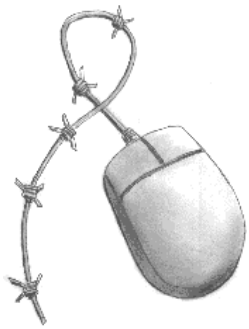
Participating With Safety Briefing no.2

Backing Up Your Data



Written by Paul Mobbs for the
Association for Progressive Communications, March 2002.

This briefing is one of a series on information security. It looks at:

- Why, when and how you should back up data and information on your computer
- Storage, security, costs and legality
- Organising your information
- Methods of backing up
- Issues to consider: longevity, security, recovery and redundancy
- Installation disks
- Backing up with Linux

Why 'back up'?

Information on your computer is vulnerable: hard disks can fail, computer systems can fail, viruses can wipe a disk, careless operators can delete files, and very careless operators can delete whole areas of the hard disks by mistake. Computers can also be damaged or stolen. For these reasons backing-up your data is essential. This involves making copies of essential files on your system and keeping them on another computer, or on some form of storage media.

To ensure that you can back up easily it is important to organise the information on your computer. The aim is to make different users, and/or different areas of an individual user's work, easily identifiable and easy to find on the disk. This is achieved by setting up a series of directories on the disk that contain different types or areas of a user's work. This is generally called a 'workspace' - the area of a disk that contains a user's work. This also helps segregate a user's unique work from files or information that are held by a number of people, or that are loaded onto the computer from other sources, such as CDs, so you don't back-up files unnecessarily.

When to back up

Backing up data can be a difficult task where there are lots of files to copy. The first rule of backing up is therefore to minimise the amount of data you have to back up by regularly removing useless or out-of-date files from your system. It's also important to get into the habit of backing up in different ways for different reasons to increase the reliability of your backed-up data.

You should consider backing-up:

- When you've done a large amount of work over a short period - in which case you should back up

all the contents of your 'workspace';

- When you've completed a major body of work - you should clean up the directory containing the files (to get rid in files that are not needed) and just back up that directory;
- On a regular basis, back up your whole 'workspace' and the essential system files.

How to back up

How you back up depends upon the type of equipment you have and the amount of data you have to back up:

- Backing up your immediate work is very simple. It is unlikely to be a huge volume, and so some sort of disk storage is easy to organise. Within an organisation, where you are more likely to have a network of computers, it also possible to back up users' work to a single computer somewhere on the network.
- On a regular basis it is a good idea to back up not just the current work, but also copies of the word processor's user dictionaries, email or Internet download directories, and other important system files, as well as all the users' data.
- Make arrangements to hold copies of backed-up information at other locations. These back-ups are not going to be as up-to-date as your regular office-based back-ups. But in the event of a catastrophe, such as the theft of computers, a fire, or a deliberate raid by the state or some other organisation, you can at least put a large part of your computer-based work back together.

Storage

Another issue with regard to backing-up is how data is *stored*. Many people use programs that compress large amounts of data into a smaller space, and which stitch together many small files as one big file in the process. Whilst this is a useful way of backing up data in the short term, in the longer term this has security and reliability implications.

Security

Backing up data has other security implications. Whilst the information on the computer can be protected behind different type of security barriers, data on back-up disks is more vulnerable because it is 'open' - simply stored on some form of storage media without any barriers to access other than the box or room the storage media is stored in.

> *In circumstances where you have particularly sensitive data to back up, it is often safer to have one set of back-ups containing your ordinary data, and a smaller set of back-ups containing more sensitive data for which you make separate storage arrangements.*

Costs

Fundamentally, your policies on backing up are just a matter of money. If a computer holds all the work you have produced over 2 years, and the cost of producing that work is £5 to £10 per hour, the value of the information held on a computer could be £15,000 or more to reproduce. On a computer worth only £1,000, and the cost of disk storage or CD's at £1 to £10 each, backing-up makes clear economic sense, without the impact that losing large amounts of your work can have on your business or campaign activities.

Legality

There are legal implications to backing up. As copyright law becomes more restrictive it affects your legal rights to back up the contents of your hard disk.

Where information has been created by someone else, be that a book or an email, it is technically their property. The backing-up of information, where it is solely for your own use, is a grey area in the law in many countries. But where more than one person has access to data, for example as part of a community organisation or group, it can be argued that making copies of other people's information without their permission is an infringement of copyright. In some countries this will be a criminal offence, whilst in others it will be a civil offence where the copyright owner must prosecute.

The easiest way to steer clear of copyright problems when backing-up is to segregate the information on your system according to whether it is 'open' information, or information that may need restrictions on its use. You can then either not back up this information, or back up to a disk or media that will not be copied or distributed to other persons/organisations.

In general you should try to avoid backing up:

- Any software - be it the installation programs or installed programs - where there is any possibility that a copy may end up in the hands of someone else;
- Any web pages or email where the page/email contains a specific copyright message;
- Any reports, and especially books and multimedia works, unless they have no clear copyright restrictions, or unless they are specifically circulated as 'open content'.

Organising information

Effective backing-up depends upon how well you organise your information.

> *If your computer has lots of unimportant data mixed in with your most important files you risk damaging files by deleting or editing the wrong ones. You will also waste time and money backing up far larger quantities of information than you need to. Organising information in this way will also mean that you use your computer system more effectively and efficiently.*

To ensure good practice in organising your information you should:

- Use directories within the hard drive(s) of your computer to hold data for each user, and different areas of that user's work. In this way you can back up single files, directories that contain an entire project, a user's entire workspace, or the data for all users who use that particular computer.
- Make sure that files shared between many users are kept separate from a user's individual files.

In more detail, this means:

- Always have a directory for each user who uses the computer, and perhaps a 'guest' directory for occasional users;
- Try to keep finished work and work in progress apart - finished work should be backed up for long-term storage, where work in progress should be backed up regularly, and keeping older/completed files out of current work area saves space;
- When starting a new project, always create a directory for it and store all information related to the project in that directory - in this way you can keep a regular back-up of the project by copying the whole directory in one go;

- Where a project contains a large volume of files try subdividing them into more sub-directories to allow backing-up on different disks - this is dependent upon the capacity of the storage media you are using; and
- If using a network for access and backing up, always try to keep the files that everyone shares separate from user's. Get your users into the habit of accessing and updating shared files in one location - this prevents confusion arising over different versions of the same file stored in different areas of the network system.

Organise your backed-up data so that it mirrors the organisation of the work on the computer. This means that in the case of a file or a whole disk being damaged or corrupted, the work can be restored easily; it also means data can be simply copied in a way which re-creates the working environment that your users find familiar.

Methods of backing up

There is no specific right way of backing up. It will depend upon

- what form your data is in,
- how much data you wish to back up, and
- what hardware you use to back up.

There are various options for backing up small and large quantities of data. The critical factors you need to consider are the costs and capacity involved in each option and the period for which data can be held without any degradation or corruption.

Table: Comparisons of capacity and cost

	Low cost	High cost
Low capacity	Floppy disk - 1.44MB E-mail - 1 to 2MB	Secure server - 1 to 25MB
Medium capacity	ZIP disk - 95MB CD recordable (CD-R) - 600MB	JAZZ disk - 900MB CD re-writable (CD-RW) - 550MB
High capacity	QIC tape storage - 500 - 4,000MB	DAT tape - 2,000MB or more DVD recordable (DVD-R) - >4,700MB Removable h/disk - 4,000 to 80,000MB+

All figures are in megabytes - MB

Floppy disks

All PCs come with a floppy disk drive. Until recently the floppy disk was sufficient for backing up data. But over time, programs have become more complex, files have grown larger and the size of hard disks has grown from a few megabytes to a few gigabytes. Using floppy disks to back up the entire contents of your hard disk is no longer a viable option.

Despite their small capacity, floppy disks are still a good way of backing-up small amounts of data - for example backing up your day's work. But floppy disks have become increasingly superfluous with the development of the Internet. Whereas previously people sent floppy disks via the post to move data over long distances, today the same volume of data can be sent as a file attached to an email.

Local networks

Backing up over a local network, to another computer in the same room or building, can be done at high speed and involves the transfer of large volumes of data. Backing up over a network can reduce costs, because you only need buy one high-capacity back-up device, such as a tape drive or CD-burner, which everyone can share. But you still have the problem of the data being stored on machines, and those machines being accessible within your office.

When backing up over a network you are still backing-up to another hard disk. But the statistical likelihood that two computers will have their hard disks fail at the same time is very low. You will be able to recover information from at least one of the hard disks involved.

> *The only problem you would have with a local network would be if all the computers in your office were stolen. For this reason keep any computer you use for backing up more securely than other computers on the network; locked in a ventilated cupboard or purpose built security cage, for example.*

The Internet

A more secure option is to back up over the Internet to a secure server, or to another person with whom you have an arrangement to hold data with. These machines may be in your own country, or more likely, in another country where the laws on privacy and data protection provide far greater protection for your information.

The reason this provides greater security is that it removes the data from your location to somewhere else; even, if necessary, to a different legal jurisdiction that gives better protection to personal information.

There are two options:

- With a secure server you can access the system at any time to store or retrieve data. You should be able to store data in an encrypted format so that only you can access it, but the actual online session should also use some form of encryption so that the transactions themselves are secure. For those states where storing data in an encrypted format is a problem, or where the possession of certain type of information is a problem, secure servers are a simple solution to data security.
- If you have an informal agreement with other activists or groups in another country, you can exchange information via the Internet and they can look after your files for you.

> *The only problem with this is that the transfers are not automated, so you rely on those looking after your data to store it carefully, and return it to you when you request it. In many cases you could do this by default, by sharing your information for other groups or activists to use themselves; in this way your data is more secure because others will hold and be able to use it in the event of your work being restricted or prohibited.*

The issue to consider with backing up over the Internet is the *amount of data* you can transfer. With *dial-up* connections only a few megabytes are realistic. With *broadband* connections, depending on the available upload speed, that can rise to around ten or twenty megabytes.

ZIP/JAZZ drives

ZIP and JAZZ drives are *high-capacity removable* disk drives.

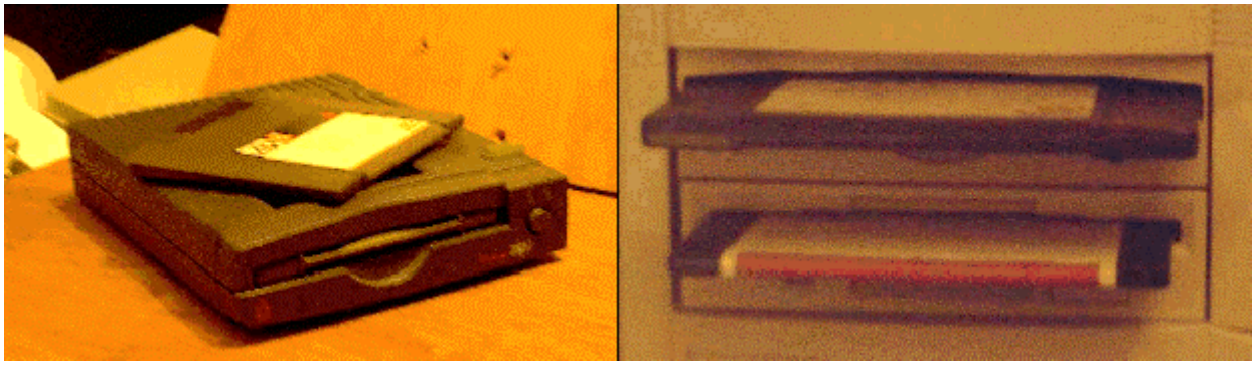
ZIP drives are 100 megabyte disk drives that plug into your parallel port (without affecting your use of the

printer) or can be fitted internally like an additional removable hard drive.

JAZZ drives are similar, but they have a 1 gigabyte capacity.

The only issue between the two is how much data you have to back up, and how safe you can keep it. JAZZ drives are good for backing up data from *devices that generate huge files* - such as digital video. If you lose or damage a JAZZ disk you will lose ten times more information than if you had lost a ZIP drive.

ZIP drives provide a very good *short-term* back-up medium, although their cost and smaller size makes them less good for long-term storage because of the number of disks you might generate.



Left -an external ZIP drive. Right, top -an internal ZIP drive,
with a floppy disk drive below for comparison

QIC Tape/DAT tape drives

Tape drives or DAT tape drives are usually fitted inside the computer, although external units are available. Some versions plug in like another disk drive, whilst others plug into the computer's parallel port or USB port. Tape drives use a long tape in a cartridge to copy the whole contents of a hard disk onto tape. There are two types:

- QIC ('quarter inch cartridge') drives, which work in a similar way to audio cassette players. QIC tapes vary in capacity from 40 megabytes to 4 gigabytes.
- DAT ('digital audio tape') drives, which work in a similar way to video recorders. DAT tapes can store 2 gigabytes or more.

With tapes you usually back up an entire hard drive rather than parts of it. You can then restore all or parts of the disk at a later date. They are a *cheap* solution, given their capacity, but they are not very reliable after a long period of use. They are also not as convenient as other options such as CDs because you must have a tape drive to use them. They are also slower, and on standard PCs require an additional interface card because they usually use the SCSI drive interface standard.

Writable/re-writable CD ROM drives

Writable and re-writable CD ROM drives can store up to 550 or 650 megabytes per CD (they may say 600 or 700 megabytes, but 50 to 100 megabytes are used for directory information).
Writable drives (CD-R = 'CD recordable') are good for making *permanent archive back-ups*, but it's a bit of a waste because you can only write once to a CD; you cannot over-write it after that. This can also present a security problem because as you progressively update your backed-up data you'll probably want to dispose of older CDs. Writable CD ROMs, are now the cheapest form of backing up, with the recent fall in the price of drives.

Re-writable CDs (CD-RW = 'CD-read/write') are better, because you can add and overwrite files as you go; they're better than CD-Rs for small-scale backing-up of recently completed work. But the more times that you re-write a CD the more likely that data may be corrupted. This is because the re-writing process slowly degrades the recording polymer inside the disk.

Re-writable CDs, are a competitive alternative to tape drives because of their better reliability and versatility. But tape drives still win on capacity.

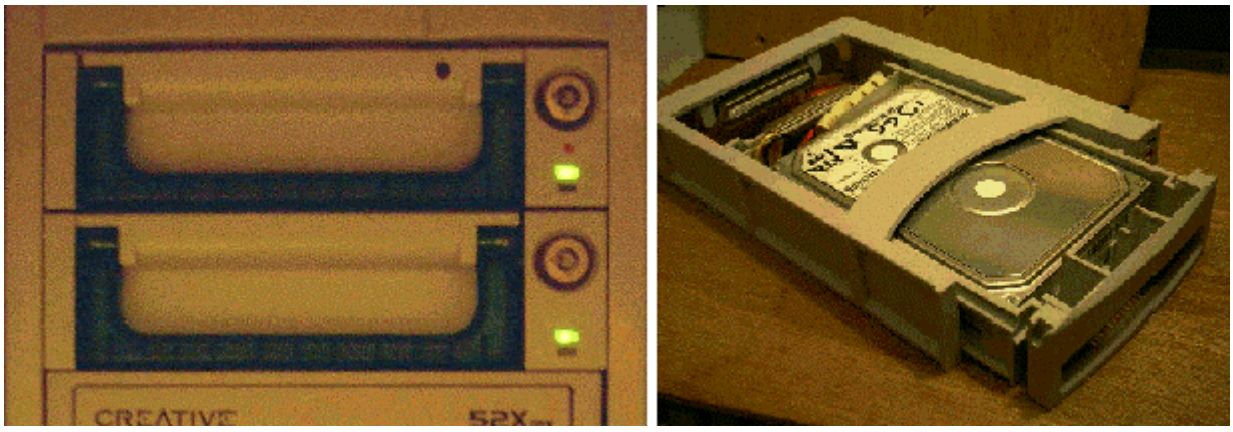
Writable DVD drives

Recently writable DVD drives (DVD-R) with a capacity in excess of 4 gigabytes of data are now available. But they are extremely expensive to buy and run, although costs will fall over the next few years. For the foreseeable future they will be outside the scope of many small organisations. The only exception may be those involved in digital media, for whom writable DVD not only represents a back-up option, but also a means of distributing multimedia productions.

Removable hard drives

Removable hard drives can hold tens of gigabytes of data and are very effective in terms of security. You need more technical expertise to manage them than other options, however.

When you have finished using your computer you can remove the hard disk in a protective caddy and lock it away. You could also back up your data to the hard disk and then store it in another location. This method is very space-effective, because a large amount of data can be stored on the disk. The metal shielding of the hard drive's case and the small size also make it more secure to transport and store.



Left -Removable hard drives, with key locks, installed in a computer.
Right -a hard disk rack and caddy, with the hard disk installed, before installation in a computer

The real benefit of removable hard drives is if you have a local network; you can have one dedicated machine that backs up to a removable hard drive, back up from all other machines to this computer, and then remove and store the hard disk elsewhere.

A hard drive can contain a working operating system, not just data. You can therefore set up a hard disk in the computer with a basic operating system, and then be able move it to another computer in an emergency. Another advantage is that you could encrypt the entire hard drive for additional security.

A comparison of the performance of different backing-up options

	Longevity	Security	Recovery	Redundancy
Floppy disk	A few years, if stored in shielded container. Very good for small daily back-ups of current work.	Need good physical security, or encryption of data	Good, if drive is aligned. In worst case, data can still be recovered from parts of the disk even when corrupted.	Easy to create copies at low cost
Email	Relies on recipient's information security procedures	Low security in transit unless encrypted - relies on good storage security by recipient(s)	Relies on recipient to send it back.	Easy to email many recipients at once. You might have complications over the version/date of the backup.
Secure server	Relies on server operator's information security procedures	Low security in transit unless encrypted - relies on good storage security by server operators	Good if server is online for the majority of the time, but it is still reliant on you having Internet access to recover data	Not good as sole means of backing-up sensitive data, but very good as a means of quickly backing-up off site
ZIP disk	A few years, if stored in a shielded container	Need good physical security, or encryption of data - ideally should sign if stored off site.	Good, if drive is aligned. But overall problem that not everyone has a ZIP drive	Easy to create copies, although it can be time consuming and expensive.
JAZZ disk	A few years, if stored in a shielded container	Need good physical security or encryption of data - ideally should sign if stored off-site.	Good, if drive is aligned. Greater problem that JAZZ drives are rarely used	Easy to create copies, although it can be time consuming. And expensive
CD recordable	Perhaps ten years or more, if stored to reduce damage to disk	Need good physical security, or encryption of data, but compact way of storing data off-site	Good, but you must verify reading after creating the CD on an ordinary CD-ROM drive	Easy to create copies at low cost
CD re-writable	Limited by regularity of use; each re-write degrades the storage media. Useful for daily/weekly back-ups.	Need good physical security, or encryption of data - ideally should sign if stored off-site, but this is difficult on such a large disk.	Good at first, but may become difficult with regular re-writing.	Not easy to copy because they can only be read from a CD-RW drive, unlike CDs. OK for historical backing-up
DVD recordable	Perhaps ten years or more, but the media has not had long term real-world testing	Need good physical security, or encryption of data, but compact way of storing data off-site	Good, but you must verify reading after creating the DVD on an ordinary DVD drive	Easy to make copies, but a very expensive option. Good for huge quantities of data - e.g. digital video
QIC tape	A few hundred back-ups before wear may become a problem	Need good physical security - difficult to encrypt/sign. Inclusion of sensitive files such as encryption keys makes them a security problem.	Good, but can be slow. Also not widely used outside of the business world.	Not an issue. Made for single back-ups of a whole hard disk - making copies is time consuming
DAT tape	A few hundred back-ups before wear may become a problem	Need good physical security - difficult to encrypt/sign. Inclusion of sensitive files such as encryption keys makes them a security problem.	Good - main benefit is faster transfer rate than QIC. Problem is DAT drives are not common because of the higher cost	Not an issue. Made for single back-ups of a whole hard disk - making copies is time consuming
Removable hard disk	Perhaps ten years or more if stored correctly	Good physical security required. Encryption of hard disk simple to arrange	Good, and at high speed. It is also highly portable between computers - it is a drive.	Expensive option, especially for large quantities of data. But good for infrequent

A removable hard drive thus provides you with a really good disaster recovery system; in the event of a catastrophe you simply need another compatible computer, plug in the back-up hard drive, and all your data and applications will be available as before.

> *Using a removable hard drive with the Windows operating system is not really feasible; it is not very portable between machines, and usually requires you to provide new drivers to work with different hardware. The new Windows XP system would not work at all if you had to install the drive in a different computer system. With Windows your best option is a 'pluggable' hard drive - this plugs into a USB port and can be move from computer to computer provided each computer has the correct software and drivers installed.*

The Linux operating system, on the other hand, lends itself well to using a removable hard drive. It is very portable, and even on computers with very different hardware it will reconfigure itself and carry on working as before.

Longevity, security, recovery and redundancy

When backing up data you are trusting that the system you are using will return that data to you when you require it to. There are four important issues you must consider to ensure that your data remains available for when you need it most:

- **Longevity** - how long the data remains viable;
- **Security** - protecting the data from damage or theft;
- **Recovery** - being able to read your back-ups; and
- **Redundancy** - making sure that you have enough copies, should something nasty happen to one of them.

Longevity

Longevity is an important issue if you want to keep data for a number of years. The behaviour of materials you use for back-ups is critical; you will have a lot of data stored in a very little space, and degradation of the materials can destroy your entire back-up.

All magnetic media - floppy disks, ZIP/JAZZ disks, tapes and hard disks -are vulnerable to damage by magnetic fields. This includes:

- Strong sources of electromagnetism, such as your computer's monitor, telephones and loudspeakers;
- The Earth's magnetic field.

A floppy disk left unshielded will have its contents corrupted after a few years because it becomes damaged by variations in the Earth's natural magnetic field, and by proximity to magnetic fields from electrical equipment.

In order to ensure the long life of data stored on magnetic media you should regularly *'refresh'* the information on the disks. Information stored on *ZIP or floppy disks* should be copied back onto a computer every year or two. You should then reformat the disk, and write the data back to the disk again.

> *Keep all your magnetic media inside metal storage cabinets, as far away from electrical sources as possible. If you intend to store the disk for long periods of time but you do not have a suitable metal case, put a layer of metal (metal foil, for example) around the disks to significantly reduce the influence of magnetic fields.*

With *magnetic tapes/DAT tapes* refreshing is not an issue because you tend not to use them for file storage, but rather backing up a whole hard disk. Also, the roll of magnetic tape shields itself from local magnetic fields (although stronger fields will still degrade its content).

> *Tapes that are to be stored for long periods should be wound every year or so by putting them back in the drive and doing a test read. This prevents the surfaces of the tape sticking together.*

Read-only CDs, as well as *writable CDs* such as CD-R, CD-RW and DVD-R, are not susceptible to magnetic fields.

> *CDs are sensitive to light, however, particularly the ultraviolet component in sunlight, because it degrades the polymers/plastics in the disk.*
>
> *All writable CD media are also sensitive to heat. Heat can degrade the film in the CD that the data is stored in.*
>
> *CDs should therefore be stored in light-proof, strong containers, and they should be protected from extremes of temperature, as well as regular changes in temperature.*
>
> *All back-up media should be stored in conditions that are free from damp, at a fairly constant temperature. Regular swings of temperature to extremes can damage the plastics or polymers in them. Swings in surrounding temperature can also cause damage through condensation, as warm humid air condenses on the cold surfaces of the storage media.*

The storage area should be free from vibration because this will cause mechanical stress on the polymers. It's also important to protect hard disks, if you use them for backing up data, from static electricity; this is most easily done by storing them in anti-static bags.

Security

Backed-up data can be a significant liability; in the event of a break-in or raid, back-ups provide a very portable means of taking copies of your data. The only way to protect backed-up data is through physical barriers. You can encrypt your back-ups, but this has implications for recovery of the data (see *Recovery* below).

To prevent anyone removing your back-ups:

- Keep them in secure containers, which are themselves securely installed/fixed to prevent them being carried away;
- Segregate your back-ups according to the sensitivity or importance of data. In this way you can keep your most sensitive data under better security than the rest;
- Consider setting up 'decoy' storage areas by putting your less important data or old back-ups in clear view, whilst finding more secretive/secure places for more important data.

The *most secure option*, for data and for important paper-based records, is to *keep copies in another location*. Although off-site back-ups will not be as up-to-date as those kept in your office, they will survive any attempt to deprive you of your computers and data, especially by the state or other organisations trying

to stop your work.

With the actions of the state especially, the confiscation of computers and data effectively stops your work. The most effective way for other organisations to stop groups or individuals working is to burn down their offices (a tactic seen in states such as the USA). Providing that you have off-site back-ups, and you can get access to computers, you can carry on your work fairly soon after any catastrophic loss of data or equipment.

Creating a record or signature

The simplest form of backing-up record is a directory listing. The simplest way to create this is to open an MS-DOS prompt window and then change to the directory you require, remembering that any directory names with spaces in need to be surrounded by quotes. For example you could change to the D: drive and then the directory called 'data back-up' using the following two commands -

```
d:                    (then press return)
cd "data back-up"     (then press return)
```

Then, to create the directory listing as a file, assuming you wanted to store the file in your C: drive, you enter the command -

```
dir > c:\list.txt     (then press return)
```

This command creates a directory listing and then directs the list to a file (list.txt or whatever you call it) rather than to your DOS window. The important details this file contains are the files size, in bytes. A different number would indicate a change in the file.

Creating a checksum is difficult on Microsoft systems because a checksum utility is not supplied as standard. Some anti-virus programs have an ability to create checksums of files, and some PC security programs also create checksums. When you generate a checksum file you not only get a file size, you also get a number that is uniquely created according to the contents of the file.

The simplest option on Windows systems, and the most secure, is to create a signature of the file using PGP or other encryption programs that have a file signature function - such as *PGP Free*. To create a file signature you go through the following steps.

1. Run your PGP/encryption program that has a signature function.
2. Select the 'create file signature' option (or however it is described) from the menu.
3. You then select the file you wish to sign.
4. You then enter your key password. You are also usually given the option at this stage to 'detach' the signature and store it as a separate file. Unless you are able to detach the signature the file itself is signed, and this is likely interfere with the reading or interpretation of the file.

No matter what type of record you create, with the exception of signatures which are more secure, you must store the records separately from the back-up. For signatures, you must store copies of your key pair safely, and find a way of remembering the correct password, in order to verify the signatures of backed-up files.

Verification is more of a priority on media that can be *edited*, such as magnetic media and CD-RW disks. Read-only disks cannot be edited (although you could have a duplicate substituted for your own copy).

A final security feature you should consider installing on your back-ups is some form of *signature or verification* of the back-up's integrity. If the back-ups are not encrypted, have no strong physical security, or are stored on a computer outside your direct control, verification helps you ensure that your back-up has not been tampered with. This does not prevent the back-up being read. It just ensures that no one can change the content and introduce erroneous data or a computer virus.

Verification can be as simple as having a directory listing, that you keep on a different disk, in order to verify the information stored about the file - in particular the file size.

The most secure way of doing this is keeping some sort of *checksum* or *digital signature*:

- A checksum is a simple numerical analysis of the file(s') contents. There are programs that generate checksums, but they are not supplied with Windows as standard. Some virus checkers will also generate checksums for you, stored as a separate file.
- A digital signature is an analysis of the file(s') content which is protected with a cipher.

Checksums can be forged. Digital signatures cannot be forged unless the people modifying the data have the key for the cipher.

Digital signatures can be produced by most PGP encryption programs; the file can be signed and the signature added to the file, or the signature can be detached and stored separately.

The signature or checksum should be kept separately from the data so that it can be modified itself (although a signature cannot be falsified, it can be corrupted to cast doubt on the authenticity of a back-up). By running the signature or checksum against the back-up you can check not only for any corruption of the data, but also, if the risk exists, for whether any of the data on the backup has been tampered with.

Recovery

You must be able to recover your data. You should also always plan for data recovery *not* taking place on the computer on which the back-up was made. This has implications for the way that you make your back-ups.

You should always verify your back-up immediately after creating it. Some CDs, for example, can contain errors created as part of the CD burning process. They may not be detected by the CD-RW drive but may prevent the reading of information on an ordinary CD-ROM drive. By reading a newly created back-up in an ordinary drive you ensure that the CD has been correctly created. Problems can also occur with floppy disks because the heads of the disk drive are misaligned. Whilst the floppy disk will work well on the machine it was created from, other floppy drives may be unable to read the whole disk.

The format in which you store data can be very important - file formats become outdated and unreadable, and some forms of data storage are more vulnerable to corruption than others.

If using *proprietary file formats* (for example word processing files) you should make sure that you are using a widely compatible format. The most recent format of a particular application may make compatibility with other people's systems more difficult. Likewise, using less well-used formats (a particular problem for Linux users) may make compatibility a problem. If you work with a wider community of officers or computer users, you need to arrive at an agreement on what file formats you will use to ensure that data is available to the widest possible number of users.

Data compression, using programs such as PKZip, GNUZIP, etc., is a very useful way of squeezing a lot of data into a smaller space.

> *There are also high risks involved in using compression:*
>
> *If you have one small error in the stored information, you tend to lose all or large portions of the compressed data. This is because the data is processed as a long stream of continuous data, and an error in reading causes the decompression process to fail; and*
>
> *With compression you are usually compressing more than one file, and whereas an error in the storage media might only cause the loss of one file, with compressed files one error can lose all the files contained in the compressed file.*

> *Encrypting backed-up data can be very risky. Whilst data encryption represents a very secure way of holding data, it is risky because you always need the following to decrypt your data:*
>
> *A small file which contains the encryption key (the 'secret' or 'private' key); and/or*
>
> *A password or passphrase.*
>
> *For security purposes, always keep your encrypted data, the private key and the password/passphrase separate. If not, the security encryption gives is diminished. If you lose the password or secret key you will never be able to recover the data.*

Like compression, *encryption* works on a continuous stream of data. Some encryption systems will also initially compress data in order to reduce the size of the created file. Errors in storage or reading from the storage media will prevent you recovering anything from the encrypted file.

> *Never, therefore, rely on an encrypted file as your master back-up. Encrypted back-ups are a way of protecting against poor physical security, but you should always try to keep a plain back-up in a more secure location.*

Redundancy

The laws of probability dictate that, at some time, your backed-up data will fail. You can minimise the likelihood of this happening, by adopting the 'good practice' tips outlined above. But you must always assume that at sometime your system of backing up will go wrong. To cope with this, keep 'redundant' copies of backed-up data, to be used in the event of a problem with your latest back-up.

There are two ways of keeping redundant back-ups:

- *Historical back-ups* - this involves keeping a regular back-up of information, but not deleting older back-ups. This is very easy to do with CD-R disks - you just keep backing up to cheap, write-once disks, and keep the older disks in case more recent versions are damaged.
- *Duplicate back-ups* - this involves copying your back-ups, and storing them in case your primary back-up fails. This means you have to be able to copy your back-up media, or you must back-up twice.

Duplicate back-ups have the advantage of always being more up-to-date than historical back-ups.

If you use CD-Rs to back up, you will effectively make historical back-ups because you will always have the older disks left over from previous back-ups. After a while you will need to carefully dispose of them; or you could use them as off-site back-ups.

> *Duplicate back-ups take more effort to create. There is also a problem of inheriting errors in the original file. With duplicate back-ups, any error that crops up on your system (either an error with the file or something like a file virus) will always exist in your backed-up data. With historical back-ups you always have the option of going back to find an uncorrupted version of a file.*

At other times, for example when creating up-to-date copies for storing off-site, you will need to make copies of back-ups, or back up twice.

In practice it really doesn't matter which option you use - you can choose either, or use both together.

> *You must evolve your own procedures for backing-up that fit the technological capabilities of your system and your own needs. It doesn't matter how you do it, as long as you do it regularly, the information is stored securely, and the resultant information is usable should you ever need to retrieve it.*

Installation disks

So much of backing up relates to the data created by your use of computers. One issue is seldom considered - backing up your system or installation disks.

When you set up your system you install software from floppy disks or CD ROMs. These are important for three reasons:

- If you use or treat them incorrectly, they won't function, and you won't be able to install your software;
- If you lose them, you'll possibly have to replace them, or upgrade, which could involve great expense; and
- If you use them incorrectly you could infect your installation disks with a virus - which then means you may never get rid of it because you will infect your system each time you install those applications.

> *CD ROMs (unless they are copies of software on re-writable CDs) are immune from virus attack because they can't be written to. However, all magnetic media - for example the floppy 'boot' disk that comes with operating systems, can be infected with viruses. Software that's downloaded from the web as an executable file can also become infected with viruses.*

In practice your installation disks are not very vulnerable to casual theft, since the unique registration of the software makes them traceable. But if someone wanted to disable the work of an activist or campaigns group they would not only seek to disable their equipment - they would also take the software disks too. Taking or damaging the software disks makes it far harder for you to start over again because you have to buy new software. For activists, who are often using older systems and old software, this can also mean having to pay for new, more expensive software, because copies of the older systems they were using are difficult to get hold of.

Increasingly, people are downloading software, either as whole programs, or as upgrades or patches for existing systems. All such files should be backed up as soon as you have downloaded them.

The major problem with backing up most software on CD-ROM is the copy protection systems on recent applications and operating systems. Older software, and software based around the Linux operating system does not restrict the creation of back-up copies. But proprietary systems, such as Windows, do not permit copying. There are two options:

Backing-up with Linux

Much of the discussion on backing-up above is also relevant to Linux. There are various additional options when backing-up with Linux, but the major problem will be the compatibility of hardware. Many of the newer DAT drives and CD burners are not compatible with Linux, unless drivers are specifically included for use with Linux systems.

In general backing-up with Linux is more flexible. There are many graphical utilities for backing-up data. But learning how to use the console-based utilities can be very helpful as a means of being able to work with Linux under different distributions (different Linux distributions and graphical interfaces change, but the console commands stay constant).

If using removable hard disks, you'll find that Linux is more portable is included on the disk as an operating system. It is also far easier to set-up hard disk encryption using recent distributions of Linux than it is with windows systems.

It's very easy to create checksums with Linux - this is a good quick and easy method of verifying integrity, and is less hassle than individually having to sign files. There are two commands - `sum` and `cksum` - available to do this. For example, from the command line/console window if you enter `sum * > file.txt` you create a checksum for all the files in a directory and store it as a file. Many Linux distributions install `gpg`, Gnu Privacy Guard, as standard. This means file encryption and signing are easily available.

Another useful command is `dd`, which copies and converts file. But it can also be used to make images of disks in order to create exact copies of a disk as a 'disk image' - the content of a disk stored as one large file. You can then copy an image back to another disk to create an identical copy, or store them on CD-ROMs. All you must do is specify the path of the CD or floppy drive involved, and the path of the directory.a filename to store the image. For example -

`dd if=/floppy.path of=image.filepath` - creates an image of a floppy disk
`dd if=/cdrom.path of=image.filepath` - creates an image of a CD-ROM
`dd if=image.filepath of=/floppy.path` - creates a floppy from a floppy image

Some CD-ROM burners actually provide the option of burning a CD from an image. This is an efficient of making copies of CDs if you only have one CD drive - the CD-RW burner - installed (many mass produced computers only have one CD-RW drive installed to prevent the copying of CDs from a CD-ROM to CD-RW drive).

Another useful command is `tar`. This creates a single 'tarball' (hence tar) archive file from many smaller files and directories. But it does not compress the file, making it less problematic if corrupted. Tar is used especially for backing-up to tape drives. For example:

`tar -cvf /tape.drive.path /home` - backs up all the user's home directories to tape
`tar -cvf backup.tar /data` - backs up the 'data' directory to a single file called 'backup.tar'
`tar -xvf tarfile.tar` - restores the files/directories in the file 'tarball.tar' to the current directory

Tar has many switches and options enabling quite complex handling of data before creating/extracting an archive. For details you should see the tar manual page (open a console window, or go to the command line, and type `man tar`)

- You can try and circumvent the copy protection systems to make your back-up, for example by making an 'image' of the CD-ROM and writing the image to another CD, if you have the software to do this. Not only may this not work with some software, however, but increasingly this type of action is becoming illegal under new copyright laws;
- You could buy another copy of the application, from a computer fair or shop, without a licence certificate; this technically makes it worthless, however, because you can only obtain a copy of a disk with a licence. You keep this copy so that, if you ever need a back-up, you can use this disk with the licence you obtained with your primary copy.

The law surrounding the making of back-up copies is vague. The practice for many years was to ignore the taking of back-up copies ignored provided that the copy was only used when the primary copy was corrupted, and that subsequent usage of the back-up was in accordance with the license for the software. But recent amendments to copyright legislation on many countries now not only make the copying of disks illegal, they also prohibit the use of techniques to circumvent the copy protection systems installed on the disk. This raises the question as to how the public are supposed to protect their (often very expensive) software from damage or corruption. Currently the balance has been tipped entirely in favour of software producers.