numbers to mean the same thing, for example, 10, 9 and 7 are all maximums for some people. It is therefore unclear if averaging the ratings for each product will accurately reflect people’s overall satisfaction with each product.

Another possible strategy could be to statistically adjust each individual’s ratings so they cover a more typical range. This may increase the accuracy of the results, however it may also have the opposite effect, causing the results to be skewed.

SET 3D

1. The data and information contained in formally published books is, in general, more accurate than data and information sourced using Internet. One reason for this is:
   (A) The Internet is susceptible to viruses that can easily corrupt data.
   (B) In general, formally publishing a book requires significantly more effort.
   (C) It is often difficult to determine the source of data available via the Internet.
   (D) Once a book has been published its contents cannot be altered.

2. Surveys and interviews are used to:
   (A) collect data from secondary sources.
   (B) collect data directly from people.
   (C) collect data from all members of a population.
   (D) collect data from a random sample of the population.

3. A census can be best described as:
   (A) a survey that is completed by a random sample of the population.
   (B) a survey conducted every four years by the Australian Bureau of Statistics.
   (C) a survey that is completed by all members of a population.
   (D) a statistical analysis technique that summarises the results of a survey.

4. Selling a database containing personal information on individuals could be allowed if:
   (A) the company selling the database actually collected the data directly from the individuals.
   (B) none of the data is of a sensitive nature.
   (C) the data is not necessary for the purchasing company to carry out its functions.
   (D) the individuals, whose personal information is in the database, have been informed of any organisations who will purchase the data.

5. It is generally accepted practice to include instructions for paper-based forms:
   (A) at the start of the form.
   (B) where they are needed within the form.
   (C) as a separate reference document
   (D) (A) for general instructions and (B) for specific instructions.

6. The amount of space left for answers on paper-based forms:
   (A) should be the same for all questions.
   (B) should reflect the amount of information required.
   (C) can be changed as an individual completes the form.
   (D) should be adjusted to enhance the overall look of the form.

7. Carpal Tunnel Syndrome:
   (A) is another name for RSI.
   (B) is the most common form of RSI experienced by data entry operators.
   (C) is caused by muscles being held in a static position.
   (D) can be easily corrected by improving the design of workstation furniture.

8. When developing surveys many researchers have a theory that they wish to be supported using evidence from the survey. Surveys created for such purposes:
   (A) are likely to be biased if the researcher designs the survey.
   (B) should be designed by individuals who do not have an expectation that one outcome is more likely than another.
   (C) should collect data from a random sample of the population or from the entire population.
   (D) All of the above.

9. Which of the following contains only positive interviewer characteristics?
   (A) Remembering the last person interviewed more positively, letting the candidate direct the interview.
   (B) Well-prepared questions, talking too much, putting the subject at ease.
   (C) Careful listening, politeness and generosity, focusing on the topics to be covered.
   (D) Personal warmth and engaging manner, not allowing enough time for the interview.

10. The main aim of adjusting furniture and equipment correctly is to:
    (A) reduce the amount of stress experienced by users.
    (B) ensure all muscles are maintained in a relaxed state.
    (C) reduce the number of repetitive movements performed by users.
    (D) increase the amount of time users can spend at the keyboard.

11. The design of paper forms for data collection shares many aspects common to the design of user interfaces, however there are significant differences. Describe differences in the way paper forms should be designed compared to computerised user interface forms.

12. Imagine you are conducting interviews to fill a position for a data entry operator. Devise a list of questions suitable for such an interview.

13. Explain how the symptoms of carpal tunnel syndrome result from repetitive overuse of the fingers.

14. Most libraries maintain a computerised catalogue of each of their resources, however the resources themselves are not held in digital form.
    (a) Explain why most libraries do not digitise all their resources.
    (b) Describe how data held in non-computer based library resources can be located.

15. List and describe any social and ethical issues apparent in each of the following scenarios:
    (a) A researcher is conducting a survey to determine the current population distribution of an endangered species of bird. The researcher sends out a survey form to each landholder within the region in which the bird has previously been encountered. The landowners are requested to note the number of individual birds of the species they encounter, together with other details in regard to each sighting.
    (b) A credit card company sends out letters offering to increase the credit limit for a selected number of their cardholders. The cardholders who are offered the increased credit are selected based on their income, past purchasing history and poor payment history; these are the most profitable customers for the credit card company.
    (c) Mary works for a telephone sales company. She is required to work 10 hour shifts, after every 2 hours she is scheduled a 10 minute rest break. Mary's job entails making phone calls and recording the result of each call into a database, she is only permitted to talk to her supervisor during each shift.

CHAPTER 3 REVIEW

1. Hardware devices for collecting image data include:
   (A) scanners, digital cameras and camcorders.
   (B) keyboard, mouse and barcode readers.
   (C) barcode readers, microphones and digital camcorders.
   (D) CCDs, pressure switches and USB ports.

2. LEDs are used to assist the collecting process in many:
   (A) keyboards, mice and scanners.
   (B) mice, barcode scanners and flatbed scanners.
   (C) digital still cameras, camcorders and web cameras.
   (D) analog to digital conversion processes.

3. Microphones collect audio data and organise it into:
   (A) analog sound samples.
   (B) digital sound samples.
   (C) digital electrical energy.
   (D) analog electrical energy.

4. CCDs output:
   (A) digital electrical energy.
   (B) analog electrical energy.
   (C) digital light.
   (D) analog light.

5. Flatbed scanners do not require aperture and shutter speed controls because:
   (A) they use a Bayer filter to control the light collected, hence modification of the light is simply not needed.
   (B) images collected using flatbed scanners do not emit light.
   (C) only a single image is being collected at a time so it is not necessary to change these settings.
   (D) flatbed scanners produce their own light source and the image is always a constant distance from the CCD.

6. Application software:
   (A) provides the interface between hardware devices and software.
   (B) performs a specific set of tasks to solve specific types of problems.
   (C) is used to control and coordinate the functions of a computer system.
   (D) is used to manage the hardware and software resources of the system.

7. Users interact with computer systems via:
   (A) collection devices.
   (B) the user interface.
   (C) the keyboard.
   (D) application software.

8. Collecting data from Internet users via a web page:
   (A) requires that each user's computer must be a web server.
   (B) requires a data entry form to be stored on or created by a web server.
   (C) means that each user must install the appropriate data entry software on their machine.
   (D) is not possible as web pages are only able to display information.

9. Data integrity is a measure of:
   (A) the accuracy of the data.
   (B) the validity of the data.
   (C) the ability of the system to update its data.
   (D) how often the data needs to be analysed for errors.

10. Health concerns for participants entering large volumes of data include:
    (A) vision problems, general muscle strain and RSI.
    (B) work routine and design and adjustment of equipment.
    (C) lack of job satisfaction leading to various workplace injuries.
    (D) privacy, copyright and security issues.

11. You have been assigned the task of collecting an image of each student attending your school for inclusion in the school's database. Describe suitable collection hardware and software necessary to achieve this task.

12. Text data is commonly collected using the keyboard, however it can also be collected using voice recognition and optical character recognition (OCR). Research and describe the hardware and software needed to collect text using voice recognition and OCR.

13. Draw a diagram to illustrate the essential features of an ergonomically sound workstation.

14. List and describe possible reasons for inaccuracies in data as a consequence of the collecting process.

15. 'The collecting information process does not operate in isolation. Many of the other syllabus information processes must occur as part of the collection of data'.

    Do you agree? Justify your answer using at least three specific examples.