

A Guide to Document Management

From



Compact Data Management Ltd
6 Leons Way
Tollgate
Beaconside
Stafford
ST16 3HS

Tel: 01785 220846
Fax: 01785 220876
www.compact.uk.com

Introduction

The introduction of the desktop PC into the office was predicted to reduce our dependence on paper. However, like the prediction made in 1900 that London would be under 6 feet of horse dung by the year 2000, it has not happened. For companies involved in providing document management solutions both these are a blessing.

Most companies produce an ever increasing volume of paperwork and this brings major headaches. The disadvantages to storing information on paper include:

- Significant space required for storage
- Paper is easily damaged
- Documents can easily be lost or misfiled
- The required paperwork is often not readily accessible
- Documents cannot easily be filed by more than one criterion
- Valuable time can be spent on collating documents
- It can be difficult to share information with others
- Relying on single source data can make disaster recovery very difficult

Electronic document management may be the answer for you and your company but careful consideration needs to be made of how you want to access information.

This guide sets out some of the main advantages, and disadvantages to storing data electronically. Some of the various systems available are also discussed together with their relative merits.

The Basic Principle

Most systems involve taking the hard copy originals and scanning them to produce electronic images of the pages. The most common format for these images is Tagged Image Format Files (TIFF). These may be single or multipage images.

If you have the time, this can be done using a standard flat-bed scanner but commercially it is done using high-speed rotary scanners. These are similar to a photocopier with an automatic document feeder and can scan up to 200 pages per minute.

The disadvantage of these machines is their cost. A typical rotary scanner can cost £18,000+ and so is not the sort of machine that most companies can afford to purchase unless they have a significant scanning requirement.

Once the documents have been scanned, they need to be indexed in some way so that they can be found easily. This is similar to filing documents in alphabetical order. Each

image is viewed and information such as invoice number or supplier name is read and entered into a database.

This can either be carried out manually or automatically by using barcodes or Optical Character Recognition (OCR). OCR is the process by which each character is examined by the computer and converted into a computer readable character code.

Once a database has been created, documents can be easily found and sorted by any of the keywords in the database. For example, invoices might be indexed by invoice number, date and customer name. Here is the first advantage of electronic storage: documents are now filed in several different ways and are therefore more easily found than in their original form.

An alternative to creating a database is simply to re-name the images by the title of the document. An example might be for sales invoices where each image is given a filename corresponding to the invoice number. Obviously this has the disadvantage of providing only one means of finding a particular document.

After indexing and scanning, the images and their associated database, if there is one, are written to CD, DVD or similar medium to be returned to the customer. Typically a CD will hold 16,000 images. That's the equivalent of two, four-drawer filing cabinets.

There are a host of different image viewing software packages and database formats available on the market.

TIFF versus PDF

Many people are aware of the Adobe PDF format for electronic documents. Essentially, PDF's are a widely used format that can be read by anybody who has Adobe Acrobat Reader. This small piece of software is available free of charge from Adobe and PDF's are frequently used for distributing documents on the internet.

A disadvantage of PDF's is that they may be slightly larger in size than the equivalent TIFF file but it is a widely accepted format.

Systems are available for the indexing of PDF files and OCR can be used to create "Image and Hidden Text" format PDF's. In these, the text of a document is not only viewable but it can be edited as well. Text can even be copied and pasted into other formats such as Microsoft Word. Obviously this format should not be used for those documents that must remain exact copies of the paper originals.

PDF's are ideal for those people who want to publish information on the web.

ERM/COLD Files

Computer **O**utput to **L**aser **D**isc or **E**lectronic **R**eports **M**anagement – is technology that dramatically decreases the number of paper printouts that a company might normally produce. A COLD system intercepts the stream of data that is normally sent by a host computer to a printer. This information is intelligently indexed and earlier systems were designed to store the data on optical disk rather than on paper. With the reduction in costs for hard disk storage, files now tend to be stored on a central server. COLD technology revolutionizes the way companies view report documents.

Compact can take print spool files generated by mainframe or other systems and convert these into facsimile documents that can then be retrieved using our standard system.

Advantages

- A dramatic reduction in cost associated with the creation of microfiche and printed reports.
- Reduces repetitive and time-consuming tasks of administrative personnel, enhancing customer service by fulfilling requests for information instantly.
- Instead of taking up large amounts of space storing microfiche or paper copies, all the data is kept online and near line on high capacity storage.

Choosing A System To Suit Your Needs

As mentioned earlier, there are a large number of solutions on the market but most follow one of two patterns.

Either, images are linked to a search database of some form and viewing software allows you to find documents by using a number of keyword search fields.

Or, images are given specific filenames that relate to a document and are then stored in a directory-tree format. An example might be a series of directories relating to month in which the files are stored by invoice number.

Obviously, the database system offers greater flexibility and ease of search as the number of search fields is virtually limitless.

Most viewing software performs the same functions; to allow you to see the document image on screen, to be able to print a copy and to be able to email or fax copies to other people.

Systems may be stand-alone with images stored on CD, stand-alone with images transferred to hard disk or shared over a network. Systems are now available where images are stored on a remote webserver and retrieved via the internet.

Costs for software can vary vastly but we can offer software for free thereby ensuring that there is no capital cost involved in setting up a system.

Search systems are available for documents in PDF format. These may be either as manually indexed fields, data-merged fields or full text search using OCR.

It is possible to install your own in-house system but the costs involved for hardware, software licensing and manpower make these suitable only for those companies with very high volumes of scanning. Many companies have installed systems only to find a short time later that they have needed assistance from an outside bureau.

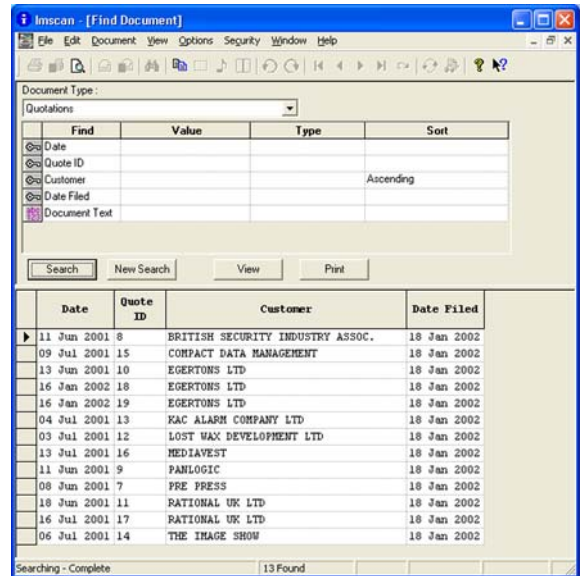
[Compact](#) can provide you with advice on developing an in-house system if required.

IMSCAN Document Retrieval System

IMSCAN looks after your business documents. It manages the storage and retrieval of document images to provide fast, detailed access to the business information you want to see.

All IMSCAN Documents are held in TIFF format.

This versatile document retrieval software will be supplied to you free of charge. The system allows you to find documents quickly and easily after which they can be viewed, printed, faxed and e-mailed direct from the system by way of a simple, intuitive retrieval screen accessible from anyone's desktop.



Whilst images can be easily viewed, they cannot be tampered with so you can be sure that your important company information remains intact.

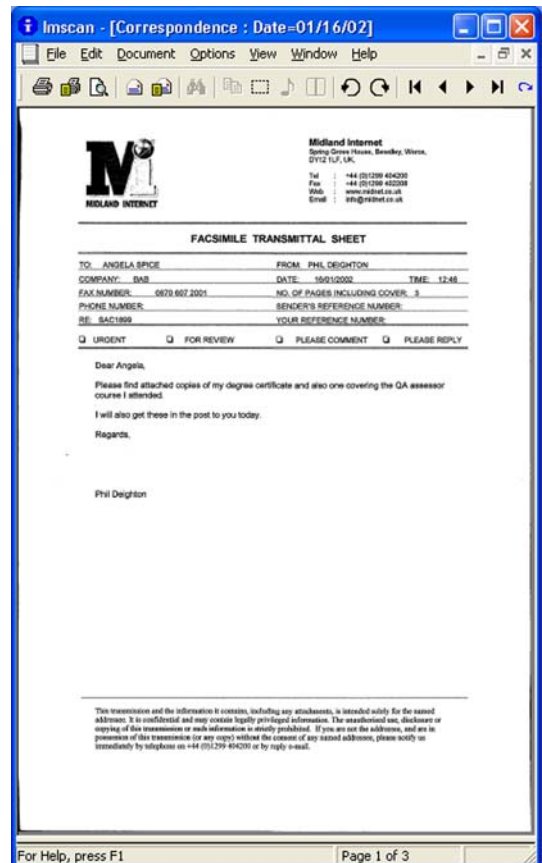
Images held within the system are compatible with most e-mail systems, Intranets and even Internet distribution vehicles.

Indexing methods include manual indexing, single or multiple barcode reading, OCR (Optical Character Recognition) and database lookup/import routines from other systems.

Methods can also be combined to provide maximum flexibility and efficiency with any particular document type.

If documents are retained on CD, the system finds a document and tells you which CD it is on.

Please note that the above describes only the basic retrieval system. A full server based system is available incorporating workflow and document routing. This system is also capable of storing a wide range of document formats including PDF, Word and WAV formats.



PDF Plus Software

Compact have developed a unique document management system that allows users to easily search the complete text of a document and display a PDF image of the selected document.

Up to 5 keywords may be used to select documents and only documents containing all the keywords will be returned.

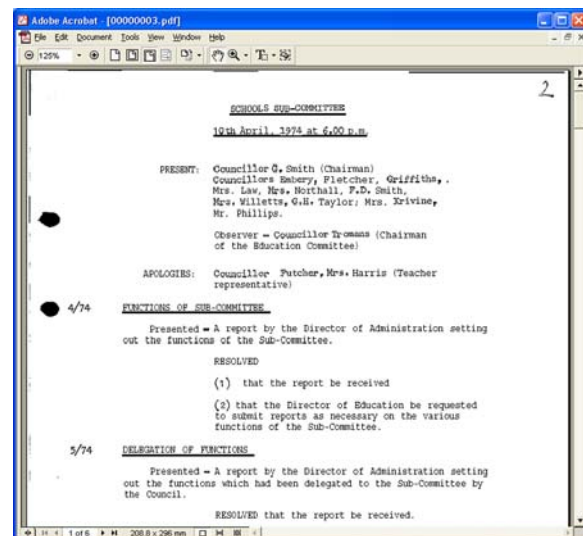


Full text of each document is displayed in compressed format to allow the user to correctly identify the required document.



Documents are then displayed using the industry standard Adobe Acrobat Reader and may then be printed or emailed using standard system tools.

Each batch of documents scanned is added to the central database thus providing full search facilities.



Cost Estimates

Costs for a project are normally based on:

- The cost of document collection
- The cost of document preparation
- The cost of scanning
- The cost of indexing
- The cost of supplying CD's or other media
- The cost of software

Beware that some companies may quote low scanning costs but may charge very heavily for other services. For example, the cost of CD's can vary from £5.00 to £100.00 each.

Evaluating Your Paperwork

A quick survey of the paperwork to be scanned will enable [Compact](#) to provide you with a more accurate quotation in a shorter time-scale.

Listed below are the main items that they will need to evaluate the project:

1. What is the approximate number of pages to be scanned?

As a guide a typical 4-drawer cabinet contains around 8,000 sheets and for papers stacked on their edge on a shelf one linear foot holds around 1000 pages.

2. Is the work to be carried out in a single operation or do you need a regular collection of documents?
3. What is the quality of the paperwork?
4. A useful way of categorising paperwork is as follows:

- A Quality - A4, single/double sided, uniform colour background, single orientation (e.g. reports etc.)
- B Quality - A4, single/double sided, various colour backgrounds, various orientations (e.g. sales invoices)
- C Quality - Various sizes up to A3, various colour backgrounds, various orientations (e.g. purchase invoices / expense claims)

5. Is there any colour to be scanned?

6. Are there any drawings?

7. How do you want to be able to find documents?

Any number of index fields is possible but as each set of data will probably have to be manually entered, more search fields means a higher cost. Longer search fields may also increase the cost as this is charged per keystroke used to enter the data.

8. Is any of the search data already available in electronic format?

Many companies can match existing data with scanned images. Normally a unique field is manually entered for each image and then this is matched with the existing data.

For example, your system may already have a list of customer names and addresses for each invoice. If proofs of delivery are to be scanned in, the invoice number on each POD is entered and matched to the existing data. You can then search by invoice number, customer name or address but will only be charged for the cost of entering the invoice number plus a small charge for data-matching.

9. Do you want the documents returned to you or do you want them destroyed?

[Compact](#) offer secure destruction of documents. This includes the shredding of documents before they leave our premises.

10. Do you want to publish information on the internet?

If you do, then you may want to consider PDF format.

Disadvantages of Electronic Storage

The first disadvantage to electronic document storage is often seen to be that of the cost of scanning and indexing. However, when this is compared to the storage costs for paper and the time saving achievable for document retrieval, in most cases overall cost reductions are achieved by its use.

Security may be a great consideration and care should be taken to ensure that the system chosen can only be accessed by authorised personnel.

Care should be taken to ensure that electronic records are acceptable to regulatory and government bodies. For example, at the time of writing, the Financial Standards Authority

does not accept electronic records in all cases. However, Inland Revenue and Customs and Excise together with most other government bodies are perfectly happy with them.

Summary

The most important factors when choosing a scanning company and a system are based on your needs and their ability to fulfil those needs. If in doubt, ask.

[Compact](#) have had nearly 30 years experience of document management and are now one of the UK's leading data conversion specialists.

With clients ranging from internationally renown Aerospace Manufacturers to the Public Sector and Health Authorities, they have a vast experience of handling critical and sensitive material.

Backed by ISO9002 accreditation and a wide-ranging technical expertise, Compact can offer a solution to suit your needs. All systems can be tailored to exactly fit to your requirements.

The attached evaluation sheet may be of assistance in assessing your initial requirements and ensure that you are provided with an accurate quotation for the project.

Paperwork Evaluation Document

Company:

Department:

Contact:

Telephone:

Date:

Evaluation Point	Criteria	Result
Number of documents to be scanned	4-drawer cabinet = 8,000 1 foot shelving = 1,000	
Document quality	A Quality – A4, uniform colour background, single orientation (reports etc.)	
	B Quality – A4, various colour backgrounds, various orientations (sales invoices)	
	C Quality – Various sizes up to A3, various colour backgrounds, various orientations	
Is colour scanning required?	Yes/No	
Drawing sizes		
Approximate number of drawings		
Indexing details	Number of fields required	
	Average number of keystrokes required	
Preferred format	TIFF	
	PDF	
Is text search required?	Yes/No	
Are copy CD's required?	Yes/No	

Additional Information

Return the completed document to:

Compact Data Management Ltd
6 Leons Way
Tollgate
Beaconside
Stafford
ST16 3HS

Tel: 01785 220846
Fax: 01785 220876
Email: sales@compact.uk.com