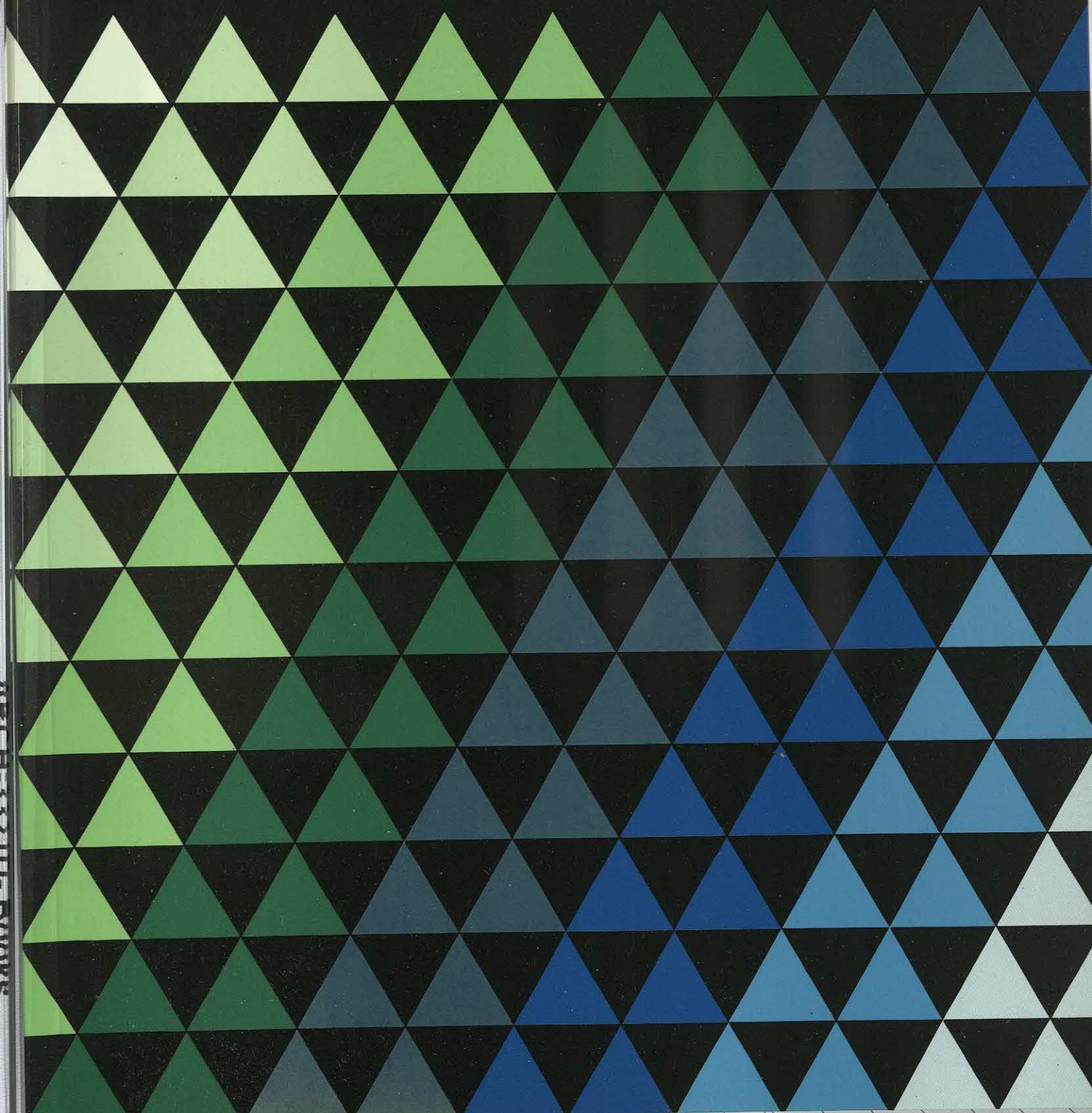
Electronic Document Management

BUTLER COX FOUNDATION

Research Report 70, June 1989



BUTLER COX FOUNDATION

Electronic Document Management

Research Report 70, June 1989

Butler Cox plc

LONDON
AMSTERDAM MUNICH NEW YORK PARIS

Published by Butler Cox plc
Butler Cox House
12 Bloomsbury Square
London WC1A 2LL
England

Availability of reports
Members of the Butler Cox Foundation receive three copies of each report upon publication; additional copies and copies of earlier reports may be purchased by members from Butler Cox.

Photoset and printed in Great Britain by Flexiprint Ltd., Lancing, Sussex.

BUTLER COX FOUNDATION

Electronic Document Management

Research Report 70, June 1989

Contents

A Management Summary of this report has been published separately and distributed to all Foundation members. Additional copies of the Management Summary are available from Butler Cox.

Report synopsis

Electronic document management (EDM) is a new combination of technologies that offers organisations the chance to take a further step in the transition from paper-based to electronic working. EDM is now a practical proposition, and cannot be ignored. It does, however, demand skilled staff to exploit it to the full, places extra responsibilities on systems departments, and requires a heavy investment. The report provides guidance on how to choose applications where EDM will be most beneficial. If these guidelines are followed, the benefits will usually justify the costs.

Chapter 1

Electronic document management is now a practical proposition

The quantities of paper circulating throughout organisations are growing rapidly, making the task of identifying relevant information more and more difficult. Developments such as office automation and facsimile have increased, rather than decreased, the amount of paper-based information, at a time when competitive pressures make it vital to process information as quickly as possible. In fact, the use of information technology has not, by and large, changed the methods used to handle information.

Although many organisations make extensive use of computing and office automation systems, these systems address only a small proportion of the total information held (just 6 per cent, according to a recent survey in the United States). For example, much of the data held in data processing systems is operational information (such as customers' names and addresses, or part numbers), rather than management information, and word processing systems usually contain items such as internal directories and routine correspondence, which have little value in the decision-making process. As a consequence, decision-making still relies heavily on paper-based documents.

Much of the information contained in documents cannot easily be transferred to conventional data processing or office automation systems, because the documents take a variety of forms — maps, legal documents, handwritten papers, photographs, press cuttings, and so forth. Although electronic systems exist to handle each of these document types, they handle them in isolation from each other — data processing for procedural activities, word processing for internal text handling, and CAD/CAM for design work. Thus, until now, electronic information-handling systems have been fragmented, and have often had to be used in conjunction with paper records to make them really useful.

In this report, we explain how electronic document management (EDM) removes the limitations imposed on information handling by conventional technologies. It shows that EDM technology and applications can now bring together all types of information, and make it accessible to users in a common, electronic form.

Despite the obvious advantages of electronic systems that can handle the bulk of paper-based information, EDM technology is not, as yet, being widely used. Our research indicates that potential users have concerns in a range of areas — from the performance of the technology itself, to the legal implications of its application. (The research team and scope of the research carried out for this report are described in Figure 1.1, overleaf.) The concerns are understandable because the concept of EDM is still quite novel. Even so, we believe that the technology has now reached the stage where investment in EDM is a practical proposition; EDM systems providing significant benefits are now in operation, but a large investment is required to gain the benefits. For the right type of application, however, a substantial payback can be achieved within a year or two.

As EDM suppliers better understand the application areas in which their products will deliver the greatest benefits, and direct their marketing efforts accordingly, the use of EDM systems will increase. Confidence in the technology will also grow as potential user organisations realise that they too can achieve the benefits gained by pioneering users. Like any new technology, however, the introduction of EDM will place added burdens on systems managers. To ensure that the benefits of EDM

Figure 1.1 Research team and scope of the study

The research for this report was led by Judith Wainwright, director of the management and office systems consultancy practice within Butler Cox. She was assisted by Rebecca Morgan and Fergal Carton, both consultants in Butler Cox's London office, with a special interest in office automation. Further research was carried out by Onno Schroder (Amsterdam), Lothar Schmidt (Munich), Michel Lederman (Paris), and John Cooper (Sydney). Charles Chang did the research in South-East Asia.

Independent research was undertaken on our behalf in the United States by Lance Brilliantine of Consulting Support Systems Inc. Special guidance on the legal implications of EDM was sought from Said Mosteshar, the managing partner of Mackenzie and Company, a firm specialising in international information technology and communications law.

The research programme was carried out between September 1988 and January 1989. During this period, existing suppliers of EDM were launching major marketing initiatives, and traditional information systems suppliers were announcing new EDM products. To obtain information about the current usage of EDM, we sent a questionnaire to all Foundation members worldwide and received 150 responses. We conducted interviews with 44 Foundation members and held a research workshop with members who were considering whether to use EDM technology. Our objective here was to identify common areas of concern and to explore the perceived costs and benefits.

We also conducted interviews with advanced users outside the Foundation, consulted experienced implementors both within and outside Butler Cox, and spoke to 16 suppliers, both major systems suppliers and niche-market suppliers, in the document image processing area. We also drew on our consultancy experience and the extensive literature available on the subject. (The bibliography at the end of this report lists the publications and articles we recommend for those members who wish to explore the subject of EDM in more detail.)

are achieved for the whole organisation, systems managers will have to plan and coordinate the introduction of EDM into those business areas and those applications where it can make the greatest contribution. They must take up this challenge, because EDM technology has now reached the stage where business managers will begin to demand that it be installed. Unless the systems department takes the lead, there is a danger that EDM applications will be installed in a piecemeal and uncoordinated way.

EDM uses existing information technologies

Our survey of Foundation members revealed widely varying interpretations of EDM. Some members define it as narrowly as a proprietary document-management system; others take a wider view and define it as storage and access to the complete contents of a document, not just the data recorded on that document. For the purposes of this report, we have chosen a definition that encompasses the concept of document management, but that does not prescribe the technologies that might be involved in an EDM system. Our definition is that:

> "EDM is the application of information technology to the problems of managing the flow of paper-based information. EDM systems therefore provide facilities for storing, accessing, and retrieving documents in electronic form. An EDM system uses a variety of technologies to enable all the types of information normally found in a document to be recorded, manipulated, and managed electronically, using common access methods."

An EDM system has four main features

From our definition, it follows that an EDM system has four main features. The first is that there must be *linkages* between the different technologies used to handle the different types of information; free-standing office automation systems or document image processing systems are not, by themselves, EDM. The second is that information must be recorded, manipulated, and managed *electronically*; thus, micrographic systems are not EDM. The third element is the use of a *common access method*. This does not mean that all information is held in a common form, but that a standard method is used to access the different types of information. Fourth, the information flow must be *managed*. In addition to providing access to the information, EDM systems provide facilities for controlling and sequencing the information and for generating management statistics.

EDM links five types of information technology application

The five main types of information technology application that form a powerful combination when linked to create an EDM system are:

- Data processing.
- Office automation, including desktop publishing and personal computing.
- Full-text retrieval, either by a keyword index or free-text searching.
- Document image processing.
- Workflow software.

Systems managers are very familiar with *data processing* and *office automation* applications, both of which are well established. *Full-text-retrieval* applications are also in common use, although not as widely as data processing and office automation applications. Nevertheless, the power and limitations of this type of application are well understood. *Document image processing* is the newest type of information technology involved in EDM. It is still not well established, and its usage is less widespread than conventional data processing and office automation systems. Systems departments are still learning how to implement and exploit the features offered by document image processing. (In September 1987, the Foundation published a Position Paper on document image processing, which described the technology and its possible areas of application.) Originally, document image processing was perceived as a free-standing technology, and used in this way, it was not easy to realise benefits from it. The full potential of document image processing can best be exploited by using it in conjunction with other types of IT application to create an EDM system.

Document image processing is, in fact, the critical component of an EDM system because it provides the means of storing and processing the bulk of paper-based information previously excluded from other types of IT applications. Figure 1.2 illustrates how EDM links document

Figure 1.2 Electronic document management links document image processing and other information technologies

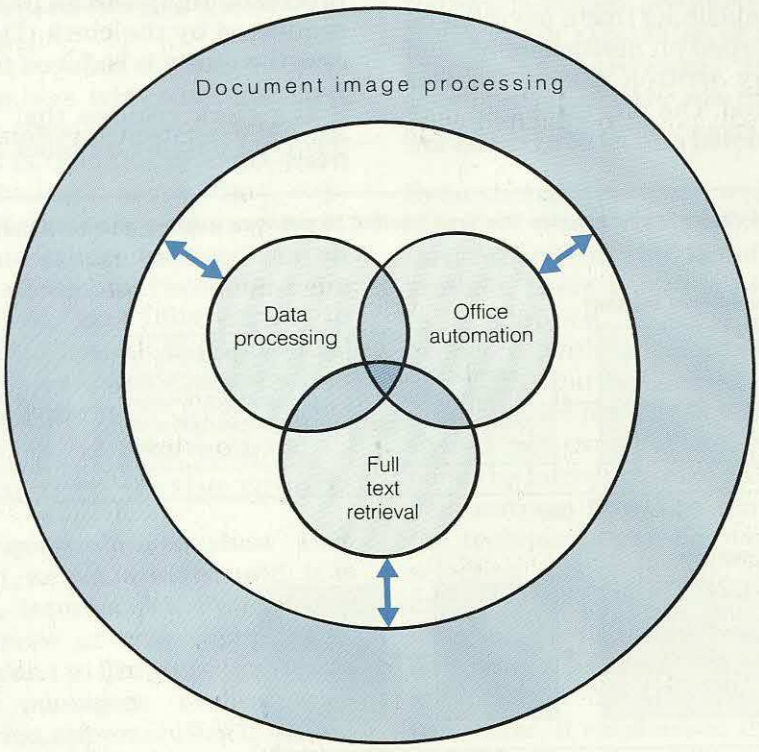


image processing with data processing, office automation, and full-text-retrieval applications. Simple linkages between these last three types are commonplace — an office automation system can store information ready for full-text retrieval, for example, or a data processing system can download data to an office system ready for processing by a spreadsheet. These linkages, however, provide better access to and manipulation of only a small proportion of the total information held by an organisation. It is only by individually linking the three well-established types of application to document image processing that all types of information can be made available online.

The fifth component of an EDM system is the *workflow software*, which controls the flow of documents into and out of the system and highlights areas for exceptional action.

All types of information are available with EDM

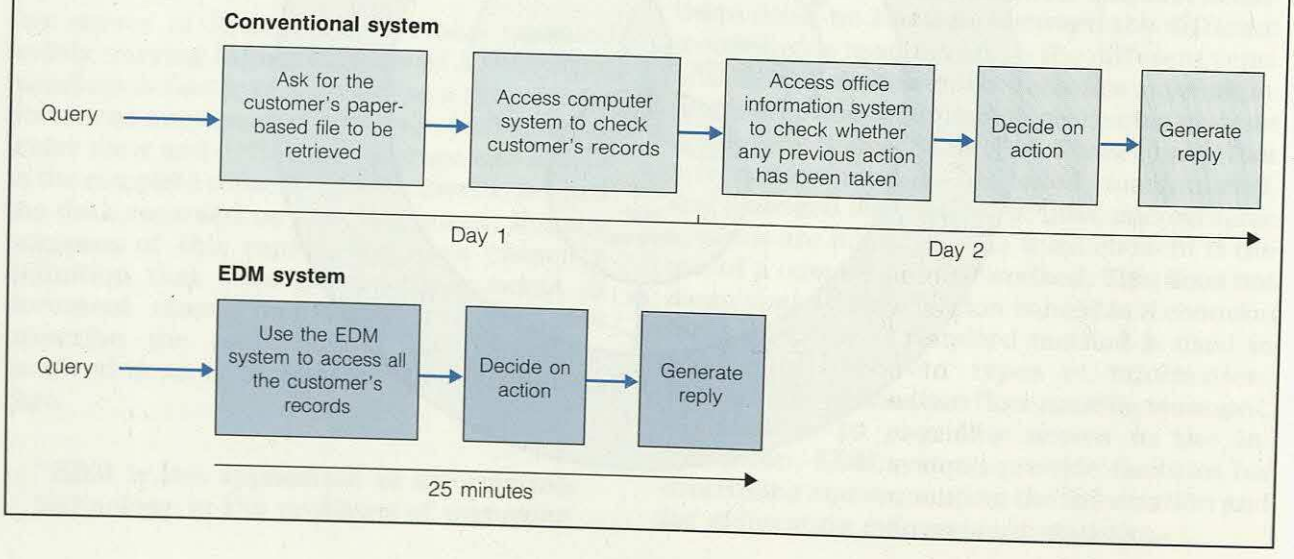
The major advantage of an EDM system is that it permits different types of information to be retrieved in an integrated way. Consider, for example, the work of a clerk in a customer-services department. The clerk receives a letter from a customer complaining that a payment he has made is not recorded on his statement, and that he has already written asking for this problem to be resolved. The processes that take place using conventional and EDM systems are shown in Figure 1.3. In a conventional, automated system, the clerk would probably:

- Check the customer's reference number and ask for the customer's paper-based file to be retrieved.
- Enter the customer's reference number into a computer system to gain access to the customer's records to check if the payment has been received.
- Access the office automation system to check whether a reply has been sent to the previous letter.
- Decide on the appropriate course of action, which will probably result in the use of the office automation system to generate a letter.

This whole process would probably take a minimum of two days. Using EDM, however, the first three steps are reduced to a single access of the EDM system, dramatically reducing the total time needed to process the query and respond. Entering the customer's reference number into the EDM system enables all the relevant information to be accessed. The payment record, an image of the handwritten correspondence from the customer, and the word processed reply will all be available in the order requested by the clerk. The time taken to process the query is reduced from days to minutes.

If the clerk realises that the problem can be resolved only by a higher authority, the query

Figure 1.3 Use of EDM dramatically reduces the time needed to process a query and respond



can be automatically transferred by the EDM system to a supervisor's terminal. In addition, supervisors can use the system to generate statistics on outstanding queries or individual workloads and to distribute the workload between the clerks to optimise throughput. EDM therefore improves both the productivity of the clerks and the quality of customer service.

Foundation members recognise the potential of EDM

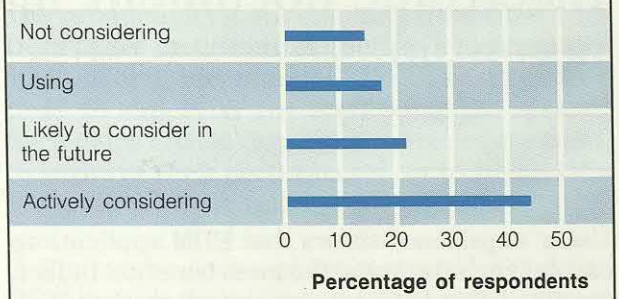
In our survey of Foundation members, we asked if they used (or planned to use) EDM systems. Figure 1.4 shows that more than 85 per cent were using or considering using such systems, although fewer than 20 per cent had already installed EDM systems. Despite the high level of interest, the newness of the technology means that use of EDM is still at the pioneering stage. Nevertheless, the survey results show that those who have already installed EDM systems have benefited from their investment, and that most of them are considering either expanding their existing applications, or installing systems in new areas. In addition, the benefits actually achieved were, in many cases, greater than those expected when the system was originally installed.

However, a few members who have installed EDM systems have no plans to extend them. The main reason for this is that the systems were implemented in business areas that subsequently became fragmented, as management structures and organisations changed. In all cases, these organisations had recognised the intrinsic value of EDM, and fully expect to exploit EDM in other areas. The potential benefits of EDM are therefore well understood by Foundation members.

Purpose and structure of the report

Responses to the questionnaire show that although EDM is not, as yet, widely used, it is an area of great interest to Foundation members. The purpose of this report is to explain what EDM is and to illustrate to systems directors, business managers, and systems project managers what opportunities it offers, and how those opportunities may be converted into tangible benefits.

Figure 1.4 There is widespread interest in EDM, but few Foundation members yet have systems installed



(Source: Survey of Foundation members)

The nature of the benefits provided by EDM is discussed in detail in Chapter 2. Those who have already invested in EDM systems are achieving cost savings, providing a better customer service, and seeing marked improvements in staff morale. For those who have not yet adopted EDM, we indicate the areas of business and the types of application where EDM is likely to be of most benefit. In Chapter 3, we describe how the components of the technology are linked, concentrating on those areas with which systems managers are likely to be least familiar — that is, the capabilities of the document image processing hardware, the trends in its development, and the role of workflow software in delivering effective systems.

Even though there are well-documented case histories (some of which are described in this report) of the benefits that EDM can provide, there is still considerable resistance to its adoption. To some extent, this is to be expected with a relatively new technology like EDM, but we believe that the perceived problems are not as serious as they appear, and should not cause systems managers to postpone the introduction of EDM. In Chapter 4, we demonstrate how the current limitations of the technology can be overcome by carefully selecting, planning, and designing the applications for which EDM is to be used. Chapter 5 provides advice on how EDM should be managed to ensure that the potential benefits are realised for the whole organisation. In particular, it emphasises that the introduction of EDM will create additional responsibilities for the systems department.

Chapter 2

Electronic document management has great potential

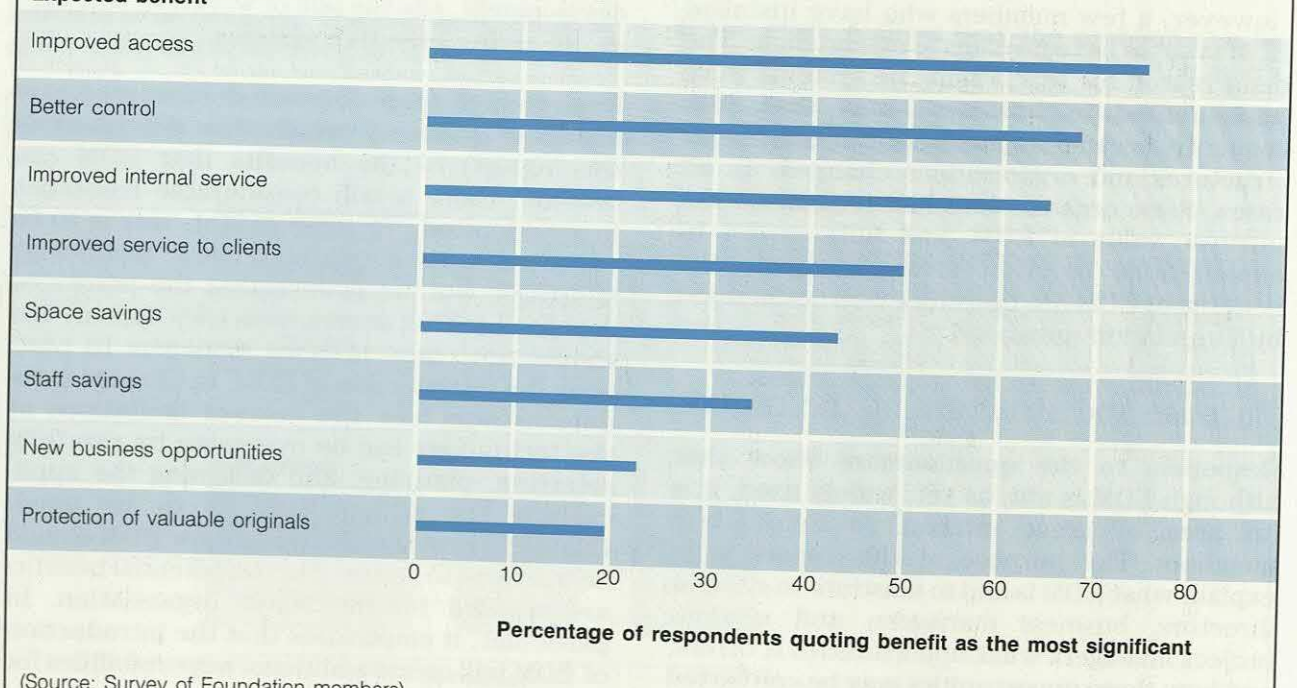
Users' experience shows that EDM applications can deliver substantial business benefits. In fact, provided that applications are selected according to appropriate criteria, the benefits can be much more significant than many Foundation members believe, judging by the results of our survey. Admittedly, to exploit EDM to the full, organisations will have to invest on a large scale, but this has not deterred the pioneering users, many of whom are now receiving substantial returns on their major investments.

In this chapter, we describe the benefits that derive from the introduction of EDM, and identify the types of application and businesses where these benefits can most readily be gained. The typical costs associated with introducing EDM are described at the end of the chapter.

EDM provides significant business benefits

In the questionnaire sent out at the beginning of the research, we asked Foundation members about the benefits they expected to gain from using EDM. The responses are shown in Figure 2.1. The benefits expected by members who have not yet implemented EDM systems correspond closely with those achieved in practice, although the emphasis is different.

Figure 2.1 **Most Foundation members believe that the greatest benefit of an EDM system is improved access to information**



(Source: Survey of Foundation members)

Potential users of EDM placed more emphasis on the less tangible benefits, such as improved access and control, and various internal service improvements, than on new business opportunities, improved service to clients, and cost savings. Those who have not yet implemented EDM tend to underestimate the savings that can often be achieved in practice.

Early users of EDM systems have reported both tangible and intangible benefits, relating to operational efficiency and quality of service, respectively. We have categorised the benefits under four main headings — reduced costs, improved service (which leads to competitive advantage), improved staff morale, and enhanced security. All the EDM users we interviewed had achieved the forecast benefits and, in many cases, had significantly exceeded them.

Reduced costs

We asked EDM users to compare the total annual running costs of their EDM system with their costs prior to the introduction of the system — that is, when computerised records were being used in conjunction with separate word processing, and paper and micrographic records. They reported annual savings of up to 40 per cent, resulting from increased staff productivity and reduced requirements for storage space.

Productivity gains

The productivity gains resulting from EDM arise from faster processing of queries (which reduces staff effort), less manual filing, and a reduced need to supervise the staff using the system. Such gains enable organisations to reduce costs, to improve their performance, or to expand the business without taking on extra staff.

A file can be accessed far more quickly via an EDM system than if it had to be retrieved from an on-site manual filing system, or transferred from a remote site. Queries and transactions can therefore be processed faster. Citibank Savings, a major international banking and finance house, whose system is described in Figure 2.2, introduced an EDM system for mortgage processing. As a result, it has been able to increase staff productivity and, at the same time, dramatically improve the service it provides.

The ability of EDM systems to make information from a wide range of sources available via a single workstation can also have a significant impact on staff efficiency. Glaxo Inc in the United States has introduced EDM to increase the productivity of its research scientists, and to improve the efficiency of its submissions to

Figure 2.2 Introduction of EDM can result in better service for customers and improved productivity

Citibank Savings

Citibank Savings is part of the major American financial-services organisation, Citicorp. It is pioneering the use of EDM in the United States and recently introduced an EDM system in its UK mortgage business. Once a loan has been approved by the underwriter, EDM is used to support all subsequent aspects of mortgage processing — funds transfer, insurance, customer service, and redemption.

A review of the service provided and of the requirements for storing, retrieving, and moving files prompted the interest in EDM. The review revealed that 35 per cent of requests for customers' files could not be met without some delay and that files at the off-site storage facility took at least 24 hours to retrieve. On-site storage and microfiche were considered before deciding that EDM was the most efficient way of achieving customer-service and productivity benefits.

All customer documents are now scanned and indexed in the EDM system and are available for simultaneous inspection within 20 seconds of a request being made. The combination of EDM with standard-letter generation, interfaces with the IBM 3090 mainframe, and improved paper flows (driven by the EDM system software) has resulted in considerable improvements in customer service, in productivity (as high as 40 per cent in some areas), and in reductions in off-site storage costs.

Citibank believes that the success of its EDM system is due to:

- Optimising the complete mortgage-processing system prior to looking at EDM, by conducting a detailed organisation and methods (O&M) study.
- Producing a user-driven software specification.
- Being flexible in its approach and looking closely at what the software could deliver.
- Exploiting workflow software to maximise system benefits.

the US Food and Drugs Administration. Not only has the submission process been speeded up, but major cost savings have been achieved. As a consequence, Glaxo has gained a competitive advantage by getting drugs registered earlier. How the organisation has achieved this is described in Figure 2.3.

The amount of time and effort required for filing can be substantially reduced with the introduction of EDM. Manual filing, or filing into a micrographic system, is time-consuming and costly. Costs of between $5 and $10 per 100 sheets filed are not unusual, when the full costs of storing and retrieving by conventional means are considered. (Approximately 70 per cent of the costs of filing relate to activities undertaken *after* initial filing.) Using EDM, almost all the manual processes can be eliminated, and the chances of losing information through misfiling are minimised. A recent study at a large teaching hospital, which has to file 325,000 case notes and 1.3 million pathology forms each year, showed that the introduction of EDM could result in annual savings of $0.5 million. This would mean that the proposed system would produce a payback within 2.5 years.

In manual systems, users typically file documents under a single heading, because maintenance of cross-references is too cumbersome and labour-intensive. EDM systems allow database techniques to be used for indexing items and provide a variety of ways for accessing a particular item — for example, by reference number, name, or date. Filing errors are still possible, but the risk of losing a record completely in an EDM system is virtually eliminated because alternative ways of accessing the document are always provided.

The improved management and control facilities provided by EDM systems bring further productivity gains. For example, work outstanding from before a specified date can be automatically presented by the system for action. As a consequence, outstanding queries are cleared up faster, thereby reducing the need for progress-chasing enquiries. One user of a new banking EDM application told us that the introduction of the system had resulted in a noticeable reduction in the number of customer complaints. He attributed this to the fact that the system enables the supervisor to spot bottlenecks earlier than had been possible

Figure 2.3 A large pharmaceutical company realised annual savings of $2.6 million with the introduction of EDM

Glaxo Inc

Glaxo Inc is the US research division of Glaxo, a major international pharmaceutical developer and manufacturer. It needed to provide better access for its research scientists to a range of literature, internal research reports, regulatory submissions (of up to 90,000 pages each, supplied to the licensing authority when permission to market a new drug is sought), product literature, and the medical press. Abstracts of all documents were already stored in a free-text-retrieval system, but the scientist often required the full text, and delays were experienced in getting the full documents to the correct person. Frequently, several scientists wanted simultaneous access to the same documents. A micrographic system was investigated, but it was decided that micrographics neither met current needs, nor did it have the potential to expand and offer additional facilities. The decision was therefore made to install an EDM system.

Major benefits have resulted in two areas. The time taken by managers and scientists to retrieve documents has been reduced — analysis of activities showed that the reductions translated to annual savings of $900,000. Space is no longer required in Glaxo's new research and development facility for storing archival information. This space saving is worth $1.7 million a year. The EDM system cost about $2.11 million to install and implement. Glaxo estimates that the annual rate of return on the project is 57 per cent, and that the payback period is just over two years.

In order to maximise the benefits and minimise the problems associated with transferring the huge number of existing documents to the EDM system, Glaxo analysed the use made of existing information. As a result, only the most frequently accessed documents and the most current documents are being transferred to the system. By the end of 1989, it is estimated that the system will hold the images of 2.5 million pages.

The project-team members were drawn from research scientists, the IT department, and external consultants. This proved to be a highly effective combination, and Glaxo believes that the system could not have been implemented in any other way. The system is well received by the scientists who work faster and 'smarter' as a result.

Given the success of this initial project, Glaxo is seeking to extend the system and to integrate it with other IT systems so that regulatory submissions can be prepared faster. The ability to deliver facsimile from the EDM system to remote locations is under consideration and the copyright issues are being addressed.

before, and to reallocate work among the staff. Supervisors can also identify where individuals are having difficulty with particular types of work, and can, if it seems appropriate, organise training for them.

The productivity gains available from EDM mean that some organisations are adopting the technology as part of their plans to launch a new service. This means that the administrative procedures to support the new service can be designed on the assumption that EDM will form part of the standard working methods. During the research, we met an organisation that was involved in the launch of a new insurance product. All justifications for the launch, and the product market plans, were based on the assumption that EDM would be used. The estimates of the long-term support costs for the product were much lower than they had been for earlier, similar products, because an EDM system would be able to handle the increased business created by market growth without the organisation having to employ extra staff.

Space saving

EDM removes the need to store large quantities of paper in expensive office space. This is a particularly important advantage in locations where office-accommodation costs are increasing rapidly. A major corporate bank, which was moving its headquarters, decided to install an EDM system rather than use expensive office space for document storage at the new site. Another company in our research faced the choice of either moving to a larger building to accommodate growth in its current systems, or installing an EDM system. An impending change of location often provides the opportunity to review the current filing systems to assess whether there is a role for EDM.

Improved customer service

EDM can help organisations to improve the service they offer their customers. In many industry sectors, companies now compete on the basis of time rather than of cost. Recent research indicates that many of the world's most successful companies have found speed of response to be a new source of competitive advantage, and that there may be a positive correlation between 'time-based' competition and higher business growth rates. Time-based competition depends on faster throughput, which in turn results in better response to clients. By providing improvements in productivity and a faster turnaround, EDM has a major part to play in time-based competition.

Moreover, the improvements need not be directed just at external customers. Improvements in internal services will often also benefit the ultimate customer — researchers with easy access to a wide range of source material will be able to make the results of their research available earlier; medical staff with access to records held in various registries can deal with patients more efficiently; contractors with access to drawings and legal documents can quickly settle queries raised by project managers.

Improved staff morale

Several of the organisations we interviewed told us that the introduction of EDM often resulted in significant improvements in staff morale. EDM removes paper 'mountains', which results in cleaner and tidier offices, and there are fewer customer queries. A better working environment, and the perception that the job is being handled more efficiently, create a greater sense of satisfaction for the staff using and benefiting from the EDM system.

The experience of British Airways, the UK national air carrier, is a good illustration of the effects of EDM on staff morale. The airline employs more than 9,500 cabin crew; managers of the crew are on duty at Heathrow, Gatwick, and other main bases in the United Kingdom, to handle enquiries from arriving and departing cabin crew on such matters as performance appraisals, adverse or complimentary reports from passengers, and forthcoming training programmes. More often than not, crew members are seeking information just before or just after a trip. The conventional paper-based filing system meant that it was difficult to access the required information, and crew members inevitably had to wait, at a time when they were under pressure. British Airways installed a prototype EDM system and discovered that, in addition to increasing the productivity of both administrative and management staff, the cabin crew felt that they were being cared for because their queries received immediate attention, and were answered accurately.

Enhanced security

An additional benefit of EDM systems is that they enhance the security of the information

stored in them. In conventional filing systems, the high costs of storage often make it uneconomic to store duplicate copies of paper records. Copies of EDM files, however, can be held off-site, in the same way as security and back-up copies of data processing files are held off-site. In practice, most organisations retain the originals of documents stored in EDM systems, usually in low-cost off-site storage. In future, as electronic working becomes the norm, we believe that many organisations will be prepared to throw away the paper copies of many records.

Benefits depend on selecting appropriate types of application

EDM provides the greatest benefits when unstructured, live (as opposed to archived) information of high value is being handled, and where frequent and easy access is critical. Where the information is structured, and hence, predictable, other information technologies will usually be more appropriate. In the early days of EDM, document image processing systems were typically installed to handle archived material and to support research-library applications. (In effect, the early systems were seen as an electronic replacement for micrographic systems.) Because these systems had just one workstation, users wanting to access the information had to queue up. As a consequence, the technology did not provide significant benefits, and the costs required to introduce it were high. These early experiences have shown that the benefits of EDM will be achieved only if the technology is applied to appropriate types of application.

Applications involving live information

Live information is that which is either currently used in the day-to-day running of an organisation or could be at some future date. Examples include medical records, police records, credit-card vouchers, and correspondence relating to insurance claims. On the other hand, copies of invoices, payment-history records, and stock records are not required (and are not likely to be required) for the day-to-day running of an organisation; these are not live files. The applications that handle live information can be subdivided into two categories — transaction and reference applications, each of which has quite different characteristics.

Transaction applications

We define a transaction application as one in which paper-based information forms part of the standard procedure for progressing a transaction through the system. All the information handled by the application is of a similar nature, and is often recorded on standard forms. The contents of documents are not predictable, however. Transaction applications usually deal with high volumes of information that has a short active life, but that may be reactivated at a later date. Examples include insurance claims, mortgage applications, leases, and descriptions of accidents. While an insurance claim is being processed, frequent access to the information is required over a period of a few weeks; thereafter, the need to access the information all but disappears. Once a mortgage has been granted, no further access to the information may be required until the mortgage is closed many years later. The typical benefits resulting from using EDM for this type of application are improved turnaround, reduced clerical effort, and reduced filing space.

Reference applications

Reference applications are those in which high-value information from a variety of sources is held and accumulated. Such information has a long life. The format of the information cannot be predicted, nor can the patterns of access to it, because they are dependent on changing circumstances over long periods of time. Applications of this type include patent records, research literature, company annual returns, and medical records. Patent records, for instance, will be valuable even after the lapse of the actual patent, and the information will retain its research value for many decades. Typical benefits from applying EDM in this area are better productivity from professional staff, and reduced filing space.

Applications involving unstructured information of high value, to which frequent and easy access is essential

EDM systems deliver the greatest benefits when they are used to handle unstructured information of high value, to which frequent and easy access is essential. Figure 2.4 illustrates the

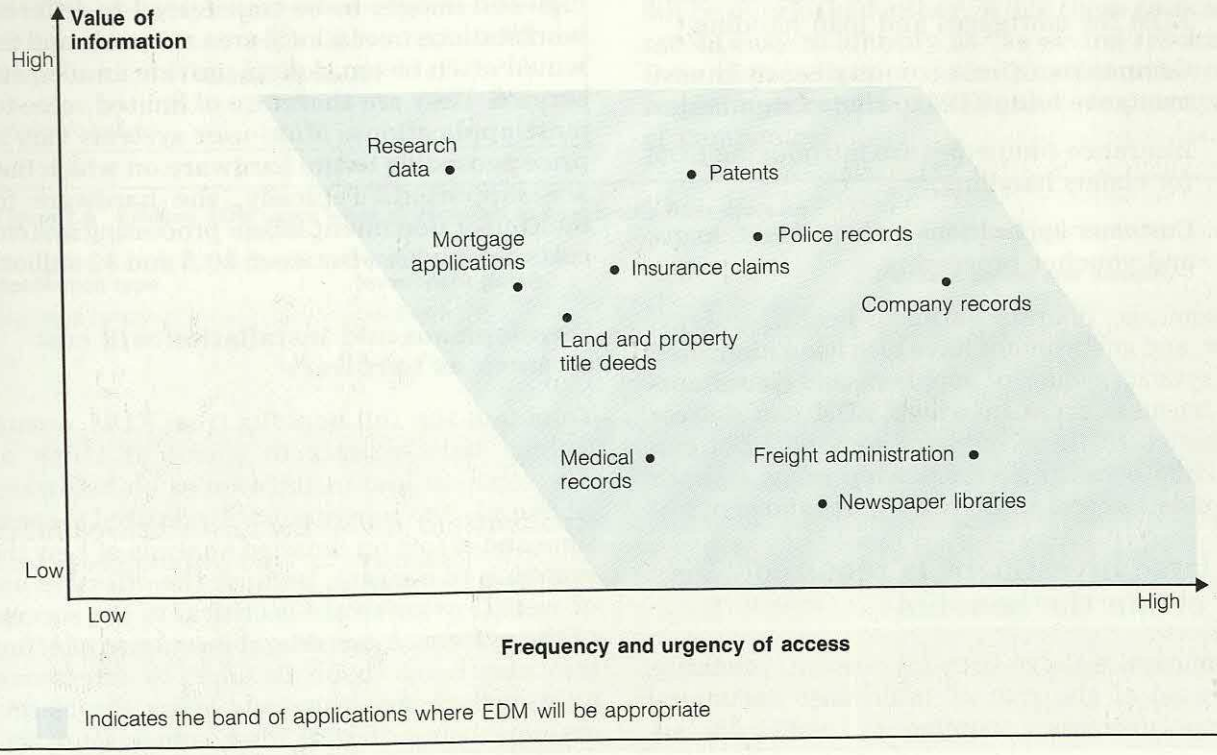
band of applications where EDM would be more appropriate than other manual or automated systems.

The main characteristics of *unstructured information* are that the detailed contents of records cannot be predicted, and their physical layout varies widely — examples include press cuttings using different typefaces and different layouts, photographs, hand-written application forms, medical records, diagrams, and signatures. Information has a *high value* if it contributes directly to the efficiency and effectiveness of an organisation by enabling important decisions to be made (or made better) and actions to be taken. New mortgage applications are of high value in the competitive mortgage market. Historic payment schedules on completed mortgages are of low value. *Frequent and easy access* to information is necessary if time is of the essence or if the information may need to be accessed simultaneously by several people. The first example quoted below illustrates an appropriate application for EDM, where the characteristics discussed above apply; the second example illustrates an application where other systems would be a better choice.

The first example concerns a new system being designed by a police force to provide electronic access to photographs of suspects, and to their criminal records. Records are selected using a range of criteria, such as physical description, types of crime committed, and so on. This information is of high value, as the faster a suspect can be identified, the greater the success rate in solving crimes. There is obviously a need for easy access to the information. Moreover, records about a particular suspect may be of interest in more than one case at a time, although access to an individual's records may well be very infrequent. Because the information is unstructured, of high value, and with a requirement for easy access, the application is suitable for EDM.

The second example concerns the storage of invoice copies, which are retained for fiscal purposes. The documents are structured, and the data contained in them is also probably held in coded form on magnetic media. A copy of an

Figure 2.4 EDM is most appropriate for applications where unstructured information is of high value and where access to it is frequent or urgent



invoice does not contribute to any decision-making process, nor does it trigger any operational action; it is therefore of low value. There will not be much need to access the documents, and it is unlikely that more than one person will ever need to access a single document at the same time. Therefore, although an EDM system could be used to store the documents and provide access to them, other technologies such as data processing or micrographics would be more appropriate and more cost-effective.

Benefits will be easier to obtain in particular types of business

Investment in EDM will be justified in business sectors where the concept of time-based competition is understood, where large volumes of paper are handled, and where the information is live, unstructured, of high value, and needs to be accessed quickly and/or frequently. The finance sector was one of the first to see the advantages of EDM. All the Foundation members in this sector who responded to the questionnaire are actively considering, or already using, EDM. The systems being introduced are typically transaction-based applications:

- Personal financial-services firms are using EDM for mortgages and loan handing.
- Companies offering equity-based investments are using EDM to control new issues.
- Insurance companies are introducing EDM for claims handling.
- Customer applications include credit, loans, and voucher processing.

Chemicals, pharmaceuticals, health, energy, law, and government have also been identified by systems managers, suppliers, and consultants as business areas in which EDM can deliver benefits. In these industry sectors, EDM can provide benefits across a wide range of both transaction and reference applications.

A large investment is required to obtain the benefits

EDM requires large-scale investment, primarily because of the cost of multi-user document image processing equipment, which is an essential component if unstructured information is to be handled and easy access is to be provided. As the technology is relatively new, major systems design effort will be necessary if the full benefits of EDM are to be realised and the costs are to be justified. However, as the technology matures and comes into widespread use, prices will inevitably begin to fall.

Hardware costs are substantial

Hardware costs for a full EDM system vary widely, depending on the equipment supporting the document image processing element of the system. In general, the cost of different types of system depends on the number of potential users. Figure 2.5 shows the costs for a selection of full EDM systems available from a range of suppliers, and plots these against the number of users who can be supported by the systems. The costs shown in the figure are for the minimum configuration considered viable by the individual suppliers with whom we discussed this issue.

EDM systems may be based on personal computers, minicomputers, or mainframe equipment. Personal-computer-based systems are, with few exceptions, single-user systems; a multi-user system of this type would require digitised images to be transferred to different workstations over a local area network, and this would often be too slow to provide an adequate service. They are therefore of limited value for most applications. Multi-user systems vary in price according to the hardware on which they are supported. Typically, the hardware for multi-user document image processing systems costs somewhere between $0.5 and $2 million.

Development and installation will cost as much as hardware

To obtain the full benefits from EDM, organisations must expect to spend as much on development and installation as on hardware. Above all, developers must be prepared to spend time and effort on detailed analysis of how the system is to operate, because the effective use of workflow software is critical to the success of the system. A considerable amount of effort will also need to be devoted to integrating existing data processing and office automation systems, using EDM as the common interface.

The issues to be resolved when designing an EDM application are discussed in more detail in Chapter 4.

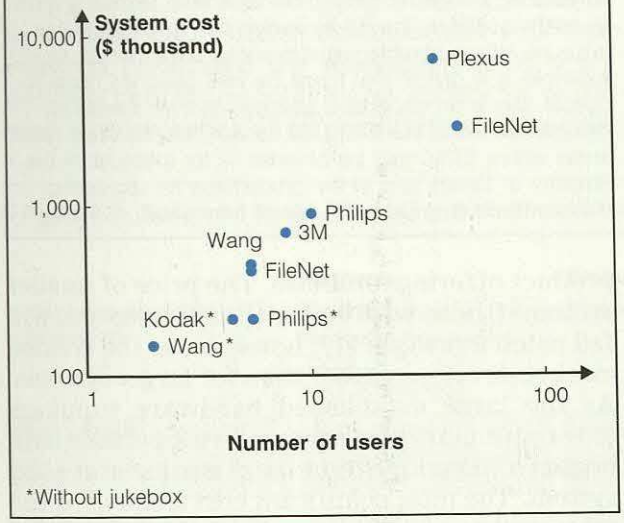
The benefits make the investment worthwhile

While recognising the areas in which they might benefit, 40 per cent of Foundation members in our survey were concerned that the costs of EDM could not be justified. This concern is unfounded, however. In practice, the benefits can be realised and, by ensuring that applications are selected according to the right criteria, the payback period may be very short.

Although the investment costs are undoubtedly high, many of the EDM systems already installed have produced a payback in between one and three years. The user organisations we interviewed had all been required to cost-justify their systems, and to show that they would produce a positive cash flow within a few years, in order to gain approval to proceed. Figure 2.6 shows the payback period and benefits quoted for a range of application types. In all cases, the promise of productivity improvements was the deciding factor in obtaining authorisation to implement the EDM system.

Organisations already using EDM are prepared to invest in further use of the technology because of the proven success of their existing systems. Experience shows that, initially, managers question the role, cost, and justification of suggested implementations. Once the benefits of the early systems are realised, requests for subsequent funding are regarded more favourably. As costs fall, it will be easier to justify the required investment, because the benefits are in areas of rising cost — staff, recruitment, and space. The cost of losing information, or of it not being available, must not be overlooked either. While these costs are not as easy to identify as the saving resulting from staff reductions, such omissions have a major impact on the success, and even survival, of companies.

Figure 2.5 The point at which an EDM system can be cost-justified depends on the potential number of users



*Without jukebox

Figure 2.6 Existing EDM users were all required to cost-justify their systems

Application type	Investment (total)		Payback period and benefits
Mortgage handling	Hardware Design and implementation Software	$1.80 million $0.16 million $0.30 million	Payback period: 3 years Benefits: Staff reductions Reduced staff turnover Space reductions
Research records	Hardware Design and implementation Software	$1.08 million $0.27 million $0.54 million	Payback period: 2.1 years Benefits: Better use of professional staff time Space reductions
Loan handling (including OCR)	Hardware Design and implementation Software	$2.50 million $0.60 million $0.70 million	Payback period: 2 to 3 years Benefits: Reduced need for bulk data entry
Hospital records	Estimated total based on pilot	$2.50 million	Payback period: 2 to 3 years Benefits: Reduced storage costs Reductions in administrative time

Figure 2.7 EDM may make it possible for organisations to offer new services

The Companies Registration Office, Republic of Ireland

The Irish Companies Registration Office (CRO) is the repository of statutory records relating to companies registered in the Republic of Ireland. These records (articles of incorporation, annual returns, accounts, and other associated documents) are provided by the individual companies and made available to the general public, who pay a small enquiry fee.

At the time of writing, the CRO has successfully installed, and nearly completed, a pilot EDM system and, based on its experience, was likely to proceed with a full implementation. The move to EDM was prompted by the fact that the existing computer system could address only 30 to 35 per cent of the enquiries made, because most information was still held on paper files. Retrieving the paper files was labour-intensive and the failure rate was as high as 25 per cent, because of misfiling or because another person already had the file. Furthermore, the volume of information and the number of enquiries was increasing while the staff levels were fixed, because of a ceiling imposed on the number of civil service staff.

The aim of the pilot was to assess whether EDM could:

— Improve the service to the public by reducing the high failure rate, reducing the time lag, and making enquirers more self-sufficient.

— Improve the efficiency of the CRO staff, because they would not be diverted to answer queries that members of the public could answer themselves, once they had access to the full records.

— Save money by moving the paper records to less expensive archival storage, rather than storing them at the central location.

The pilot has demonstrated that a full system would pay for itself. There would be staff savings and increased productivity, resulting in an increase in fees from the general public. EDM would also make it possible for the CRO to offer additional services — online access for larger users, and a facsimile enquiry facility. The CRO's users are prepared to pay more for an improved service.

The pilot illustrates that the benefits derive not just from introducing image recording, but also from exploiting workflow software; the full operation has to be taken into account. When scanning and indexing are involved, for example, it is critical that forms be well designed because optical disc is an expensive storage medium. Based on this pilot scheme, the Irish Civil Service has identified other areas where EDM may be of value — for example, in the Registry of Deeds and in the procedures for registering births and deaths.

Benefits also derive from the ability to offer new services. The Companies Registration Office in the Republic of Ireland installed an EDM system to manage information that has to be maintained by law. Its introduction has also enabled the registration office to provide clients with a new service. EDM has changed the role of the office from one where it reacted to requests from the public to look up the information, to one where it now offers services. The experience of the Companies Registration Office is described in Figure 2.7.

Prices will fall, but not in the short term

The total price of larger EDM systems (which we define as those with more than five users) is unlikely to change significantly over the next five years; thereafter, the price is likely to fall as the technology becomes more mature and sales volumes increase.

Initially, the price of larger EDM systems will fall slowly, as functionality increases and product offerings stabilise. The price of smaller systems (those with fewer than five users), will fall much more quickly, however, as the volume market develops earlier than for larger systems. As the large established hardware suppliers enter the market, increased competition will begin to have an effect on the price of a total system. The most significant cost element of the total system will continue to be that associated with customising the basic package. While basic features will be packaged, the more sophisticated EDM features made possible by advances in hardware and software will still have to be tailored to meet individual needs.

In this chapter, we have reviewed the benefits that EDM can provide and indicated the types of application and business where the benefits will be easiest to obtain. We have also shown that the benefits are large enough to justify the investments required to install EDM systems. The benefits, however, depend on making the best use of the technology that forms the central core of an EDM system. In Chapter 3, we describe the components of the technology in more detail.

Chapter 3

The technology is already advanced and is continuing to evolve

In Chapter 1, we showed how EDM links document image processing with data processing, office automation, and full-text-retrieval systems to provide the means for electronically managing the full range of information, and for integrating documents originating in their most pervasive form (paper) with all other forms of information. Data processing made it possible to manage data; office systems and full-text retrieval facilitated the management of text; desktop publishing added the ability to work with various types of information at the same time. Figure 3.1 shows how document image processing provides the final component of total information management.

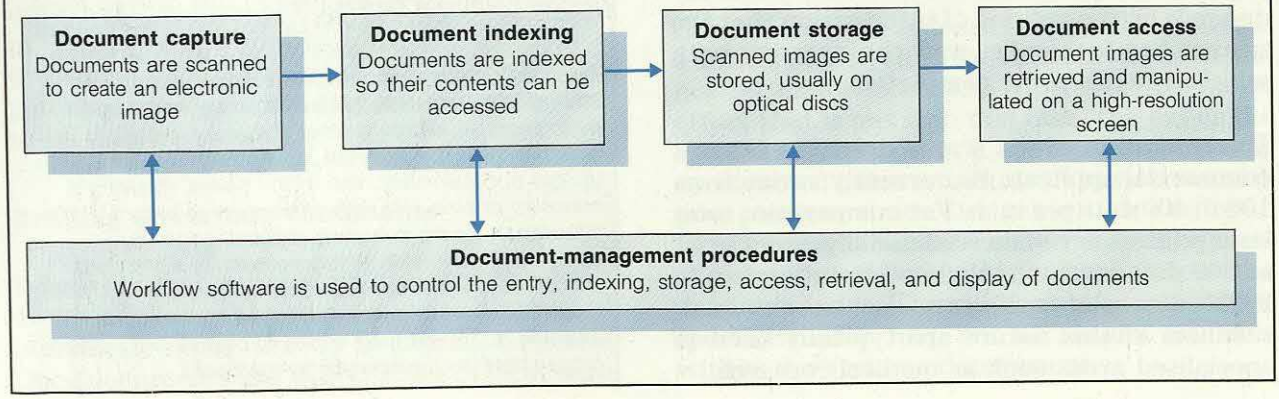
EDM systems consist of five elements, illustrated in Figure 3.2:

- Cost-effective document-capture facilities.
- The means for indexing the information.
- Cost-effective and secure storage for the information.

Figure 3.1 Document image processing is the final component of full information management

	Capture	Storage	Retrieval	Distribution
Structured data (data processing)	Data entry Scanning (bar code, OCR) Unit is data item	Magnetic media	Full database search	Networks EDI
Text (office systems)	Word processing Unit is text document	Magnetic media	'Directory' index Free text retrieval	Electronic mail Shared filing
Paper records (document image processing)	Scanning Unit is scanned image	Optical media	Keyword indexing Serial indexing	Facsimile Shared filing

Figure 3.2 EDM systems consist of five elements



— Easy access for users.

— Workflow software, which manages the flow of documents and information through the system.

In this chapter, we look at the technology that is used for each of these elements, describe how it works, and illustrate how it is linked to the others to provide a total EDM system. We also assess how these technologies are likely to change, and what impact such changes will have on the capabilities of EDM systems.

Capturing the document

Documents that are to be stored in an EDM system exist in one of two forms — hard-copy (paper), or electronic. Paper documents are first scanned to create an electronic copy of the document image. Once scanned, the electronic image of the document is compressed to minimise its storage requirements. The electronic copy can then be used instead of the original form of the document, thereby dispensing with the need to retain the original. Documents that are already in an electronic form (for example, a word processed document) may usually be input directly to the EDM system.

Scanning from paper

Paper documents are entered into an EDM system by digitising the images of the pages. This is done by scanning the pages (using the same process as is used to transmit a facsimile) and storing the images as digitised 'bit maps' in computer memory. The scanning and digitising process is described more fully in Figure 3.3. Documents can be scanned at speeds ranging from less than one second to one minute per A4 page. The speed of the scanner normally depends on the volume of information that the system has to handle, and the urgency with which it needs to be completed.

The resolution of the scanned images in most commercial applications currently varies from 100 to 400 dots per inch. For comparison, most laser printers in commercial use at present print at 300 dots per inch. Higher resolution (up to 2,000 dots per inch) can also be obtained; scanners of this nature are typically used in specialised areas such as medical research.

Before the document image is stored, it is compressed (by removing redundant information) to minimise the amount of memory required to store the image. This is essential because an uncompressed image of a text document requires about 500 times as much storage as that of its coded equivalent (that is, text stored in normal word processing format). While an A4 page of text stored within a word processing system occupies about 2.4 kilobytes of storage, the equivalent document held in uncompressed image format currently requires more than 1 megabyte of memory. After compression, an A4 digitised image typically requires only 50 kilobytes of storage. The compression process is described in Figure 3.4.

Converting into revisable text

In the majority of commercial applications, the text documents are stored as images and cannot be amended or edited in the same way as word processed documents. However, OCR techniques can be used to 'read' the document images, if they are text, and to convert them into revisable text, using one of two techniques — template matching or pattern matching, which are described in Figure 3.5. Converting an image into a machine-readable format is a slow process. Users must balance the advantage of being able to edit scanned documents against

Figure 3.3 Digitising a document image involves the creation of a bit map from hard copy

Scanners use a matrix of photodiodes to scan a page that has been illuminated by a light source. Light reflected by the white space on the page generates a voltage in the photodiodes, while light absorbed by the dark areas of the page does not. This combination of on/off voltage levels creates an analogue data pattern, which is digitised into a matrix of binary data ready for storage in random-access memory. The digitised image is created by dividing the page into tiny dots or picture cells (pixels), and information about each pixel is then recorded. The resolution of the scanner dictates the amount of information that needs to be stored about each page. At standard office-document resolution (300 dots per inch), for example, an A4 page will be broken up into approximately nine million pixels, requiring a minimum of nine megabits of memory to record all the information for the complete digitised page. By comparison, the storage requirement for a standard word processed A4 page would be only 2.4 kilobytes. If more than one bit is required for each pixel (for example, to record grey scales or colour), the storage requirement is correspondingly increased.

Figure 3.4 Compression techniques are necessary to minimise the size of image files

Compression techniques reduce the amount of image data that has to be stored by minimising the amount of information required for non-critical parts of the page — for example, white space. Developers of document image processing systems adopted the standard compression routines previously developed for transmitting facsimile images. As facsimile machines communicate via normal telephone lines, methods for reducing the amount of data required per page were necessary if transmission speed was to be reduced to an acceptable level. These compression standards (Groups III and IV) typically reduce the amount of data by a factor of between 15 and 30. This means that the image of a standard A4 page would be reduced to between 30 and 70 kilobytes for storage.

Figure 3.5 Two OCR techniques are used to convert images into a revisable form

There are two methods of character recognition — template matching and pattern matching. With the former, more basic technique, each typeface that is capable of being read will have a matching template stored in memory, thus limiting the number of different fonts and point sizes that can be read. Pattern-matching programs, on the other hand, use artificial-intelligence techniques to create their own templates and to recognise the topological features of each character. This form of OCR may, in theory, recognise any character, regardless of typeface. Both forms of OCR can retain information about a typeface, and once a character has been recognised for a given document, it is 'remembered' for the remainder of the text.

the reduced speed of throughput. It typically takes a matter of minutes to scan and process a single A4 page so that it can be stored in a machine-readable format. Furthermore, since the accuracy levels achieved are never 100 per cent, extra time is required to correct the errors. Systems of this type are used mainly in engineering applications (for technical drawings), and in legal applications (for potential evidence that may subsequently need to be fully retrieved).

Reading electronic documents

Where documents are already stored in a coded electronic form, they do not need to be scanned, but can be input directly to the EDM system. Once documents are stored in an EDM system, there is no distinction between those documents that were initially paper-based and those that originated in an electronic form, although the system usually sets an internal indicator so that it can 'remember' the original document type. To enable a document originating in an electronic form (such as word processed text, electronic mail, telex, or a file downloaded from on online database) to be manipulated within the EDM system, it may be necessary to convert it from its original format to a standard text format, such as ASCII. Because EDM systems store documents in a facsimile-compatible format, many systems can accept facsimile documents as direct input. Figure 3.6, overleaf, shows how different types of documents may be input to an EDM system; regardless of their initial format, all types are stored and indexed in the same way within the system.

Indexing the document

Once the document has been captured, it is indexed so that the contents of the document can be accessed. Depending on the application, indexes can be as simple or as sophisticated as the user requires. In their simplest form, they may be a straightforward listing of all the files stored in the EDM system; more sophisticated indexes provide both information about the document content (title, summary, key words), and information about the way the document was created (date of creation, author, date of last revision, number of pages, and so forth). Most EDM systems use relational databases to hold the index information, which means that the indexes can be accessed using the powerful and flexible searching facilities provided by such databases. Indexes can be created either manually, through direct entry of the index terms, or automatically by using bar-coding or OCR techniques.

Constructing the indexes for an EDM system is a critical activity, and the effort involved should not be under-estimated. However, the more effort that is put into this task, the easier the system will be to use. Inappropriate or poorly constructed indexes will result in systems with poor response times. Designing the index for an EDM system to minimise retrieval times requires a knowledge of both records management and database design. The right combination of these specialist skills may not be readily available within the systems department.

Using relational database management systems

EDM indexes are typically based on modern relational database systems (such as Informix, or Ingres). These allow users to search large volumes of information rapidly, using specified keys. This approach represents a compromise between the simplicity of the document-filing facilities of traditional office systems, where most files are held under a single reference, and the functionality of full-text-retrieval systems, where all the information is reviewed during a search. Using database techniques forces the index designer to think through exactly what information will be required to create a useful index, and often leads to a better understanding of the way in which the information stored in the EDM system is used within the organisation. In turn, this can lead to more appropriate procedures for retaining and storing documents within the system.

Figure 3.6 Documents may be input into an EDM system in several ways

Documents are indexed and stored in the same way, regardless of the way they are input.

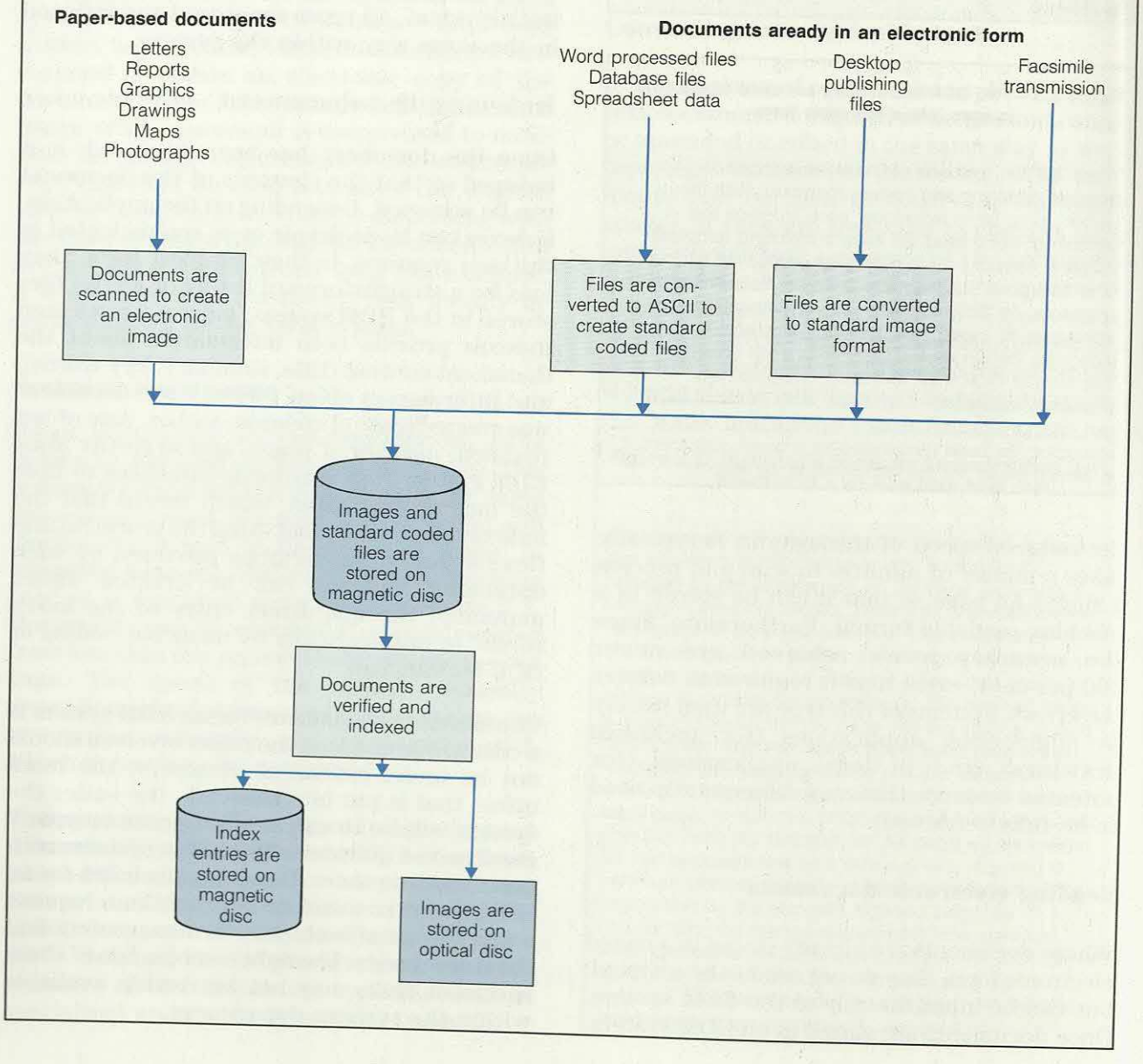


Figure 3.7 shows how a business letter might be indexed within an EDM system. This example illustrates one of the difficulties of good indexing. As the figure shows, the index is composed of a combination of factual information about the document (its author and date, for example), subjective information about its importance (urgency), and selected keywords that indicate its content. The usefulness of these latter items usually depends on the skill (and experience) of the indexer. Indexing is labour-intensive; using experienced staff for this task can greatly increase the cost of the indexing process. A balance therefore has to be sought between the cost of generating the index and the value of the index items. As we illustrate in Chapter 4, the cost of creating an index for an EDM system can be so high that it is not worthwhile moving from an existing paper-based system.

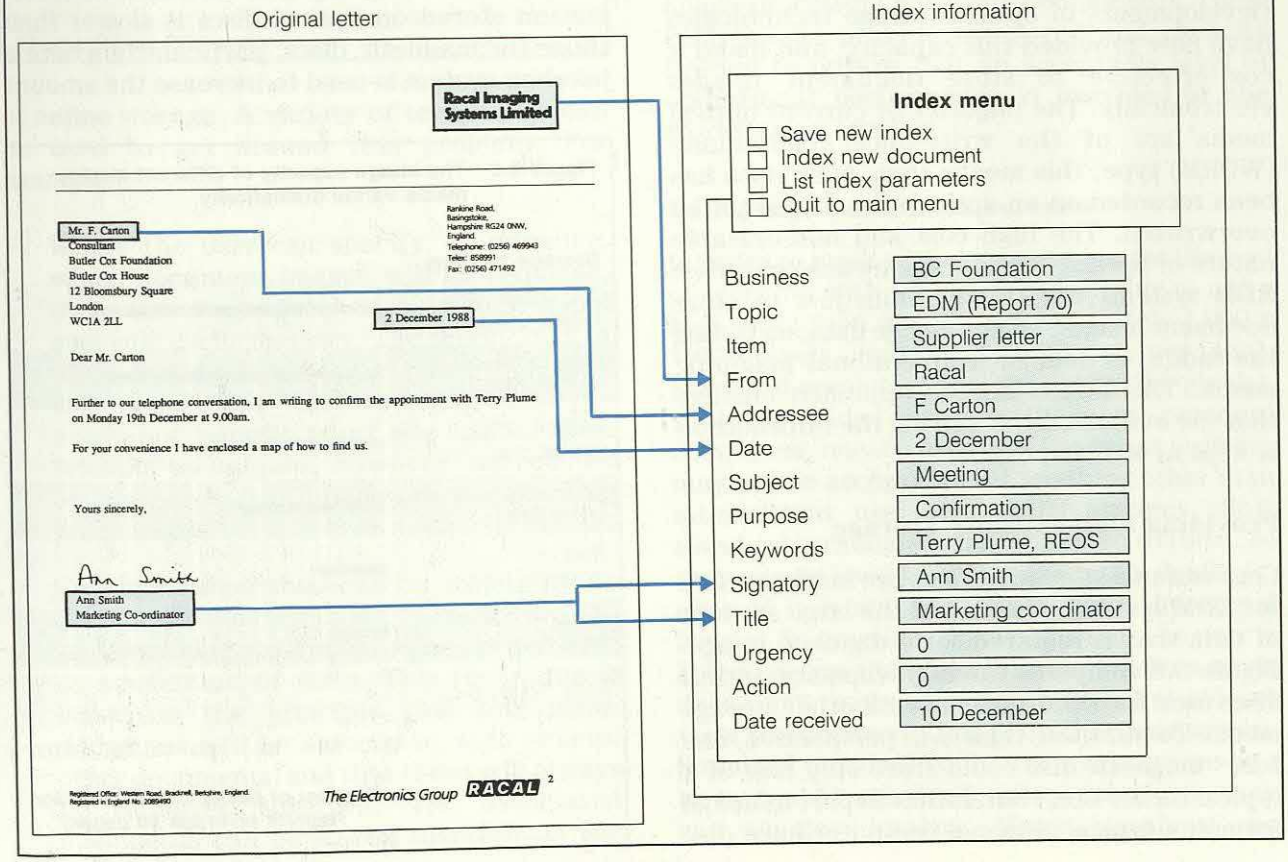
Using an existing database structure

It is often possible to customise an existing database management system so that it can be used to maintain the indexes for an EDM system. Adopting this approach has two main advantages. First, it provides compatibility with existing systems, and makes it easier to provide technical support for the EDM system. Second, it makes it possible to link document images with relevant data files held in the main database, using the same indexes. For example, an insurance company's central client database might include not only information about the premiums received from each client, but also index fields for an EDM system, which could provide access to electronic copies of policies, claims, and correspondence with the client.

Indexing automatically

Automatic indexing techniques, usually based on bar-code reading or OCR, are now being developed. Such techniques selectively scan sections of the document to identify the index information. The data scanned will generate index references for that particular document, removing the need to key in some (or all) of the

Figure 3.7 Good index design optimises the retrieval of information



document-reference details. By this means, documents can be scanned and the required indexes generated with no manual intervention. To date, bar-code reading and OCR techniques can be used only with standard forms and other internally generated documents that are specially printed with electronic indexing in mind. There are some technical complications with both techniques, however — in particular, they reduce the speed of throughput, requiring greater processing power, and they require additional manual procedures to handle those documents that for a variety of reasons, the automatic indexing system is unable to cope with.

In the future, it may be possible to use natural-language processing and other artificial-intelligence techniques to index certain types of documents automatically.

Storing the information

Once the document has been captured and indexed, it is ready to be stored by the EDM system. The storage of document images requires very-high-capacity storage media. Developments in optical-storage technologies have now provided this capacity, and make it cost-effective to store document images electronically. The majority of current optical media are of the write-once ready-many (WORM) type; this means that, once data has been recorded on an optical disc, it cannot be overwritten. The high cost and non-erasable nature of optical-disc storage means that most EDM systems use this medium just to store document images; all non-image data, including the index, is held on conventional magnetic media. This mixed-storage approach ensures that the overall cost of storing the information is kept as low as possible.

Providing high-capacity storage

Conventional storage devices are inappropriate for EDM systems, because of the large amount of data that is required for a digitised image. Figure 3.8 compares the high-capacity optical discs used for EDM systems with other storage media. To put these figures in perspective, one 5.25″ magnetic disc could store only half of a typical Butler Cox Foundation Report in image format; a typical personal computer hard disc (20 megabytes) could store the equivalent of 10 reports. A single optical disc could store somewhere between 200 and 1,000 reports. Devices called jukeboxes, which allow access to multiple optical discs, could, in theory, extend this to the equivalent of 200,000 reports.

To date, EDM systems have relied on WORM discs, which cannot be updated or erased. Erasable optical media are now becoming available, however, and these will have a significant influence on the design of EDM systems. Where users require online access to large volumes of documents, a jukebox can be used to load a disc automatically, from a library, into an optical-disc reader. When there is a need to access a document that is not held on the currently loaded disc, the system automatically searches for the correct disc, takes it from the storage rack, and inserts it into the reader. Figure 3.9 indicates the capacity and access speeds of two different jukebox systems provided by FileNet Corporation.

Compensating for slower access times

The access times required to retrieve information stored on optical discs is slower than those for magnetic discs, particularly where a jukebox system is used to increase the amount

Figure 3.8 The image capacity of different storage media varies dramatically

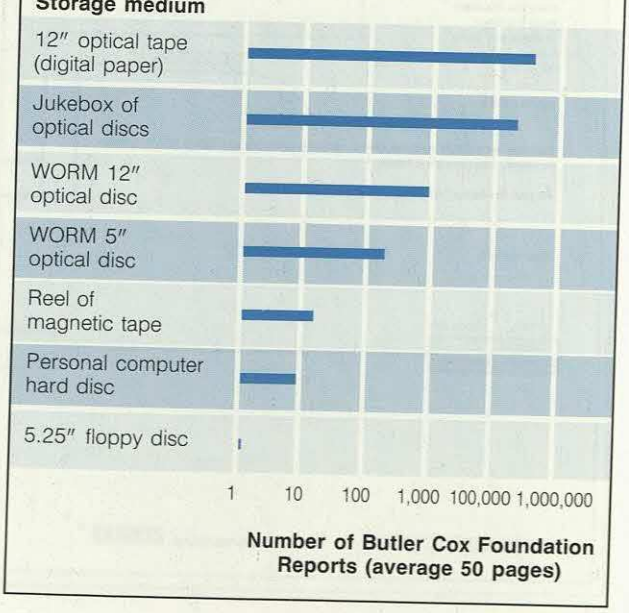


Figure 3.9 The optical-disc jukebox provides massive storage capacity

	Number of discs	Number of drives	Capacity (gigabytes)	Capacity (million pages)	Average retrieval time (secs.)
Configuration 1	64	4	166	3	15
Configuration 2	204	4	530	10	16

The above information relates to FileNet Corporation's jukebox systems, which use 12-inch optical discs, each with a capacity of 2.6 gigabytes. A typical jukebox system might have 64 optical discs, with four disc readers. This would give a total online storage capacity of approximately three million documents. It is also possible to connect more than one jukebox to the same EDM system, thus giving a theoretical maximum of 80 million documents held online. FileNet's largest current installation has online access to 20 million documents. Jukebox response times are determined by the mechanical nature of the operation and the slow retrieval times for optical discs. Average retrieval time for such systems is quoted by suppliers as 16 seconds, although this would deteriorate on larger, multi-user systems. (One user organisation quotes one minute as the average retrieval time for documents that are not prefetched).

(Source: FileNet Corporation)

of online storage. A variety of techniques can be used to get around this problem. For example:

- Where the user can specify, in advance, which document images will be required, these can be prefetched to a high-speed magnetic cache memory. The prefetching of these documents can take place while the user carries out other tasks. To gain the maximum benefit from the use of prefetching techniques, however, users must be prepared to amend their working procedures.

- Clustering algorithms can be used to group documents that are likely to be used in the same working session onto one optical disc, or a selection of discs. This technique is based on the principle that any given document will be associated with several other documents, and that these will always be retrieved together. The associated documents can be stored together on the same optical disc, thereby reducing access times.

Providing easy access

We have seen that EDM systems facilitate the capture, indexing, and storage of a variety of document types. What makes EDM more than just an archiving system is its ability to retrieve and manipulate the document images on a screen in realtime, removing the need to access the original paper record.

Users retrieve document images by accessing the index database, selecting the required document from the results of the index search, retrieving the image file from the optical disc, and displaying it in a window on the screen of their workstation. In multi-user environments, it is also necessary to route the image to the correct workstation.

The advances in technology that have made on-screen manipulation of document images possible are the availability of the high-resolution screens and the development of specialised image-processing hardware components. These components provide the speed and capacity needed to ensure that an acceptable level of operational performance is provided at the workstation level.

Using high-resolution screens

In order to work with document images (instead of working with paper), it is desirable to use a high-resolution display (approximately 1,000 x 1,000 pixels, equivalent to 75 dots per inch). Standard-resolution screens (640 x 480 pixels), such as those used with most personal computers, may be used, but the image legibility may not be acceptable for anything other than intermittent use. Most EDM systems allow standard terminals to be connected to them, so they can be used to access the index database, even if it is not possible to view the actual images. In practice, this means that the user is able to access the index information for a given document, without necessarily being able to view the document image on the same screen. Standard personal-computer workstations can be upgraded, relatively inexpensively, so they can display images. Many suppliers are

recognising the importance of EDM, and personal-computer workstations are now increasingly being provided with high-resolution displays as standard.

Providing image-processing hardware

Large amounts of processing are required each time a document image is retrieved or manipulated — screens have to be redrawn, for example, when moving to the next page of a document, and the data for uncompressed images (in excess of 1 megabyte) must be processed. EDM systems provide similar image-manipulation facilities to those provided by desktop publishing packages — in particular, the ability to scale, rotate, and cut-and-paste all or part of the image — and these facilities also require substantial processing power. Image workstations contain additional hardware to provide the required processing power. This is usually in the form of an add-on circuit board for the standard workstation, containing the additional circuitry necessary to handle digitised images. Typically, this board has:

- A high-speed processor capable of performing the millions of calculations per second required to redraw images on the screen in an acceptable time. It also runs the image-compression algorithms so that this processing does not have to be performed by the workstation's own processor. The system is then free to accept the next user's request, and the flow of work is not affected by the processing effort required to draw or compress the image.
- A large amount of memory, so that uncompressed images may be loaded into memory and manipulated.
- A compression algorithm, which is stored on a programmable read-only memory chip. This converts the image into its compressed storage format.

Annotating documents using 'overlay' files

The text included in document images cannot be edited using word processing software. Some suppliers have attempted to overcome this potential limitation by attaching revisable 'overlay' files to each document. These files are, in effect, electronic notes that can be attached to each document. When viewed with the document image, it looks as if the original document has been edited, although within the system, the initial document image remains unchanged. This feature can be used to append notes or other information to the document image. Examples include the addition of a signature where a document needs signing-off before further processing, or the addition of comments or notes to a particular item of correspondence, recommending action. This information does not normally alter the index entries for the document, although the existence of the overlay files would be indicated within the system. The overlay files can themselves be viewed separately.

Controlling the flow of work

The basic functions of an EDM system are to scan, capture, and store images, and provide easy access to them. The flow of work through the system is controlled by special software, known as workflow software. This software provides the document-management facilities that are an essential part of an EDM system. They are analogous to the rules and procedures governing manual access to, and processing of, paper documents in a clerical system. By managing document access and flow electronically, greater control can be provided more efficiently.

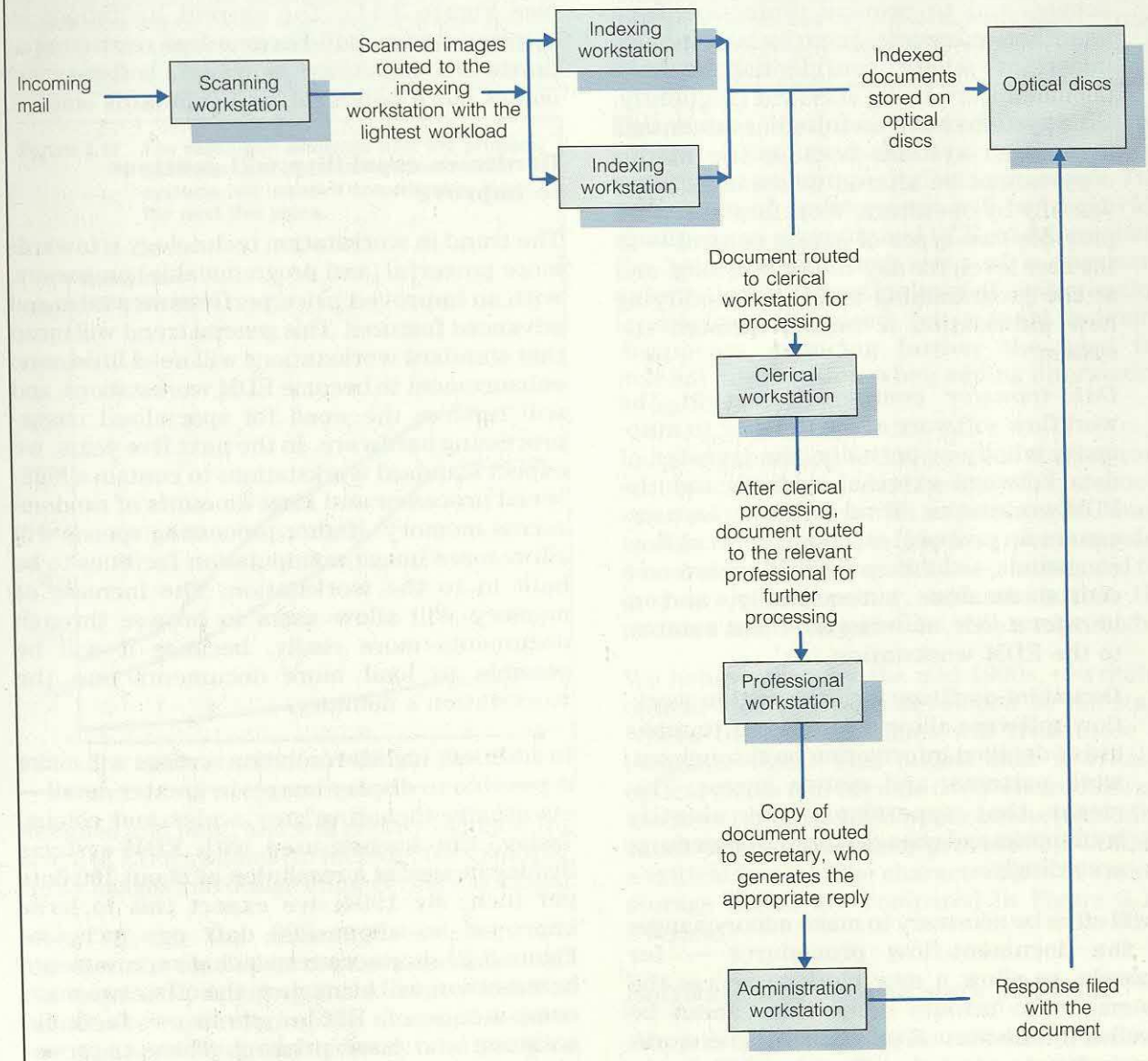
Workflow software enables specific items of information to be presented automatically to the user for action. On completion of the action, the software recognises that the document is ready for the next stage of processing and presents it automatically to the next person in the processing chain. Consider, for example, an EDM system that automates the processing of expense claim forms. The forms are first filled in by individual staff members and, once all the required information is present and valid, they are automatically presented to the relevant manager for authorisation. As each claim form is authorised, it is transferred automatically to the accounting system for payment. The routeing of the document and the triggers that cause it to move between steps are all controlled by the workflow software within EDM.

The main tasks of workflow software are therefore to control the entry, editing, indexing,

storage, accessing, retrieval, display, and manipulation of documents. It also manages the links between the EDM system and other systems, such as data processing. Workflow software provides the following functions:

– *Document routeing* controls the movement of documents between users or workstations. This may be a very basic function, or it may have some in-built logic that routes the document according to its content. For example, workflow software can use a specific piece of information on a document to determine the most appropriate person or workstation to route the document to. Figure 3.10 gives an example of document routeing. This function is critical for the success of many applications because it automates the links between the tasks that have to be carried out at the various stages of processing a document, and enables documents to be transferred between users automatically, thereby increasing productivity in many transaction-processing operations.

Figure 3.10 Document routeing controls the movement of documents between users or workstations



- *Retention schedules* are the means by which users decide how long individual documents will be retained within a system. Retention will vary according to the subject and the document type. Memoranda, for instance, may be discarded once they have been viewed, while client correspondence might be held online for several months during the life of a project. Workflow software takes account of retention schedules by automatically deleting items, or bringing them forward for review, when a specific period of time has elapsed.
- *Access levels* are set within workflow software. These are used to restrict access to certain types of documents to specific people, and to provide protection from unauthorised access. Security is extremely important where confidential business documents are being accessed frequently. EDM systems have an inherent advantage over paper systems because the master copy cannot be altered or damaged accidentally by operators. Workflow software provides two types of access control — at the user level, for day-to-day working, and at the programming level, for modifying how information is routed through the system.
- *Data-transfer commands* within the workflow software allow the user to automate, wholly or partially, the transfer of data between external systems and the EDM workstation. Thus, for example, users can set up procedures, using the workflow commands, to look up a client record on a central database automatically, and to transfer a code, showing the client's status, to the EDM workstation.
- *Document-auditing* facilities within workflow software allow a supervisor to make use of detailed information on throughput, work patterns, and system access. This means that supervisors can identify bottlenecks and change working procedures accordingly.

It will often be necessary to make minor changes to the document-flow procedures — for example, to allow a new clerk to access the system, or to reroute items that cannot be handled by the normal procedures. The workflow software must therefore be sufficiently flexible to allow authorised staff to effect these changes easily, without having to refer back to the supplier or other external resources. Equally important, workflow software should be constructed so that unauthorised staff do not inadvertently alter document flows.

Future developments

As with all new technologies, major changes will occur in EDM technology over the next few years as it matures and comes into widespread use. All technologies progress through a four-stage life cycle, which we define as emerging, pacing, key, and base. EDM is no exception. Currently, EDM is at the key-technology stage (see Figure 3.11). The current limitations on systems design will become less restricting as hardware capability improves, software becomes more powerful, and standards emerge.

Hardware capability will continue to improve

The trend in workstation technology is towards more powerful (and programmable) processors, with an improved price/performance ratio and advanced features. This general trend will mean that standard workstations will need little or no enhancement to become EDM workstations, and will remove the need for specialised image-processing hardware. In the next five years, we expect standard workstations to contain a high-speed processor and large amounts of random-access memory. Higher processing speeds will allow more image manipulation facilities to be built in to the workstation. The increase of memory will allow users to browse through documents more easily, because it will be possible to load more documents into the workstation's memory.

In addition, higher-resolution screens will make it possible to display images in greater detail — eventually including grey scales and colour. Today, the screens used with EDM systems display images at a resolution of about 100 dots per inch. By 1994, we expect this to have improved to about 200 dots per inch. As Figure 3.12 shows, corresponding improvements in resolution will be made in the other two main components of EDM systems — facsimile scanning and laser printing. These improvements will enable more efficient systems to be

Figure 3.11 EDM is following the typical technology life cycle

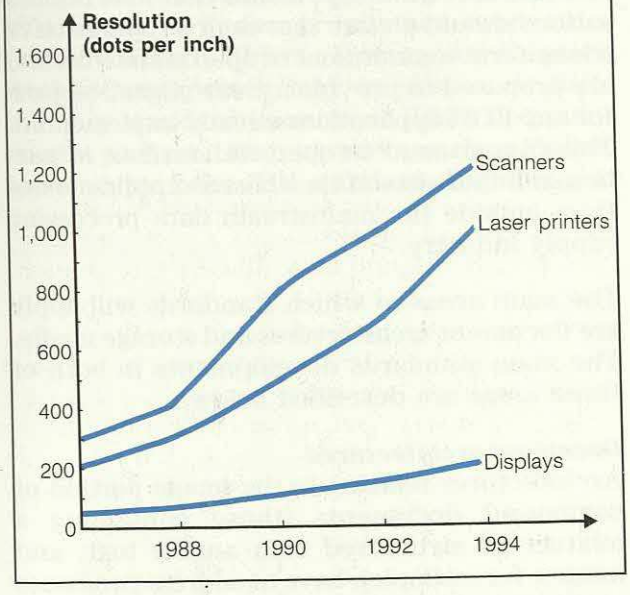
In the first stage, the technology is *emerging*, and is driven forward by pioneeering research departments and small entrepreneurial suppliers. At this stage, the technology is still being developed, although interest in it is high, and its potential is beginning to be recognised. Companies such as Philips, FileNet, and Wang were instrumental in developing EDM systems when it was at the emerging stage. Philips launched the first version of its Megadoc product in 1983; Wang launched its first PC-based image product the same year; FileNet launched its first product in 1984.

In the second stage, the technology is *pacing*. Applications are at the pilot stage, and the first commercial products become available. Established computer manufacturers become aware of the opportunities offered by the technology, and seek joint ventures with those who have developed it. During the pacing stage of EDM, FileNet entered into a joint venture with Olivetti, and its products became the basis for EDM projects for AT&T, Alpharel, 3M, and IBM.

In the third stage, the technology is *key*. User organisations realise that they can gain a competitive advantage by using the technology. EDM is at this stage at present, and it is characterised by the growth in importance of systems integrators. Vertical markets grow rapidly in this stage.

In the fourth stage, the technology is *base* — that is, it becomes incorporated into the standard product offerings of established suppliers. With EDM, standard workstations, capable of image processing, will become available (for example, the NeXT computer from Steve Jobs's new company), as will fully integrated office systems that will make it possible to control and manage data, text, image, and voice. Wang is probably the furthest ahead in this area of integration.

Figure 3.12 The resolution available with the printers, scanners, and displays used with EDM systems will improve considerably over the next five years



designed and built, and will greatly increase the scope of EDM systems (to include, for example, X-ray images and design engineering drawings).

Even though the scanning resolution will increase, the time taken to scan a document will decrease. OCR accuracy is also improving. The trend is towards storing the image of a document in semiconductor memory and carrying out the OCR processing as a background task. This means that the scanned image can be converted into revisable text without interrupting the scanning or indexing processes. The next stage of development will be to provide simultaneous scanning and OCR, to differentiate the text and non-text areas of a document automatically. Once this level of sophistication is reached, the need for manual indexing disappears, reducing further the need for manual intervention when adding information to an EDM system.

Image compression techniques will also improve dramatically from today's norm, which reduces the data required for an uncompressed image by a factor of about 40. However, the increasing demand for higher-resolution scanning, and for grey scale and colour, will mean that the average size of image files will remain high.

We believe that, by the mid-1990s, rewritable optical media will be developed to the stage where their access times will rival those of magnetic discs. This means that rewritable optical media will become the preferred storage medium for most EDM applications (excluding those where the security provided by WORM is a critical factor). The characteristics of various storage media are compared in Figure 3.13, overleaf.

Software will become more powerful

At present, most EDM software functions required by users have to be designed and

Figure 3.13 The prices and performance of current and emerging storage media vary markedly

Storage medium	Typical capacity (gigabytes)	Access time (seconds)	Data transfer rate (million bytes/ second)	Cost of storage per 50 A4 pages ($)
12″ optical tape	1,000.0	28.00	3.00	0.01
WORM (12″)	2.6	0.10	0.25	0.50
5.25″ flexible optical disc	1.0	0.04	1.50	0.10
Rewritable optical disc	0.5	0.06	1.25	Not available
Magnetic disc	0.3	0.02	3.00	33.50
Magnetic tape	0.1	13.00	3.00	0.10

programmed specifically for each application. As experience in EDM grows, more functions will become part of the standard software. To date, developments in workflow software have often taken the form of ad hoc enhancements made as a result of requests from users. This is a less than satisfactory situation, due partly to a lack of understanding by both suppliers and user organisations about the real benefits of implementing EDM, and to the relative immaturity of the technology. As the technology matures, standard off-the-shelf workflow software will emerge to address specific types of application.

Workflow software is already a critical component of EDM systems, and will develop still further. Increasingly, tools will be available to allow systems departments and end-user supervisors to develop and enhance the standard product without having to rely on support from the EDM system supplier.

Standards are emerging

As with any other rapidly developing technology, standards for EDM are at an early stage of development. In general, there are two types of standard — internationally agreed 'open' standards, and those that are proprietary to equipment suppliers. In the field of EDM, some of the proprietary product-based standards have become so well established that user organisations now regard them as *de facto* industry standards. All EDM suppliers are striving to deliver dramatic improvements in technical capability, and they will not let a lack of international standards slow down their progress. The current dominance of proprietary products and standards will therefore be eroded only gradually by emerging international standards.

For user organisations, the lack of open standards becomes a major issue when they wish to upgrade their EDM systems or take advantage of new developments. The present domination of proprietary standards, coupled with the need for user organisations to gain business benefits from EDM technology, means that user organisations should ensure that their suppliers have a long-term commitment to open standards, and are prepared to provide a clear migration path for any EDM applications already implemented. This commitment to open standards is of particular importance if the chosen supplier comes from outside the mainstream data processing supply industry.

The main areas to which standards will apply are document architectures and storage media. The main standards developments in both of these areas are described below.

Document architectures

Architectures relating to the image portion of compound documents (those containing a mixture of structured data and/or text, and images, for example), have to address two issues — the content of the image, and the means for identifying and referencing the image portion of the document. The main information technology suppliers and the International Standards Organisation (ISO) have each established compound-document architectures that are capable of dealing with both of these issues. Although the detailed standards for image-content are not yet fully developed in any of the supplier initiatives, we expect that these

will be based on the well-understood and proven CCITT facsimile standards.

A major component of image-content standards is the compression algorithm, and some suppliers are working on proprietary compression schemes that they believe will give them a competitive advantage. IBM, for example, has a proprietary compression routine that it claims gives a 30 per cent improvement over the CCITT Group IV facsimile formula. User organisations must weigh the advantage of improved compression techniques against the risks of using a non-standard technique.

Standards for identifying and referencing the image portion of a compound document are better defined. The ISO's Office Document Architecture standard, which conforms with the OSI model, covers the structure, interchange, and eventually, the contents of compound documents. The suppliers fall into two broad camps — those such as DEC, ICL, and Wang, which have developed their own proprietary architectures, but which are pledged to migrate to international standards as they mature, and IBM, which has its own proprietary architecture (called Mixed Object Document Content Architecture) which may become a *de facto* industry standard because of IBM's dominant market position. Other suppliers, such as Philips, have adopted the IBM standard. IBM has also indicated that it will provide gateways from its proprietary architecture into systems conforming with open international standards.

Storage media

The main standards issues involved in choosing optical media are disc size and format. Three WORM disc sizes are in common use — 5″, 12″, and 14″. Moreover, different disc suppliers use different disc formats. User organisations therefore experience problems if they want to use the same optical disc on different systems. This can be a major concern if disaster-recovery plans require the use of other organisations' EDM systems. The suppliers have recognised this problem, and there is now a general move to standardise on the Hitachi 12″ disc format, known as H ITOD301A. This is emerging as a *de facto* standard for optical discs.

In this chapter, we have shown that EDM technology is still developing, and any investment in an emerging technology carries with it an element of risk. Would it therefore be advisable for Foundation members to postpone decisions about whether to invest in EDM until the technology is better established and prices begin to fall? We believe not, and in the next chapter, we show how those who have already invested in EDM have gained the benefits they sought, despite the limitations of the current technology.

Chapter 4

The current limitations can be overcome

Our research shows that although Foundation members recognise the nature of the benefits that EDM can deliver, many of them are resisting introducing EDM applications because of doubts about technical performance, concerns about the lack of appropriate technical skills, a lack of confidence in a concept that is still quite novel, worries about the cost and effort required to convert existing records, and concerns about the legal implications of EDM.

In our survey, the results of which are shown in Figure 4.1, members who had not yet introduced EDM systems expressed concern about precisely the same areas in which those involved in implementing systems had encountered problems. In no case, however, had these problems caused an organisation to abandon its EDM development once it had embarked on a worthwhile application. This implies that the problems that do arise can be dealt with successfully. In this chapter, we provide advice to members on how to overcome the problems that they may encounter so that the undoubted benefits of EDM may be fully realised.

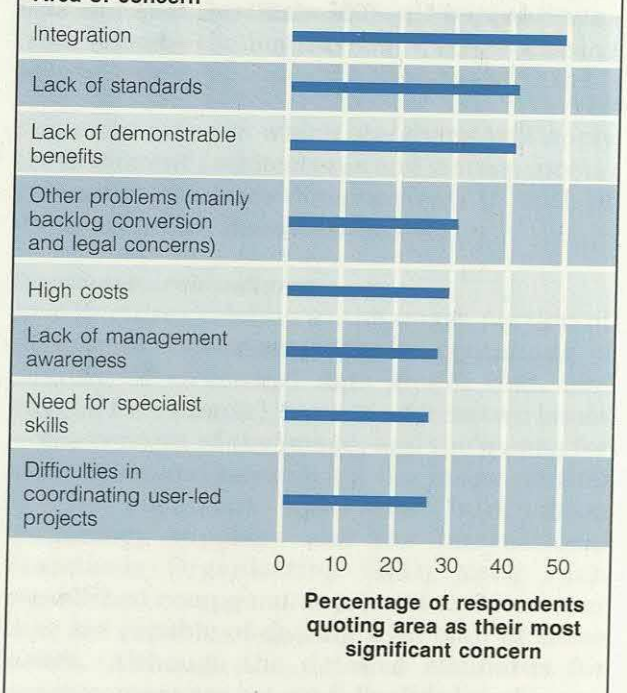
Technical limitations influence the design of applications

We showed in Chapter 3 that EDM is a developing technology, and as such, is changing very quickly. It is therefore inevitable that technical problems will arise, but by being selective about what they try to achieve, and by designing systems to exploit the current capabilities of the technology, EDM users can minimise the effects of the problems.

The most successful EDM applications are those requiring links with other systems, and those involving workgroups or entire departments, rather than individual users. This means that integration and networking are both vital considerations in planning and building an EDM application. There are, however, still limitations in the technological capability of EDM systems in each of these areas. User organisations that are prepared to be selective — integrating only where necessary, and designing networks very carefully — will nevertheless realise significant benefits from their investment.

Other technical limitations affecting design stem from the nature of EDM technology, as described in Chapter 3. Document image processing has its limitations, because of the size and unwieldiness of images, and the sheer capacity

Figure 4.1 Integrating EDM with existing systems is Foundation members' greatest concern



(Source: Survey of Foundation members)

that is needed to deal with them. Indexing and workflow impose a new discipline on users. There are limitations on the extent to which images can be manipulated. There are also some restrictions inherent in optical disc media and their functioning within an EDM system. We now describe the steps that systems designers can take to minimise the impact of these current limitations of EDM technology.

Integration problems

There are essentially two types of EDM integration. The first is that between EDM and other computer-based systems. The second is between different EDM systems and applications.

Integration between EDM and other systems

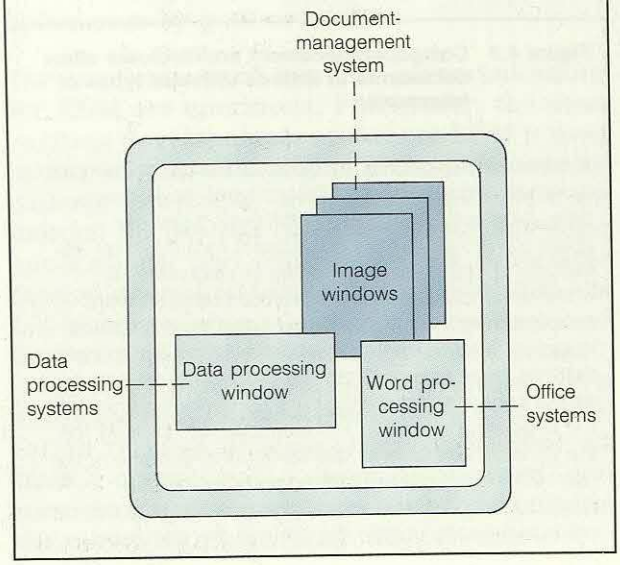
Integrating EDM with existing systems means providing facilities for users to capture, store, and retrieve image files in conjunction with the normal data and text components of a data processing or office system. This type of integration is vital for almost all EDM applications – first, because organisations want to safeguard their original investment in information technology, and second, because users of EDM applications inevitably require access to data or information from other systems. As the newest component of EDM technology, and the one for which standards have yet to be established, document image processing is the most difficult to integrate with existing systems.

Document image processing can be integrated with other systems in three main ways. The first (and most basic) type of integration provides access both to document images and to an existing database or office system. For example, a credit-card company provides operators with access to images of actual credit slips. The EDM system is linked to the existing client database, and the required slip can be identified by accessing the client's database record. Suppliers provide this type of integration by building links between document image processing and other systems, usually via the workflow software. Because the image system is an add-on facility to an existing system, it may be based on a different database system from the original system, and this can make close integration difficult. Where a new system is being implemented, or where the existing supplier can provide a fully integrated solution, database incompatability will be less of a problem. Of the main computing and office systems suppliers, only Wang can provide this capability today.

The second way of integrating document image processing with other systems is to allow users to view side by side, at a screen, document images together with word processed documents, and the results of database enquiries, or other applications. For example, a clerk may wish to view the image of a handwritten letter from a customer alongside the results of an enquiry into the customer's account status, while drafting a reply. The usual method for providing this ability is via a multi-tasking operating system and a windows-based user interface. Figure 4.2 shows how such an interface appears to the user. Windows-based interfaces are now widely available with high-performance workstations. However, providing such interfaces (and image capability) for dumb terminals is more difficult, and organisations may therefore need to invest in new workstations.

The third, and most advanced, form of providing integration between EDM and other systems is to use image as just one of the information types in a document-processing system. For example, a scientist may need to include a scanned diagram or photograph in a word processed

Figure 4.2 EDM systems use a windows interface to allow interaction with other systems and applications



report, send it electronically to a colleague for review, amend it, and print it via a desktop electronic publishing system. Suppliers are developing compound-document architectures to handle this type of mixed document (see Figure 4.3 for a description). Examples include Digital's Compound Document Architecture, IBM's Mixed Object Document Content Architecture, and the international standard, Office Document Architecture. In addition, many suppliers are developing the convertors and interfaces necessary for document images to be captured, processed, distributed, stored, retrieved, and printed alongside text, data, and graphics. Systems and applications based on these mixed environments are very powerful. They are also complex and require careful management.

Integration at the basic level (between existing systems and a new document image processing system) is essential if the hoped-for benefits of EDM are to be gained. Organisations should therefore select a supplier who can provide integration at least at this level; the more advanced levels of integration will not be a critical requirement for most organisations at this stage of development of EDM.

Integration between different EDM systems

The second type of integration, that between separate EDM applications, was not perceived by most organisations as a requirement for transaction applications. Usually, the document images in one EDM application are of little importance in other business areas. For example, an insurance company can determine that a customer has domestic, health, and motor insurance policies with the company by providing access to the relevant databases. However, the occasions on which the handwritten details recorded on a motor-insurance claim form will be of value to the domestic insurance underwriters are rare. In such an organisation, integration between the individual EDM applications used for each type of insurance and the central data processing systems is therefore critical, but integration between the EDM applications is not.

Nevertheless, those organisations with a requirement for several EDM transaction applications should establish a corporate standard for EDM, simply for reasons of economy. Different types of EDM system, and different indexing and workflow design, all require the acquisition of specialist skills, and it makes sense (and costs less) to maintain one set of skills rather than several. Furthermore, there could be unanticipated requirements to move information between EDM applications.

For reference applications, the need to integrate different EDM systems is often higher than it is for transaction applications. It may be necessary, for example, to send a document retrieved in one department or division to another. This requires that the two departments' EDM systems are compatible and that there is a document-transmission link between them. Where documents are routinely passed between departments, common layout and reference conventions will also need to be defined so that users can readily understand how documents have been formatted, and how documents from different sources have been amalgamated. This is important in any environment where complex documents need to be shared between several groups of users — for example, in research, publishing, marketing, corporate finance, consultancy, and many other business activities.

Figure 4.3 Compound-document architectures allow documents to include different types of information

A compound-document architecture allows documents to include information of different types (text, image, data, graphics). Although the different types are physically stored in their constitutent formats, they are logically stored as a single entity or document. A complete document could therefore comprise word processed text, a pie chart generated by a graphics package, a table from a spreadsheet, and a photograph that has been scanned into the system. Some suppliers aim to enhance this architecture by providing links into the originating software. For example, by 'clicking' on the spreadsheet table in the compound document, the user would be transferred transparently into the spreadsheet program, where changes made to the data will automatically update the table in the document.

Networking-capacity problems

Even after image files have been compressed, they are still very large, and transmitting them over a network therefore takes a considerable time. It is, however, possible to plan and manage local-network capacity so that images can be handled. One method is to limit the number of

network users who may want to transmit image information at the same time. The number of concurrent users who can be supported by a local area network is limited by the bandwidth of the network. Ethernets and 4 megabit Token Rings, the two most common networks used with EDM systems, can both handle between 10 and 30 active image terminals. Higher-capacity networks, such as those based on the ANSI FDDI (Fibre Distributed Data Interface) standard, are being built into many offices today, and these will allow a greater number of EDM users to use the network concurrently.

Because of the size of image files, sharing a network between image and normal communications traffic may result in poor performance for the non-image communications. As a consequence, many major organisations now have separate networks for image and data communications. Given today's network capacity limitations, some form of local high-speed network dedicated to image communications is often the most acceptable solution for an EDM system.

Wide-area networking for EDM systems poses problems similar in nature to those encountered with local networking, but on a larger scale and with greater capacity constraints. These problems have been tackled either by using high-capacity public networks, or by using high-speed private networks. For example, Glaxo Inc, whose EDM application was described in Chapter 2, has installed a private, optical-fibre network to link several of its buildings. One alternative to investing in a high-capacity network is to provide EDM hardware at each site and to transmit batches of images between them in off-peak hours. To be effective, however, this approach requires knowledge about the documents that are likely to be required at each site during the next working day.

Document-capture speed

Today's fastest scanners can operate at sub-second speeds per page. However, the speed at which the scanned image can be accepted by the image processor, compressed, and displayed on a screen ready for indexing places a limit on the speed at which images may be processed. Designers need to configure EDM systems to take account of this, and to consider providing one or more dedicated capture stations, thus leaving other workstations free for retrieval and processing. Indexing may then be done as a batch operation — that is, the documents may be scanned, uninterrupted, into the system, and indexed in a batch at a later stage. One dedicated document-capture station can usually cope with up to 500 pages per day, depending on the complexity of the indexing involved.

Another factor to consider is who in the organisation should be responsible for document capture and indexing, and where the facilities should be located. For example, it may make sense to provide a single, high-capacity, document-capture facility in the mailroom rather than allowing each department to scan its own documents.

An alternative to providing dedicated document-capture stations is to use a bureau service. A growing number of bureaux now offer document-scanning and indexing services, although costs remain prohibitively high for many applications (typical charges are between 20 and 40 cents per page). Use of bureau services may be justified, however, if there is a large, once-off requirement for high-volume scanning (to transfer existing paper files to an EDM system, for example).

It may also be necessary to consider document capture from alternative sources, as new methods of document distribution come into widespread use. The use of facsimile-compression standards in EDM and the availability of inexpensive desk-top facsimile terminals open up the possibility of direct facsimile communication to and from EDM systems. The dissemination of hard copy by facsimile compares very favourably (in terms of cost and speed) with sending the equivalent documents by post or courier.

The implications of facsimile communications for EDM are enormous. Potentially, the most exciting developments are where EDM is used in conjunction with mobile facsimile communications. The ability to send and receive commercial documents (such as orders, invoices, applications, and contracts) from a portable facsimile terminal to a central EDM system has far-reaching consequences for the provision of time-critical services, both public and private. Consider, for example, the implication for a mortgage broker, who could fax a new application through to head office from home, or from a client's house, have it processed, and receive a facsimile contract back for signing by the client, all in the same day.

Image-retrieval times

Retrieving images from an optical disc is slow compared with the times required for retrieving data from a magnetic disc. Retrieval times of 15 seconds per item are not uncommon, although some systems enable the user to browse through a multi-page document at two-to-three seconds for each page, once the images for the whole document have been retrieved. Although such retrieval times are much faster than retrieving paper files from a central filing system, they are generally not acceptable in transaction applications where a clerk is using a screen-based system to process the information.

The impact of these slower retrieval times can be minimised by designing the EDM system so that the clerk can specify the group of records that will be required. All of the records can then be retrieved from the optical storage and loaded onto a magnetic disc before detailed processing of individual records begins. Prefetching the records in this way can be done several times a day, or even as an end-of-day process, if the next day's work is predictable. Where the pattern of document access is predictable, it may be possible to automate the prefetching process. It is therefore important to study document-access patterns so that the prefetching rules can be designed to minimise image-retrieval times.

Another way to minimise retrieval times is by designing the EDM system so that related documents are clustered together on the optical discs. To do this effectively, however, requires highly specialised skills.

Indexing effort

Good indexing of the information stored in an EDM system will also help to minimise retrieval times. Conversely, a poor index will inevitably slow the retrieval process down. Because of the size of document images, the time taken to retrieve each image is significant. Thus, the index should be constructed so that individual documents can be identified, not just the batch of documents that contains the required information. Precise indexing for EDM systems is more important than for paper-based systems, because the equivalent of scanning through a pile of papers to find the required document — browsing at a screen — is slow on image systems.

However, the effort required to create a good index for an EDM system imposes an extra workload. The experience of the organisation described in Figure 4.4 illustrates that the effort required can mean that an EDM system cannot be cost-justified.

Automatic indexing techniques (which were described in Chapter 3) can help to reduce the indexing effort. At present, the techniques are used mainly by document-capture bureaux. We expect them to be used more widely as developments in OCR technology occur. Fast, reliable, automatic generation of indexes provides users with a further opportunity to streamline their document-management working practices.

Workstation requirements

In designing EDM systems, it is important to ensure that each user is given the right type of workstation, with access through it to appropriate facilities. This affects workstation policy across the organisation, and the introduction of

Figure 4.4 The effort required to create an index may be too great to justify EDM

A company in the media sector

One of Butler Cox's consultancy clients (a company in the media sector) was considering whether to invest in an EDM system to replace an existing press-cuttings library. The library had a staff of eight. The manual filing system was folder-based, with a simple index structure (each folder was relevant to a particular subject, and all cuttings on that subject were filed in that folder). Within the folders, the cuttings were more or less in chronological order. Each folder contained an average of 1,000 items. There were approximately one million cuttings in the library. When a request for information was received (on the Chelsea Flower Show, for example), one of the librarians would pull out the folder on exhibitions and browse through the contents. Browsing manually in this fashion was quite quick (usually less than one minute for an entire folder), even though the librarians had only a very general idea of what it was they were searching for (anything to do with the Chelsea Flower Show).

This company was not prepared to spend the time and effort required to create a new index structure, because the existing one worked well and everybody was familiar with it. However, without a more sophisticated index, it was impossible to cost-justify the proposed EDM system. Using the existing indexing system with EDM would have meant investing in very-high-performance equipment to achieve the required response times.

EDM may require a widespread review of workstation requirements. The choice of EDM workstation, and associated software, for each type of staff will depend on the answers to two key questions. Do staff need to access document images, or merely the index database? Do they need to revise images, or merely to view them?

Those users who are viewing (and perhaps processing) document images at their screens, as distinct from merely accessing the index database, require high-resolution display screens. In many instances, however, users do not need to view the document, but only require access to the index information held about that document. For example, a customer-enquiries clerk in an insurance company might want to browse through a client correspondence-tracking file, where the existence of a particular policy document (denoted by an index entry) will be sufficient information to allow the query to be processed. However, claims-processing staff will need access to the full contents of several documents to process each claim.

The degree to which different staff need to revise document images varies, and this affects the type of workstation that is required and the facilities that need to be accessed through it. For instance:

- In many EDM applications, the variable portions of a document are stored as data in a database, with the original document stored as an unrevisable image. This might occur, for example, where an insurance claims-processing department stores a client's transaction details in a central database, but an image of the claim form is stored on an optical disc. In this situation, users of the EDM system would not need to revise the images.
- In other applications, EDM users may need cut-and-paste facilities that allow them to transfer one document to another for further processing. In this situation, EDM users will require additional workstation facilities. At the simplest level, this might require the ability to transfer part of a document to a desktop publishing or drawing package. During our research, we encountered a banking application where users electronically cut and paste copies of scanned signatures into a desktop publishing system, which arranges them into a standard book of authorised signatures. Clearly, this type of application has to be designed to ensure the tightest security and control over who can use the workstations for this purpose.
- In some advanced applications, users may require the ability to 'vectorise' an image of a drawing (that is, transfer an image into a revisable vector-graphics format so that it can be modified by a graphics package). For example, in an oil-platform construction project, existing technical drawings are scanned and vectorised so they can be updated by an automated drawing system.

Optical-storage limitations

The use of optical discs as a storage medium for images has design implications for EDM systems. The most significant considerations are to provide the right capacity, to ensure that the operational performance is adequate, and to take account of the unrevisable nature of optical discs.

The online storage capacity is determined by the number of documents to be stored, the frequency with which they will be accessed, and the length of their active life. Beyond a certain level of online optical-disc storage, there is a sharp increase in the costs of storage, because of the high cost of an optical-disc jukebox. There can also be significant performance constraints when using a jukebox system, because of the time required to load the appropriate disc into a reader. It is therefore necessary to study closely the proposed document-flow procedures, to ensure that the retrieval times are minimised wherever possible.

Another limitation of the current generation of optical discs is that the images stored on them cannot be revised. This limitation can, however, sometimes turn out to be an advantage. Many EDM applications can be justified *because* of the unrevisable nature of documents stored on optical discs (for example, legal contracts and other documents). Other EDM applications can be considered only if the images can be revised, and it will be necessary to use a rewritable medium (optical or magnetic) in these cases.

Other concerns about optical discs expressed by potential EDM users in our interview programme related to backup and disaster recovery

planning. One problem is that the slow data-transfer speed means that creating backup copies of optical discs is a very time-consuming process. The solution adopted by many organisations is to record each item in duplicate as it is entered into the EDM system. In this way, the backup copies are created automatically during normal processing.

The problem of disaster recovery planning is, at present, less easy to solve, although it should disappear as EDM becomes more commonplace, and as standards evolve. The difficulty today is finding alternative sites with document-image processing equipment that could provide alternative processing facilities in the event of a total system failure. This can be crucial if EDM forms part of a transaction-processing environment, but is difficult to achieve today because different EDM systems operate to different standards.

Technical skills are in short supply

A full EDM system requires new links between existing and new technologies, and involves equipment with which the systems department may not be familiar. The skills required to make the best use of EDM technology and to overcome the technical limitations, discussed earlier in this chapter, may not always be available from within the systems department. The lack of relevant technical skills can slow down the implementation of desirable EDM systems.

A further complication arises because the technical design skills required for conventional data processing applications are too narrow for the design of EDM systems. Technical systems designers usually do not have experience of analysing and defining office procedures and the flow of documents between different groups of staff. The benefits of introducing an EDM application will often come as much from reorganising the flow of work between individual people, and using workflow software to introduce more efficient procedures. Careful design of the flow of work is also often a factor in overcoming some of the technical limitations of EDM. Many systems departments do not have staff with the skills necessary to carry out these tasks.

The required skills are in the traditional area of organisation and methods (O&M), which is concerned with measuring and defining working procedures. In some countries, particularly in the United Kingdom, O&M skills have not been absorbed into the systems department. In many UK organisations, the systems department has taken on the responsibility for other elements of management services, and the traditional O&M practitioner no longer exists. In other countries, the continuing need for such skills is recognised, and the O&M role complements that of the systems function. This situation is particularly evident in Germany, where most companies have an 'Organisation Division' run by a board director. This division combines, and gives equivalent status to, the responsibilities for organisation structure, workflow, and information technology. Organisations with this type of structure are extremely well-positioned to exploit the opportunities provided by EDM.

The concept of EDM is still novel

Another difficulty of designing EDM systems is that the concept is still novel to the majority of systems managers. Furthermore, because the number of operational systems is still small, the opportunity for learning from the experiences of others is limited. Senior business managers have not therefore been able to see EDM systems working in organisations similar to their own, and find it difficult to translate the benefits gained from EDM by other businesses to their own situation. Interestingly, once one or two pioneers in a particular industry sector have implemented successful systems and the benefits become obvious, business managers are quick to realise the potential of EDM technology. This is evident from the 'clustering' of EDM applications in certain industry sectors, rather than in organisations of a certain size, or in application areas chosen by other criteria that are often more relevant to traditional data processing systems.

Within systems departments, management is also often not aware of the potential of EDM. Typically, systems management takes a narrow view of 'information', and sees paper-based records as being outside its area of responsibility. This situation is similar to that in the early days of word processing; many systems managers did not consider text processing as

part of their brief, and word processing systems were installed by departmental or office managers. The need to integrate word processing with other types of end-user computing has now been recognised, but problems have often been encountered in making the office systems an integral element of the overall systems strategy. Allowing EDM to develop in a similarly fragmented way must be avoided if major long-term benefits are to be gained. We discuss this issue in more detail in Chapter 5.

Business managers need to make the case for installing an EDM system to the senior management of their organisations, because of the high investment required. This, too, is a difficult task because senior-management awareness of the potential benefits of EDM is also limited and because those presenting the case are not always in the best position to identify the full range of benefits that will accrue from an EDM application.

It can be difficult to identify the benefits if a narrow technical view is taken of EDM. Benefits come from streamlining the clerical and administrative procedures supported by EDM, and from the use of workflow software to manage the flow of information. If existing departmental or section boundaries and tasks are not changed, many of the benefits of EDM may not be gained. For example, a major bank, having installed EDM and gained benefits from it, now realises that its document-handling procedures can be speeded up even further by including the post room in the EDM operations. It currently takes up to half a day for incoming mail to be sorted by the supervisor, delivered to the correct department, and scanned prior to inputting it to the EDM system. By scanning the mail in the post room, instead of in the departments, the bank would be able to eliminate two manual processes.

Converting existing documents can be a problem

In many cases, EDM systems will not become really effective until all (or almost all) of the existing paper documents have been scanned and stored in the system. Many Foundation members therefore perceived the difficulty of converting existing documents as a major stumbling block to the implementation of EDM.

It is very important to take into account the effort, cost, and elapsed time required to convert existing records to EDM when compiling the cost-benefit case and developing the implementation plan.

Unless the cost case is very finely balanced, however, the cost and effort required to convert existing documents should not negate the overall case for proceeding with an otherwise beneficial application. Organisations already using EDM have minimised the conversion problems, and hence the cost of conversion, in several ways. Depending on the type of information being handled, three broad approaches can be taken — ignore the existing documents, convert them only when it is essential, or convert all of them as part the development project. Each of these approaches is discussed below.

Ignore the existing documents

Where documents have a short life (for example, credit-card vouchers), organisations have often decided to ignore the existing documents for the purposes of the EDM system. Queries that require access to documents created prior to the implementation of the EDM system are handled via the old, clerical procedures. This means that the full benefits of EDM are not gained immediately, but conversely, no costs are incurred in converting the existing documents.

Convert only when it is essential

Where documents have a longer life, but where only limited access is required after the initial processing (for example, mortgage applications), some organisations are adopting a policy of converting documents only when they are needed for EDM processing. When a client contacts the company with a query, there will typically be a period of increased activity during which several accesses will be made to the client's file. If the file is not already stored in the EDM system, it is converted at this stage. Once a predetermined percentage of the files has been converted by this ad hoc procedure, the remainder of the files will all be converted at the same time. In this way, the organisation gains the immediate advantages of EDM without having to undertake a large initial conversion exercise.

Convert all existing documents as part of the development project

In some EDM applications (research environments, for example), users need all the information to be available on the EDM system before they can derive major benefits from it. In this situation, all of the documents have to be converted to EDM as part of the development project. The scanning of the documents can take place while the full system is being developed, but the indexing task cannot begin until the index structures have been fully defined. The elapsed time required to scan and index a large number of documents should not be underestimated, and should be taken account of in the EDM implementation plans. The danger is that the development work will be complete, but the system cannot be implemented because the conversion of the existing documents has not yet been completed.

There are legal uncertainties

The move from paper records to electronic media raises legal questions — what will be regarded as adequate evidence of title and veracity of information in a court of law? Is it legal to scan research publications and hold them electronically? The legal status of electronic documents was a significant concern of those Foundation members who replied to the questionnaire, and of those participating in the focus group. Because of this, we sought expert advice on these questions. We were advised that, in most cases, legal concerns should not hold EDM applications back, provided that systems are run and maintained in a secure and controlled manner, and that the legal issues applying in each case are identified and planned for. Organisations need to be aware of three particular areas in which legal issues may arise with the use of EDM — best-evidence rules, copyright, and data protection.

Best-evidence rules

One area of concern is whether an electronic copy of a document can be used as evidence in a court of law. The problem arises because most countries have not updated their statutes and common law to take account of modern computer-based methods. In many cases, photocopying and microform techniques are not included either. The accepted rule is usually that litigants must provide the best evidence available. Today, a photocopy of the original is often the best evidence that can be produced. In most cases, this is accepted as valid evidence, subject to explanation as to why the original is not available. The status of electronic images as best evidence is expected to be almost identical.

The litigant may be required to prove that the document provided is an accurate reproduction of the original. This means, in practice, that EDM system designers must be able to demonstrate how the information is protected within the system (that is, who can access it and how access is controlled), and what security exists to ensure that it has not been tampered with since it was originally loaded into the system. While EDM systems use WORM technology, this will be comparatively simple to do, because the original image cannot be amended. The introduction of rewritable optical storage will mean that additional checks have to be added to ensure that the image currently stored is an accurate copy of the original.

We understand that it is rare for the outcome of an entire case to be determined by the veracity of an individual document; it is more common for such a document to be used as corroborative evidence. If a document is challenged, counter-evidence, equally well substantiated, must be provided. One area in which problems may arise relates to handwriting, where forgery of a signature is suspected. A handwriting expert needs the original document to make a proper comparison. However, this type of case is very rare.

Although there are very few precedents, we believe that an EDM document from a well-documented and well-managed system is unlikely, under most circumstances, to be challenged in the courts. Organisations must consider the likely risks when designing systems, and ensure that their implementations have the ability to satisfy legal requirements. If there is any doubt, they should seek professional legal advice.

Copyright

Once information is available electronically, it can be moved between companies and between

countries automatically. Where the information stored in an EDM system has been scanned from published documents, the information is subject to copyright protection. The storage and movement of information therefore has copyright implications, and users must ensure that EDM systems do not infringe the copyright laws of the country or countries in which their systems operate.

Copyright attaches to most written material generated by, or received by, an organisation, and a licence is required to copy it. In many cases – for example, ordinary correspondence – a licence to copy is implicit in the nature of the material. In other cases, an explicit licence may be required. This applies to all published material. The need for a licence may also depend on the use being made of the material. For example, in most countries, a licence is not usually required to use published material for research purposes.

Each copyright agreement usually applies only to staff within a particular organisation. In addition, copyright is territorial, each country enforcing only its own copyright laws. If a licence is obtained in relation to use of copyright in one country, it will not permit use in another country unless the licence has been extended to the latter. Such extension usually requires the payment of a royalty. Care must therefore be exercised when transferring information electronically across national boundaries, even within a single organisation. For example, a company with an EDM-based research library containing technical publications may be able to provide electronic access to its own research staff in one country as part of the original copyright agreement. However, a separate licence agreement will be needed to provide access from the same company's terminals in other countries.

Information publishers are aware of this problem and accept that licences will have to be renegotiated to take account of the wider availability of information, as large, internationally networked EDM systems are introduced. New pricing structures will emerge in response to these changes. To ensure that an EDM system used to store published material does not infringe copyright rules, access to it must be controlled, and special arrangements made with the publisher of the material, if necessary.

Data protection

Most countries have, or will soon have, some form of data-protection legislation covering data on individuals, or 'personal' data. The objective of these laws is to protect individuals against damage that may be caused either through unauthorised disclosure of information about them or through incorrect information about them being held permanently and used. Introducing EDM potentially increases the number of occasions when personal information will be held and retrieved electronically. Experience in Germany and Austria, which were among the first European countries to legislate for data protection, shows that, in this context, no distinction is drawn between magnetic and optical storage. User organisations are now experienced in complying with data protection laws in most countries, and no new problems are presented by EDM.

Developing countries are increasingly conscious of the value of data in general, and of data on their national resources in particular. Organisations operating in these regions, especially in Africa and South America, need to keep all their systems (not just EDM) under review to ensure that local legislation is not infringed.

In this chapter, we have shown that the technical and other limitations of EDM can be overcome by designing systems in an appropriate way. There is no doubt that EDM technology has now reached the stage where it can be used to advantage by a wide range of organisations. How, then, should the implementation and operation of EDM be managed? We address these questions in the next, and final chapter of the report.

Chapter 5

Electronic document management must be managed

EDM technology has great power and potential, but organisations will gain the full benefits of their EDM investments only if they apply the technology correctly and provide adequate, appropriate support. In this chapter, we describe how to set about managing EDM, so that the applications match users' expectations and significant benefits are achieved for the organisation as a whole. The five factors that are critical for the successful management of EDM are listed in Figure 5.1, and described in the rest of this chapter.

Figure 5.1 Five factors are critical for the successful management of EDM

Major activity in managing EDM	Factors critical for success
Defining applications	Rigorously applying selection criteria Reviewing working methods
Designing with EDM in mind	Defining the new workflow Taking the technical issues into account
Choosing the right supplier	Applying selection criteria Understanding the market
Coordinating the use of EDM	Integrating systems Building a partnership between systems departments and users
Making appropriate staff available	Providing the right mix of skills Using external support where needed

Defining applications

Organisations will achieve maximum benefits from EDM only where the information needs and the activities to be handled have been accurately identified and the scope of the application has been clearly defined. EDM, however, is not always the best solution to a particular document-handling problem, and it is important to review the potential applications to see if other solutions may be more appropriate. Once an application has been identified, the existing working procedures must be reviewed, and probably amended, if the technology is to be fully exploited. EDM applications often span the existing departmental boundaries, which may well have been created to facilitate the flow of paper-based information.

Selecting suitable applications

In Chapter 2, we showed that EDM provides most benefit when it is used to handle certain types of information. In introducing EDM, managers must ensure the applications under consideration meet the criteria that we defined there — that is, that the information should be live, as opposed to archival, that it should be unstructured, and therefore not suitable for conventional data processing, that it should be of high value in terms of operating efficiency or decision-making, and that access to it will be frequent and critical.

Applications will usually meet these criteria where the following situations exist:

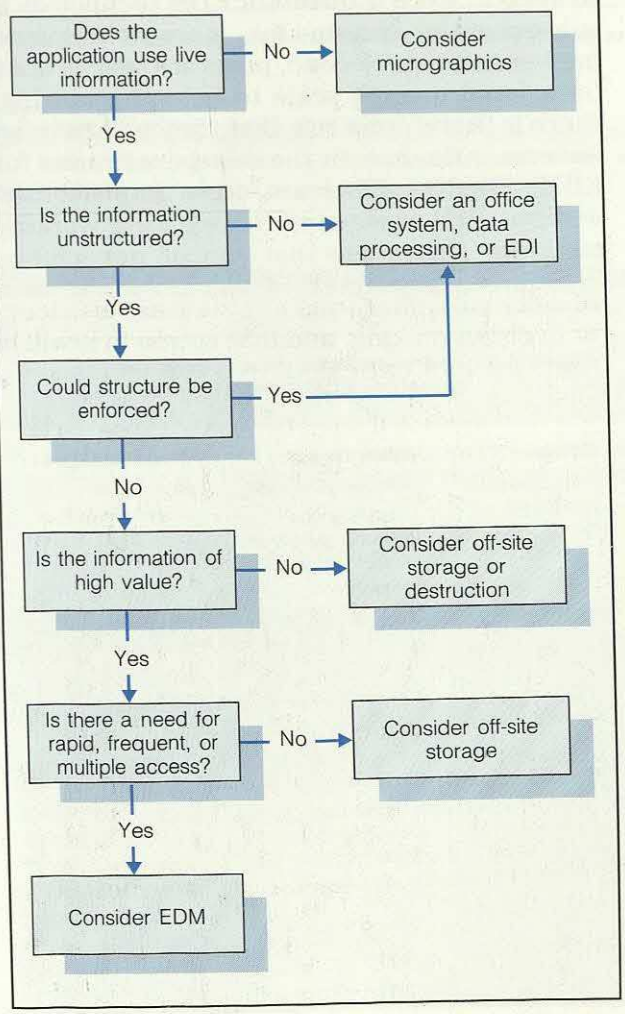
- There are large volumes of paper in frequently used filing systems.
- There are bottlenecks in the flow of paper, causing delay or difficulty in providing a service.
- A high proportion of staff time is spent finding or handling paper documents.
- Staff have difficulty finding the right documents.
- Effort is duplicated because one department is not aware of what other departments have done.
- Decisions are made on the basis of incomplete information.

– A lack of filing space leads to disorganised, cramped, and hard-to-manage files.

If more than one or two of these situations applies, a suitable EDM application may exist. To confirm that EDM (rather than other solutions) is appropriate, the criteria should be applied, in the ways shown in Figure 5.2. However, a situation that at first sight appears to be a candidate for EDM may be better addressed by other methods if the organisation is prepared to modify its working methods and the way in which it deals with its customers.

The experience of a large electrical wholesaler provides a good example of a situation where EDM was not the best solution. Analysis of its sales showed that over 40 per cent of the company's business was with a small number of major retailers, while the rest was with a larger number of comparatively small retailers. Originally, orders were received on the retailers' order forms. The information was therefore unstructured, and the application might be considered as a candidate for EDM. However, the wholesaler reached an agreement with the major retailers to introduce electronic document interchange facilities, and persuaded the majority of the smaller retailers to use a standard order form. This made it possible to process the orders from the smaller retailers by OCR techniques and improved the speed with which their orders were met. What initially looked like an EDM application was better addressed by other processing methods.

Figure 5.2 EDM may not alway be the answer to too much paper



Reviewing working methods

If the full potential of EDM technology is to be realised, system designers must seek to understand what results the organisation is trying to achieve, rather than what tasks it is seeking to automate. To do this, they need to carry out a thorough assessment of the organisation's document-processing requirements, and to review the established working methods and procedures so they can establish whether the case for investing in a potential EDM application can be made.

The organisational boundaries that are appropriate for a mixture of clerical procedures and data processing systems may not be appropriate when EDM is introduced. A good example is the bank we referred to in Chapter 4, which is considering moving its EDM scanning function into the post room. By capturing document images at the earliest possible moment, additional benefits will accrue, in terms of speed of handling, improved control, and reduced support-staff effort. To realise such opportunities, EDM system designers must be free to consider new organisational possibilities when they examine information needs and the activities undertaken.

Some form of work measurement is a vital component of the review of working methods, and is essential in identifying the potential benefits of EDM. Designers must therefore measure the mix and volumes of information processed, and the peaks and troughs of the

workflow. This implies using conventional O&M techniques, and applying them in the context of what EDM can do. When considering transaction applications, it is important to look for bottlenecks in the flow, for labour-intensive activities, and for items that are accessed particularly frequently. The EDM workflow procedures can then be designed to take account of these factors. It is also important to identify the ways in which documents are usually referenced, so that the indexes and the links to data processing systems are effective. The staff who will eventually use the EDM system should be involved in this process to ensure that the proposed new workflow is realistic, and that they will readily accept the new system.

For reference applications, slightly different measurements are needed. Here, the important facts relate to the pattern of reference to documents, and the time spent by professional staff and managers in searching for documents. Document flows may also be very important in situations where a reference system is used to support processes such as report production and publication. This applies, for example, in pharmaceutical research, where documentation accumulates for many years before a selection of it is used to support a regulatory submission.

Designing with EDM in mind

The critical design activities for ensuring successful EDM applications are to define the right document-management and workflow procedures, and to take account of the nature of EDM and its technical limitations. These two activities are, in fact, closely related as new working procedures need to be based on technical considerations.

Defining the document-management and workflow procedures

The information gathered in the review of the existing working methods will be a vital input to the design of the EDM workflow procedures. However, because it is not always possible to predict the precise impact of new technology and working methods, it can be helpful to model the new working methods. There are several ways in which this can be done. They vary in the cost and effort involved, and in the results that can be expected, but they all involve the eventual users of the EDM system. The main options and their respective advantages and disadvantages are summarised in Figure 5.3. Essentially, they are:

- Prepare a paper-based model of the new working procedures and have groups of users test it in one or more workshops.
- Prepare a prototype EDM design, with screen layouts and test data, and have users try it and give their feedback.
- Run an EDM pilot application in part of the user community, and evaluate the results so that the working procedures can be refined if necessary.

Many organisations choose the most expensive option — running an EDM pilot application. Although pilots are a good way to learn about the technology, they also have some drawbacks. First, if users like the system, they will want to keep it, even if different EDM technology is subsequently chosen for more widespread implementation. Second, pilots are seldom done on a large enough scale to be self-financing; there is therefore a risk that they will have an adverse influence on the cost-benefit case for EDM. Finally, pilots are often implemented without first reviewing the existing working methods. This means that they do not achieve their primary purpose of validating the new

Figure 5.3 There are three main options for modelling EDM workflow procedures

Option	Advantages	Disadvantages
Paper-based modelling	Relatively low-cost Users can be involved before any investment in EDM is made	Not a true representation of what EDM is like
Prototyping	A good way of introducing users to EDM Will identify most problems	Used for simulated, not live, processing May not identify every problem
Conducting a pilot application	If done well, the most thorough test A good way of acquiring skills	The most expensive option If successful, users will want to keep it Small-scale, and therefore unlikely to be self-financing Often results in less-than-optimal workflow design

working methods. Our advice to Foundation members considering pilot EDM applications is to treat them with caution, and to be clear about precisely what they hope to obtain from them.

Taking account of the nature of EDM

The main factors determining whether or not an EDM system is successful are different from those associated with conventional data processing systems. For an EDM system, the important factors concern the flow and interaction of information, while in a data processing system, it is the manipulation of the individual items of information that has to be analysed. EDM system designers must therefore address and resolve issues that are very different from those confronted by data processing system designers.

In the previous chapter, we discussed the technical limitations of EDM today and the steps that can be taken to minimise their impact. In Figure 5.4 we summarise the types of problem that careful design can minimise. The examples shown in the figure are illustrative rather than exhaustive; they show how the limitations of the technology affect the way the system is used. The essential point is that users may need to reorganise the way they work, not only to improve operating efficiency, but also to take account of the limitations of today's technology. Consider, for example, the first item listed in Figure 5.4 — clustering. When users access a specific reference, will they wish to retrieve all the items under the reference? If so, what will the impact on their work be if the items are spread across several optical discs, therefore giving poor response times? What can the designer do to reduce the problem? There are three potential courses of action:

— Space can be left on one disc so that all the associated items are stored together.

Figure 5.4 Technical problems can be overcome if they are taken into account at the design stage

System function	Potential problem	Design considerations
Clustering (leaving spare space on discs for later additional information)	How do you avoid delay in accessing associated records added to the system over a period of time?	Do you look at all records every time? Which records are needed together? Over what time is the record group accumulated? Can you predict which record the user will need next, and initiate a prefetch?
Taking archive copies of the updated optical discs	How do you duplicate your master discs in minimum downtime?	Are you going to update multiple discs because of clustering? Can you do double updating without slowing down response times? How would you handle updates to periodic security files?
Accessing information	How do you ensure that users are not slowed down by long retrieval times?	Can you predict the order in which information will be needed? Can you batch work to allow the required records to be prefetched? Can you redesign records to minimise the number that are needed for the mainstream activities?
Scanning and indexing	At what point and how do you add the required references to the records?	What are the work flows and volume peaks? Will it be possible to handle scanning and indexing as a single task, or will it be more effective to pre-scan? Could OCR be used to obtain the initial indexes? How do you cross-reference to existing computer systems?
	How do you arrange information so that scanning is minimised and disc space is not wasted on blank entries in the documents?	Are there optional areas on the form? Can these be grouped? Is too much shading being used, thereby reducing the 'white' space? Is information repeated? Is internal information held that is available from data processing systems?

— Workflow software can be used to prefetch all the items once the initial reference has been input.

— All the required items can be specified by the users as part of their working pattern. The system will then prefetch them, ready for use.

There is no universally correct solution. EDM system designers must review the options, gain an understanding not only of what the users do but how they work, and select the option that is most appropriate for the particular application. For example, a mortgage-broking company reviewed the use made of its files and found that more than 90 per cent of queries were answered by referencing only the last two items of correspondence. The clustering of documents was arranged on this basis. This resulted in faster retrieval times and better use of the disc space available.

Determining the level of indexing detail is another area in which requirements must be carefully reviewed. Indexes may be as simple or as complex as required. This issue can be particularly important where an existing manual document-handling system uses a simple index in which each entry points to a large number of documents. In this situation, the users of the existing system may not be prepared (or able) to put in the effort required to create an index suitable for EDM purposes. This aspect of EDM was a major issue in the consultancy assignment carried out by Butler Cox for a media company (which was described in Figure 4.4). The company was considering whether to convert its existing press-cuttings and product-reference library to EDM, but was not willing to index the existing library. This illustrates that the full benefits of EDM will not be realised unless there is a willingness to change existing working practices.

Choosing the right supplier

Because of the relative immaturity of EDM technology, organisations wishing to implement EDM systems today may need to purchase equipment and software from a new and relatively unfamiliar supplier who specialises in EDM. Mainstream data processing suppliers are lagging behind specialist EDM companies and office systems and equipment suppliers in their EDM capability. In choosing an appropriate supplier, or combination of suppliers, organisations should apply a consistent set of selection criteria, and ensure that they have a good understanding of the supply side of the market.

Applying selection criteria

The criteria that Foundation members should apply in choosing an EDM supplier will vary depending on the nature and scale of the applications and the information technology environment that is already installed. An essential first step must therefore always be to define the functional and technical requirements for the EDM application. This will be based on the review of working methods described earlier in this chapter, which should be carried out before selection of a supplier begins. Where a corporate-wide choice of supplier is being made, it is important to identify all the main potential uses for EDM in the organisation, and to highlight their essential features in terms of size, scope, application types, and the likely requirements for integrating EDM with other systems. The selection criteria will fall naturally out of these analyses, but the following are the most common:

— Ability of the supplier's offerings to match both the corporate or strategic plans for EDM, and the specific applications requirements.

— Ability of the products on offer to coexist and integrate with the existing and planned information technology infrastructure.

— Compliance with current and planned technology policies in areas such as workstations, standards, networking, and user interface (for example, some products require separate workstations to be used for database lookup and image processing).

— Usability of the workflow language.

— EDM skills available within the supplier — especially the understanding of EDM applications needs, and systems integration ability.

— The supplier's stability and product-enhancement plans.

Sometimes, it may be better to select a combination of suppliers, particularly when requirements are complex or unusual, and when technical skills are scarce in-house. In such cases, it may be better to select a software supplier or systems integrator, as well as a hardware supplier. The successful EDM application that was described in Figure 2.3 was designed and implemented by a consultant, a systems integration company, the supplier of the retrieval software, and the hardware supplier.

Understanding the supply market

EDM systems are based on a new combination of technologies, and suppliers come from varied backgrounds, as Figure 5.5 illustrates. Each type of supplier therefore has a different understanding of EDM requirements and market demands, coloured by his original perspective. This makes it difficult for potential user organisations to understand what is on offer, because each type of supplier positions his products according to his background. The result is that potential EDM users are confused by the often contradictory information they receive from different types of supplier.

Furthermore, the products on offer vary considerably. They include:

Figure 5.5 EDM suppliers come from different backgrounds and therefore have different perspectives

Background	Supplier	Sample products
Specialist document image processing	FileNet Xionics Plexus	FileNet DIP-X XDP
Office equipment	Philips Olivetti (FileNet)	Megadoc FileNet
Office systems	Wang	WIIS (Wang Integrated Image System)
Data processing	IBM Digital Unisys ICL Bull	ImagePlus VAXimage OFIS IMS SM DOC
Microform	Kodak	KIMS, OD 1000
Communications	Racal	REOS
Text retrieval and management	Information Dimensions BRS	BASIS Archea

- Single-user, personal-computer-based document image processing systems (such as Xionics' DIP-X, and Kodak's OD 1000.
- Multi-user, minicomputer-based systems incorporating workflow software (such as Olivetti's, whose system is based on FileNet's hardware, and Racal's REOS).
- Fully integrated document image processing systems (such as Wang's WIIS).
- Document image processing systems with the ability to include computer-aided retrieval of microfilm (such as Kodak's KIMS).
- Powerful text-retrieval packages with image-handling capability (such as BRS's Archea and Information Dimensions' BASIS).
- Systems based on suppliers' 'strategic' compound-document architectures, which provide the ability to mix images and other types of information in the same document. (DEC and Unisys both provide this type of product.)
- Large mainframe-based systems with the ability to include images as part of the main database (such as ImagePlus from IBM).

In general, the suppliers who have so far shown the best understanding of the market for EDM and its application needs, and who have the best-developed products for most types of application, are the well established specialist document image processing suppliers. A prime example is FileNet, which has improved its market position by forming alliances with Olivetti and IBM, and now claims to have over 140 systems installed worldwide.

Other well-established suppliers are Philips and Wang. Philips claims market leadership in Western Europe, with over 7,000 terminals now linked to Megadoc installations. Wang has developed a clear lead in systems integration, and WIIS is very well-integrated with the PACE fourth-generation language and development environment that is at the core of Wang's current product range. As a result, Wang can offer fully integrated management of text, data, image, and voice. Wang claims that more than 200 WIIS systems have been installed worldwide and states that it has won a $27 million contract

in France to install the largest EDM system in the world.

The EDM products available from mainstream data processing suppliers are not so well developed, however. As a result, they have opted for alliances with specialist EDM companies — such as IBM with FileNet. They have also announced proprietary compound-document architectures that include image capability, thereby indicating that they intend to add image-processing capabilities to their products over time. Once the mainstream data processing suppliers begin to offer fully integrated products with EDM capabilities, many organisations will undoubtedly opt to purchase EDM products from their existing data processing suppliers. However, it could be some time before data processing suppliers offer this type of capability. Meanwhile, Foundation members should consider alternative suppliers, although there are obvious risks in basing long-term plans on small suppliers, who may not all survive once the large established computer suppliers start to deliver the right products.

One of the most interesting developments has been in the area of reference applications, where established suppliers of full-text-retrieval software have begun to recognise the opportunities provided by EDM. A good example is Information Dimensions, which plans to develop its established BASIS package so that its powerful indexing and searching facilities can be applied to all types of information, not just text. It is therefore building links between BASIS and image-processing, office, and electronic-publishing systems. These links conform with a range of interface standards, and will eventually allow BASIS to form the foundation for a powerful EDM environment.

Coordinating the use of EDM

Throughout this report, we have emphasised that the key to success with EDM is the ability to integrate EDM systems with other types of system. Unless EDM forms part of the overall systems strategy, its full potential will not be realised. Our research indicates, however, that at present, most companies are not taking a strategic approach to the implementation of EDM. Instead, there is a tendency to implement one-off EDM applications, often initiated outside the systems department. We have also pointed out that large-scale investment is required for EDM, and this again indicates the need for business and systems management to work in partnership.

Integrating systems

There is a danger that, as in the early days of word processing, personal computers, and other office systems, EDM may be implemented by users largely outside the control of the systems department. Many of the successful EDM applications that we identified in the course of our research were introduced by the user community, with little assistance from the systems department. Although standalone applications may be implemented successfully in this way, systems managers must exert some influence if EDM users are to avoid falling into the trap that many organisations fell into several years ago in the field of office systems.

Many user-led office systems implementations were very successful in providing automated support for individual workgroups or departments. However, the overall result for many organisations was a proliferation of different approaches, which led to the fragmentation of information across organisations, incompatible systems, excessive support costs for office systems, and missed opportunities for exchanging information. Many of these organisations are still having to address the problems stemming from early, uncoordinated, office systems initiatives, and have only recently begun to take a corporate view of office systems.

Unless systems managers take the lead in planning and coordinating EDM, a similar situation may well arise. Information will be duplicated, systems will be designed to meet local needs, they will lack flexibility, and they may well be incapable of being amended to reflect organisational changes. The full returns on EDM investments will therefore not be achieved unless the introduction of EDM systems is properly coordinated by the systems department. This does not mean, however, that EDM should be planned centrally, in isolation from the user community.

Forming a partnership between business and systems managers

To date, business managers have often taken the lead in introducing EDM (all the successful

applications described in Chapter 2 were user-led). Users have a vital contribution to make to the success of EDM implementations, and the systems department must build on users' experiences if the full potential of EDM is to be realised. The systems department must therefore form a partnership with user departments, with each contributing its own particular skills and experience.

Systems specialists have knowledge of project management, and an awareness of the opportunities for exploiting links between EDM and existing systems. The systems department should also aim to provide the specialist skills needed to design and implement EDM, to set technical policies, and to coordinate EDM projects and applications.

The team responsible for developing an EDM application will also require significant contributions from users, who will have a clear view of their workflows, their information needs, the scope of the individual tasks, and the priorities that must be given to the individual aspects of a task.

Making appropriate staff available

Earlier in this report, we identified that several different types of skills are required to implement EDM systems successfully. Because EDM is new and relatively unfamiliar, many of the skills are not currently available in user organisations, and are also in short supply in supplier companies. The successful staffing of EDM projects therefore depends on providing the right mix of skills, and considering the possibility of supplementing in-house resources with external support.

Providing the right mix of skills

Designing and implementing an EDM system requires both technical skills and skills in the design of working practices and procedures. (The full range of skills required is listed in Figure 5.6.) It is unlikely that the systems department will have all of these skills available from its own resources.

The precise combination of skills needed will vary as an EDM project progresses. In the early stages, the main requirements are for good business-analysis skills, combined with knowledge of the capabilities of EDM. The next stage is usually a thorough feasibility study, for which greater technical knowledge is needed. At this stage, O&M skills are also important to ensure that the review of existing working procedures is carried out properly. The same team members used at this stage should then carry out the system design work, although additional resources may be needed to carry out detailed design work on the new system and working procedures. For reference systems, an experienced information-management specialist will be needed to design the indexing structure, reference paths, and document-handling procedures that will be built into the system.

Throughout the design stage, the team must be prepared to involve users in deciding how the new system will work. (We discussed some ways of doing this earlier in this chapter.) The team must therefore have good communications skills, and users should be members of the team on a full- or part-time basis. At the systems construction stage, there will be a need for skilled systems integrators to develop the necessary links with existing systems.

Prior to the implementation of the system, there will be a need to prepare users for the changes in working practices that will come about because of EDM. It will therefore be necessary to use people who have skills in designing and implementing organisational changes, and in designing and delivering training programmes.

Using external support where needed

Because of the need for a mix of skills, some of which are not readily available, it may be

Figure 5.6 Designing and implementing an EDM application requires a wide range of skills

- Business analysis, combined with knowledge of EDM.
- Technical EDM skills.
- O&M skills.
- Understanding of the existing systems with which EDM needs to connect.
- Information management.
- User communications.
- Systems development.
- Systems integration.
- Training and organisational development.

BUTLER COX FOUNDATION

Butler Cox

Butler Cox is an independent management consultancy and research organisation, specialising in the application of information technology within commerce, government, and industry. The company offers a wide range of services both to suppliers and users of this technology. The Butler Cox Foundation is a service operated by Butler Cox on behalf of subscribing members.

Objectives of the Foundation

The Butler Cox Foundation sets out to study on behalf of subscribing members the opportunities and possible threats arising from developments in the field of information systems.

The Foundation not only provides access to an extensive and coherent programme of continuous research, it also provides an opportunity for widespread exchange of experience and views between its members.

Membership of the Foundation

The majority of organisations participating in the Butler Cox Foundation are large organisations seeking to exploit to the full the most recent developments in information systems technology. An important minority of the membership is formed by suppliers of the technology. The membership is international, with participants from Australia, Belgium, France, Germany, Italy, the Netherlands, Sweden, Switzerland, the United Kingdom, and elsewhere.

The Foundation research programme

The research programme is planned jointly by Butler Cox and by the member organisations. Half of the research topics are selected by Butler Cox and half by preferences expressed by the membership. Each year a shortlist of topics is circulated for consideration by the members. Member organisations rank the topics according to their own requirements and as a result of this process, members' preferences are determined.

Before each research project starts there is a further opportunity for members to influence the direction of the research. A detailed description of the project defining its scope and the issues to be addressed is sent to all members for comment.

The report series

The Foundation publishes six reports each year. The reports are intended to be read primarily by senior and middle managers who are concerned with the planning of information systems. They are, however, written in a style that makes them suitable to be read both by line managers and functional managers. The reports concentrate on defining key management issues and on offering advice and guidance on how and when to address those issues.

Selected reports

8 Project Management
20 The Interface Between People and Equipment
24 Investment in Systems
25 System Development Methods
27 Developments in Videotex
28 User Experience with Data Networks
29 Implementing Office Systems
30 End-User Computing
31 A Director's Guide to Information Technology
32 Data Management
33 Managing Operational Computer Services
34 Strategic Systems Planning
35 Multifunction Equipment
36 Cost-effective Systems Development and Maintenance
37 Expert Systems
38 Selecting Local Network Facilities
39 Trends in Information Technology
40 Presenting Information to Managers
41 Managing the Human Aspects of Change
42 Value Added Network Services
43 Managing the Microcomputer in Business
44 Office Systems: Applications and Organisational Impact
45 Building Quality Systems
46 Network Architectures for Interconnecting Systems
47 The Effective Use of System Building Tools
48 Measuring the Performance of the Information Systems Function
49 Developing and Implementing a Systems Strategy
50 Unlocking the Corporate Data Resource
51 Threats to Computer Systems
52 Organising the Systems Department
53 Using Information Technology to Improve Decision Making
54 Integrated Networks
55 Planning the Corporate Data Centre
56 The Impact of Information Technology on Corporate Organisation Structure
57 Using System Development Methods
58 Senior Management IT Education
59 Electronic Data Interchange
60 Expert Systems in Business
61 Competitive-Edge Applications: Myths and Reality
62 Communications Infrastructure for Buildings
63 The Future of the Personal Workstation
64 Managing the Evolution of Corporate Databases
65 Network Management
66 Marketing the Systems Department
67 Computer-Aided Software Engineering (CASE)
68 Mobile Communications
69 Software Strategy

Forthcoming reports

Staffing the Systems Department
Managing Multivendor Systems
New Information Technologies
The Future of System Building Tools

Availability of reports

Members of the Butler Cox Foundation receive three copies of each report upon publication; additional copies and copies of earlier reports may be purchased by members from Butler Cox.

Butler Cox plc
Butler Cox House, 12 Bloomsbury Square,
London WC1A 2LL, England
☎ (01) 831 0101, Telex 8813717 BUTCOX G
Fax (01) 831 6250

Belgium and the Netherlands
Butler Cox BV
Burg Hogguerstraat 791,
1064 EB Amsterdam, the Netherlands
☎ (020) 139955, Fax (020) 131157

France
Butler Cox SARL
Tour Akzo, 164 Rue Ambroise Croizat,
93204 St Denis-Cédex 1, France
☎ (1) 48.20.61.64, Télécopieur (1) 48.20.72.58

Germany (FR)
Butler Cox GmbH
Richard-Wagner-Str. 13, 8000 München 2, West Germany
☎ (089) 5 23 40 01, Fax (089) 5 23 35 15

United States of America
Butler Cox Inc.
150 East 58th Street, New York, NY 10155, USA
☎ (212) 891 8188

Australia and New Zealand
Mr J Cooper
Butler Cox Foundation
Level 10, 70 Pitt Street, Sydney, NSW 2000, Australia
☎ (02) 223 6922, Fax (02) 223 6997

Finland
TT-Innovation Oy
Meritullinkatu 33, SF-00170 Helsinki, Finland
☎ (90) 135 1533, Fax (90) 135 1091

Ireland
SD Consulting
72 Merrion Square, Dublin 2, Ireland
☎ (01) 766088/762501, Telex 31077 EI,
Fax (01) 767945

Italy
RSO Futura Srl
Via Leopardi 1, 20123 Milano, Italy
☎ (02) 720 00 583, Fax (02) 806 800

The Nordic Region
Statskonsult AB
Jungfrudansen 21, Box 4040, 17104 Solna, Sweden
☎ (08) 730 03 00