Document Imaging: Perspective

Summary

Document imaging enables paper documents to be captured and transformed into digital images for storage and retrieval. It is used in insurance, financial services, healthcare and government applications.

Note

Document imaging technology still warrants coverage because, despite increased movement toward electronic documents and electronic transactions, much business information is still contained in paper form, and organizations still conduct paper-based transactions.

Table of Contents

Technology Basics
Document Imaging vs. Document Management
System Components
Architecture
Operating Requirements
Technology Analysis
Business Use
Benefits and Risks
Standards
Price vs. Performance
Selection Guidelines
Technology Leaders
Technology Alternatives
Insight

List Of Tables

Table 1: Document Imaging and Related Standards
Technology Basics

The document imaging market emerged in the mid-1980s with the advent of optical disk technology for storage. In its infancy, document imaging was a technology for replacing paper-based filing cabinets and microfilm archives. Its prime use was for storage and retrieval of documents in paper-intensive organizations, such as insurance companies, banks, and financial service institutions. Vendors initially offered turnkey systems based on proprietary hardware and software and targeted for use within a specific department. Document imaging systems have since evolved from stand-alone, turnkey systems to image management software that runs on industry-standard platforms. Most vendors now market solutions built on a three-tier architecture with access to image repositories provided through thin clients (Web browsers). Optical storage has been eclipsed by other storage technologies, such as redundant array of independent disks (RAID), and browser-based plug-ins have replaced dedicated image-viewing technology. The only real true pieces of imaging today are the capture and transformation of hard-copy documents into digital information. A number of vendors, however, still supply complete solutions for capturing, managing, storing, and retrieving document images and their associated text files. These include eStream WMS, Inc. (formerly Eastman Software), FileNET, IBM, and TOWER Technology. Others, such as Captiva (formerly ActionPoint) and Kofax, supply the front-end capture software or subsystems that can link to the back-end repositories. In general, the document imaging market is flat or even declining in growth and has been characterized by consolidation. The vendors that will survive are those with a strong vertical focus.

Document Imaging vs. Document Management

The demarcation between document processing technologies, particularly document imaging and document management, has blurred. The primary difference between document imaging and document management technologies is that document imaging systems tend to be used primarily for managing and archiving static documents. These are documents that do not change and that need to be managed in their final format to satisfy legal or regulatory compliance issues. Document management systems, on the other hand, are designed to manage dynamic documents that originate in electronic form. They enable documents or components to be changed or reused to form new documents. Most integrated document management (IDM) systems have the capability to manage images as well as other data types, with images, video and audio files, text files, etc., all being treated as objects. IDM solutions are generally advisable when image volumes are lower than 25,000 to 30,000 images per day. Production imaging solutions typically handle 50,000 to 100,000 images per day.

System Components

Typically, an imaging system is built from a series of hardware and software subsystems, each designed for a particular function in the process, such as capturing and converting the images from paper to electronic code; classifying the images for later retrieval (indexing); file searching (via keywords, text retrieval, etc.); managing the images and index data (metadata), and storing/archiving the images; and distributing or routing the images as part of a business transaction. The core component of an imaging system is, of course, the software for processing and managing the scanned document images. Most imaging software today is developed for Microsoft Windows environments with server software running on Microsoft Windows NT or Windows 2000 and desktop clients supporting Windows 98, Windows NT, and Windows 2000. Typically, imaging software is developed using Microsoft’s Distributed Component Object Model (DCOM) certified for Microsoft BackOffice.

Because image software applications have varying requirements, these programs typically come with a means to customize the basic image management runtime components. Customization capabilities can
Document Imaging: Perspective

be provided through application programming interfaces (APIs) that can be called from specific high-level programming languages (such as C or C++) to obtain imaging services.

Capture

All document imaging systems provide some means of capturing paper documents or microfilm and converting them to computer-readable data. The most common input device for document images is the scanner. Scanners translate the image on a sheet of paper or other surface into a digitized image that can be read by the computer. Usually, images are compressed into a standard image file format, such as Tagged Image File Format (TIFF), before they are entered into the system. As an alternative file format to TIFF, images of paper documents can be transformed into Adobe’s Portable Document Format (PDF).

Depending on the application, image capture can be distributed or centralized. A best practice is typically to perform scanning at the place where the paper originates. If all the paper comes into one place, say a centralized mailroom, then it makes sense to centralize the scanning and perform batch capture. For other applications, like loan process, it makes sense to distribute scanning and have the images uploaded to a central repository.

Users can choose how they want their documents entered into the system: typically either as a compressed digitized image or as both digitized image and text. bScanners capable of both image and text scanning are equipped with Optical Character Recognition (OCR) software, or the OCR capability is embedded in the image input software. Enterprises should bear in mind, however, that OCR software programs are at best 90-98 percent accurate. When used on clean printed text, OCR can achieve around 90+ percent accuracy.

Image capture functionality was originally a core part of the image processing and management software developed by the vendor, but gradually the market moved toward a modular, component architecture using best-of-breed technology. Although some imaging vendors, like FileNET, still develop their own capture component, most have turned to third parties, like Captiva and Kofax, for this capability.

Much like the blending of imaging and document management, forms processing and document capture technologies are converging under a single, total capture umbrella. Forms processing dramatically reduces the cost of capturing data from paper forms. In forms processing, standard scanned images are used to understand the layout of a form, extract the data, and pass it to one or more software recognition engines tuned to specific data types. Once the data is accurately extracted, the images can be sent as input to document management solutions designed to store and retrieve the forms more efficiently with many indexes.

Another method of acquiring document images is through fax integration. Most imaging systems support the reception and transmission of facsimile documents. Incoming faxes are then treated as scanned images that may be incorporated into the imaging system’s workflow.

Image capture is often outsourced, particularly if an enterprise has a large backfile volume. Costs for converting paper documents to images (based on outsourcing figures) will range from eight to 35 cents (U.S.) per page (may vary depending on requirements).

Viewing

Traditionally, imaging vendors incorporated proprietary viewers within their client software applications. Many vendors have now adopted universal viewers or moved to browser-based technologies, where the viewing components are automatically downloaded and executed from within the browser. View functions typically supported include reverse, rotate, scale, scroll, and zoom. Most products allow images to be annotated.
Data Compression

Scanned images are one of the most data-intensive forms of information that computers process. Uncompressed, each 8.5” × 11” page, scanned at 200 dots per inch (dpi), occupies 3.74 million bits (467KB). At 300 dpi, it occupies 8.42 million bits (over 1MB). Without compression, document imaging systems would need prohibitively expensive computers or would operate intolerably slowly. In data compression, mathematical algorithms are applied to reduce the size of the image by 10:1 or a greater ratio. Compression and decompression can be accomplished via software or hardware (add-in boards).

Most document imaging systems use the CCITT Group 3/Group 4 compression protocols, which were defined for facsimile communication. These compression standards encode the run length of black or white areas with an optimization toward white (presumably most of a page). Group 3, the standard for fax images, is further subdivided into one-dimensional (1-D) and two-dimensional (2-D) images. Most document imaging systems today use Group 4 compression. The recognized standard for grayscale and color imaging, including photographs, is Joint Photographic Experts Group (JPEG). In addition, some vendors have chosen to use the Joint Bitonal Image Group (JBIG) compression algorithm, an ISO standard for lossless, bitonal image compression, in addition to the standard CCITT Group 4 algorithm.

Image Indexing

After the image is scanned and verified, it must be classified or indexed. Image indexing involves the entering of character data that describes or tags an image for subsequent retrieval. This data can range from serial numbering to a lengthy, structured description. Traditionally, indexing data was entered manually from the keyboard, but with the development of more sophisticated scanners, intelligent controllers, and OCR/ICR technology, manual indexing can be substantially reduced. Most products support the traditional key-in indexing method, and many support automatic indexing through bar code scanning. Some support automatic indexing by downloading data from an external database. Indexing data is stored in a database specifically designed to hold image data. What data can be used to retrieve an image is a function of the image index database. Some allow the image transaction history to be recorded. Some index databases link the image files to established databases from other applications by pointers or common fields.

Imaging systems may not need to provide index management services if an established database can search for documents by attributes such as an applicant’s name or social security number. For example, a social security number can be assigned as a document locator number during scanning. Then, the established database can be used, say, to search for all documents that pertain to a particular region by the social security numbers of all applicants that live within a certain state or zip code area.

To increase access speeds, the image index database is often stored on magnetic disk, apart from the images on optical disk. The index database may also be stored on a separate dedicated server or on a host mainframe, apart from the imaging system.

Most document imaging systems today use a standard relational-type database to store the indices and pointers to the actual image. As a result, connection to an established system allowing the integration of images with accounting or other computer records is fairly easy.

Image Storage

When it comes to storing images, optical storage systems are gradually giving way to magnetic storage. When document imaging began in the late 1980s and early 1990s, optical disks and jukeboxes were the only medium with sufficient speed and capacity for scanned images. At roughly 50KB per page (5KB compressed), even a modest image database outstripped the capacities of the hard disks of the day. By
the late 1990s, with the advent of RAID and the rapidly increasing capacities of hard disk drives, small image databases could already be stored more economically on magnetic disk. RAID is also easier to deploy, less prone to failure, more easy to restore when it does fail and much faster (files retrieved in 10 milliseconds instead of 10-20 seconds on optical jukeboxes). Since then, magnetic disk capacities have continued to increase more rapidly than optical disks, making RAID economical for increasingly large databases. As the demand for archival optical storage has diminished, some vendors have dropped out and the survivors are struggling to keep the 5.25” and 12” optical storage technologies competitive with magnetic storage. Standards wars between vendors have also frightened buyers away from writable DVD technology. Many users nevertheless are sticking with optical storage, at least for now, and there are three main reasons why:

- **Write-once capability.** Some regulatory agencies specify that certain documents be stored on paper, microfilm or, if in digital form, in a medium that cannot be erased or overwritten. Buyers should seek expert legal advice on the specifics of this requirement, since not all optical media define write once in the same way. Some users have also begun to experiment with specially adapted write-once tape media that might satisfy the requirement.

- **System design.** Most established document imaging systems are based on 5.25” or 12” optical storage, and for many users it is too difficult to make the change. Now that terabytes of images are in place, migration to magnetic storage would be time consuming, costly and might entail major changes to the software. In the more extreme cases, the retrieval software references documents as a physical location on a specific disk.

- **Capacity.** At multiterabyte capacities optical storage can cost less to purchase than high-quality RAID systems and the multiple servers that they require. Buyers who choose optical for this reason should look closely at the cost of maintaining the optical storage system over its life, and the need to replace or overhaul the jukebox when it reaches a certain number of retrievals.

**Image Distribution**

Document imaging systems require some means of distributing the images, whether they are retrieved to be viewed, manipulated onscreen, transmitted via facsimile, routed to other users via electronic mail or as part of a workflow, or printed out in hard-copy form. In document imaging and document management systems, image distribution can be either online or offline. Document-centric workflow was designed to enable controlled routing of documents in a paper-intensive industry, where its purpose was to identify and distribute electronic images of paper to clerical staff responsible for processing that paper. The Internet and corporate intranets and extranets have increasingly become the mechanism for distributing images and work items throughout an enterprise and beyond.

**Architecture**

Document imaging software based on a client/server model was the norm for most of the 1990s. In this approach, core functionality resided on a single server or was distributed across multiple servers, while the user interface resided on networked clients. In recent years, however, most vendors have moved to Web-centric, three-tier architectures. Microsoft Windows NT or Windows 2000 are the predominant server platforms for imaging solutions, although some document imaging products support a variety of Unix servers and, occasionally, Linux as well. Most products support Windows-based desktops as well as Web browsers (typically, Microsoft Internet Explorer or Netscape) as clients. Web servers supported usually include Microsoft Internet Information Server (IIS) or IBM’s WebSphere. Communications between clients and servers is typically through TCP/IP or HTTP. Products may come with a proprietary database or support a variety of popular SQL databases, such as Oracle, Informix, Sybase, and Microsoft SQL Server.
Operating Requirements

A typical production-level document imaging system requires the following components:

- Client workstations: Windows desktop client or thin client (Web browser, typically Internet Explorer or Netscape Navigator)
- One or more servers: Intel Pentium processor running Windows NT Server or Windows 2000, or Unix workstation (Sun Solaris, HP-UX, IBM AIX, etc.)
- HTTP Server
- Relational database: typically, Microsoft SQL Server, Oracle, or Sybase
- One or more dedicated scanning workstations
- Fax server(s)
- Print server(s)
- LAN, WAN, or intranet: TCP/IP networks
- Document imaging or integrated document management server software
- Communications architecture and software, such as Microsoft Exchange or MAPI-compliant messaging systems
- Capture software: forms, OCR, etc.
- A storage device, such as RAID or magnetic tape, or an optical storage subsystem
- Optical storage subsystems for document imaging consist of the following components:
  - Jukebox or auto changer with drives and media (There may also be a stand-alone drive for recording.)
  - Jukebox management software
  - Hard disk cache either built into the jukebox or on a server
  - Unix- or Windows NT-based server connected to the network and/or mainframe (Some jukeboxes are sold with a built-in server and network interface. This is called network-attached storage, or NAS.)
  - Backup system

A production-level system also requires human resources, such as a dedicated system administrator, a database administrator, and/or a network administrator.

Technology Analysis

Business Use

In its early days, document imaging technology was justified as a replacement for paper-based filing methods. In most cases, it was used purely for storage and retrieval and archival applications. Once paper documents were digitized, however, organizations began to realize that image documents could be easily distributed or routed to others for approval or collaboration; they could be linked to other
Document Imaging: Perspective

applications; and they could be used to enable automation of other business processes, to improve
customer service operations, and to provide a competitive edge.

Imaging technology today can be used in horizontal, general-purpose applications or in highly customized
vertical applications. The Health Insurance Portability and Accountability Act (HIPAA) is driving the need
for document imaging technology within the healthcare and insurance industries. HIPAA’s goal is to
reduce administrative costs associated with and improve the efficiency of processing medical records
(claims and other patient information) while protecting patients’ privacy. Likewise, the U.S. Government’s
Paperwork Elimination Act (GPEA), which took effect in 1998, has as its goal the promotion of doing
business electronically, with the public and otherwise. This has increased the use of document imaging in
government applications. Typically, document imaging has been implemented for mission-critical
applications where the document is critical to the success of the business operation:

- Banking and Financial services—wire funds transfer, loan processing, credit card processing,
  IRA/Keogh trust account processing, remittance processing, imaged statement rendition.

- Federal government—patent and trademark tracking, maintenance documentation, personnel
  records, census information, Freedom of Information Act records.

- State and local governments—pensions, tax records, deeds, county court records, arrest records.

- Insurance—claims processing, underwriting, policyholder service, HIPAA compliance.

- Transportation—waybills and inspection reports, frequent traveler programs, accounts payable,
  passenger revenue accounting.

- Utilities—large document applications, accounts payable, manufacturing plant records.

- Manufacturing—aerospace documentation, accounts payable, pharmaceutical case reports forms,
  publishing records.

- Retail trade—accounts payable.

- Services—automobile rental agreements, package delivery.

- Healthcare—medical records, patient accounting, admitting, HMO enrollment, physician credentials,
  HIPAA compliance.

Benefits and Risks

The capability of transforming paper documents into digitized images that can be stored and routed
electronically affords organizations a number of significant benefits. Adopters of document imaging
technology have been able to reduce or eliminate the amount of filing cabinets and/or storage space
required for storing the original paper documents while gaining the advantage of immediate access to the
information. The use of document imaging technology has also helped eliminate misplaced or lost
documents/files and has allowed multiple people to access the same document simultaneously, thereby
improving productivity and increasing customer service responsiveness. For some organizations,
implementing document imaging technology has enabled them to reduce head count as well as the labor
costs associated with processing paper-based documents.

Despite the many well-publicized hard and soft benefits afforded by document imaging, the technology
has several disadvantages and inherent risks. A much-debated issue over the years has been the legality
of image storage. Currently, the acceptability of document images stored on optical media as legally
admissible evidence varies from state to state. In addition, the traditional storage media for imaging
Document Imaging: Perspective

Applications—optical disks—has not proven itself to be the best medium for long-term archival applications because of issues with the longevity of the storage media and lack of standards—readers and software may not exist in the future. It requires frequent upgrades to new technology and conversions of files. Many organizations still use microfilm in parallel or save onto CD. An issue sometimes associated with networked document imaging systems is the size of data involved. Although a single document page may be 50KB in size compressed, which will not stress a network or transmission mechanism, many documents are two, three, or more pages long. A multipage document is of limited use if the person requesting it does not get all the pages. So often, one user is requesting transmissions of 100KB, 150KB, or more. If these images are transmitted as one block of data, they can saturate a network. As color images, photographs, and video images are increasingly being distributed over networks, image sizes will grow, substantially increasing the load on corporate networks.

Standards

<table>
<thead>
<tr>
<th>Table 1: Document Imaging and Related Standards</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard</strong></td>
</tr>
<tr>
<td>ISIS</td>
</tr>
<tr>
<td>ODMA</td>
</tr>
<tr>
<td>OLE</td>
</tr>
<tr>
<td>TWAIN</td>
</tr>
</tbody>
</table>

Price vs. Performance

The cost of imaging systems varies greatly depending on the configuration (for example, the number of servers, scanning/capture workstations, and image workstations required), the number of concurrent users, and the complexity of the application (that is, how much customization and ongoing administration is needed). Typical costs range from US$50,000 for a departmental implementation to millions of dollars for an enterprise implementation. In addition, costs for back-file conversions, process reengineering, implementation services, and ongoing maintenance and support are often more than the initial system.
Selection Guidelines

Obvious factors to consider when selecting a document imaging system include the product’s functionality and technical architecture, the cost (both initial and ongoing), the vendor’s ability to execute and its financial viability, the services and support the vendor or its implementation partner can provide and, lastly, the vendor’s vision or commitment to the product and market:

- **Image Volumes**: extremely high-volume imaging requirements will steer users to a small set of vendors. Ad hoc or occasional use, low to mid-volume (under 25,000 images per day), or production (more than 50,000 images per day).

- **Scanning**: One consideration is what peripherals are supported for image acquisition. Often, the ultimate success of a document imaging implementation rides on the appropriate choice of scanners for the volumes and types of documents to be captured. Identifying the media accurately is the most critical item in selecting a scanning device, followed by the purpose for scanning, speed, and resolution. For speed, users should consider how and when the media is received; the volume per day, week, and month; peak volumes; turnaround time; and type of media. For resolution, users should consider how the images will be output. For most text applications, scanner resolutions of no more than 200 dpi are sufficient. Inputting at a resolution of 200 dpi provides fairly legible output quality; 300 dpi provides typical laser printer quality; 400 dpi is equivalent to Group 4 fax; and 600 dpi provides the same quality as the best laser printers. The higher the resolution desired, the higher the cost in storage, transmission, and hardware.

- **Quality Control**: Another feature to look for when examining systems is the ease with which users can control the quality of the documents during image scanning. Increasingly, vendors are automating this function. Most products can re-scan, re-enter, and replace badly scanned documents. The majority also support rejection of individual documents, document batch rejection, and automatic data field validation checks. Another commonly supported feature is performing confidence checks on scanned data.

- **Performance**: The speed at which images can be accessed is often an issue with imaging solutions. Look for imaging software that stores images in a standards-based (for example, TIFF or JPEG) compression format and decompresses them at the client workstation. Other methods for improving performance include downloading and displaying the first pages of large documents while the remaining pages are downloaded to the client, and caching images on remote servers and/or clients. Software products—like FileNET’s Panagon Image Services, IBM’s Content Manager, or TOWER Technology’s TOWER IDM—are optimized for handling large volumes of images.

Technology Leaders

Since its heyday, the document imaging market has undergone considerable consolidation among the vendors through high-profile mergers and acquisitions as the technology has become commoditized. A few key players currently remain in the production imaging space, while Kofax and Captiva (formerly ActionPoint) have carved out successful niches in the capture/transform subsystem market.

**Captiva**

Two former leaders in document and data capture, Captiva Software Corp. and ActionPoint, Inc., merged in July 2002. The combined entity adopted the Captiva name and is a much stronger challenge to Kofax, the recognized leader in image capture.

Captiva’s products consist of a suite of input management solutions that includes FormWare for forms processing, InputAccel for document capture, PixTools, an imaging toolkit for developers, and ISIS, the
Document Imaging: Perspective

de-facto standard scanner interface. InputAcCEL is used by Documentum, Open Text and IBM as the capture module for their document management solutions.

eiStream WMS, Inc.

eiStream was formed as a result of several acquisitions, mainly the acquisition of Eastman Software from Eastman Kodak in August 2000. The company was officially launched as eiStream in September 2001. It consists of the following entities: eiStream WMS, Inc.; eiStream Kofile, Inc.; eiStream ViewStar; eiStream Services; and various foreign affiliates. eiStream reports having 4,000 sites and 370,000 licenses worldwide, but this is for all eiStream products, including the mid-tier imaging and recognition offerings from eiStream Kofile. eiStream WMS, Inc., is among the leading suppliers of production imaging and workflow software, and targets applications in financial services, insurance and government. eiStream’s workflow and imaging products are designed to help improve an organization’s customer service, document production, forms processing, marketing and product development functions.

eiStream WMS, Inc.’s Enterprise Edition: Imaging and Workflow is a set of products and services for developing production imaging and workflow applications. WMS Enterprise Edition: Workflow can be licensed independently of the Imaging product and vice versa, or the two can be implemented together for a fully integrated work management solution. WMS Imaging and Workflow can also be tightly integrated with eiStream’s Reports Manager solution for handling mainframe report data. WMS Enterprise Edition: Imaging and Workflow has typically been implemented to support back-office applications with anywhere from 50 to 500 seats.

FileNET Corp.

Since its inception in the early 1980s, FileNET has been known primarily for its integrated document management (IDM) suite of products and has been one of the market leaders in production imaging and document-centric workflow applications. The firm has an installed base of over 3,300 systems worldwide. FileNET Panagon represents the umbrella name for FileNET’s suite of client and server software for enterprise content management and process automation. Core components in FileNET’s Panagon family are Panagon IDM Desktop, an integrated client for imaging, document management, COLD and workflow; Panagon Image Services for production imaging; Panagon Content Services for document and Web content management; and Panagon Web Services, the Web-based application development environment. Separate products include Panagon Capture, available in Professional and Desktop versions, for capturing paper-based information; Panagon Web Publisher, a toolkit for publishing documents to internal or external Web sites; Panagon eProcess Services and WorkFlo Services for more advanced workflow capabilities; and Panagon Report Manager for accessing COLD data.

IBM

Since the 1980s, IBM has been a leading production imaging and workflow vendor with one of the largest installed bases (9,000+ systems) of any vendor and a product line that targets a range of platforms from networked PCs to mainframes. Its strength as a systems and services vendor has contributed to IBM’s position as one of the leaders in the imaging and workflow markets.

In March 2000 IBM introduced IBM Content Manager, a product that combined the features and functionality of two former IBM products—the IBM Digital Library media asset management system and the ImagePlus VisualInfo production imaging solution—into a single content repository running on DB2 or Oracle in Windows or AIX server environments. Since introducing Content Manager, IBM has enhanced and rearchitected the product, making it more attractive and viable as an enterprise solution for a broad range of content management needs. IBM announced IBM Content Manager (CM) version 8.1 in June 2002. It will be generally available in the third quarter of 2002. CM v8.1 provides improved
Document Imaging: Perspective

workflow/document routing capabilities, integral search services, enhanced system administration capabilities, support for Sun Solaris servers, deeper integration with DB2 and a new viewer. In addition, it is built on a new, extensible architecture with a hierarchical data model.

TOWER Technology, Inc.

A privately held software developer and system integrator founded in 1987, TOWER Technology has a strong focus on high-volume, transaction-oriented applications. It specializes in distributed imaging and workflow applications, as well as in production color imaging, for large-scale, enterprise deployment. The firm markets the TOWER IDM suite, a set of software modules for production imaging, document management, enterprise report management and workflow (case management). TOWER now develops most of its own core technology, but has fleshed out its product suite by licensing best-of-breed components that it integrates at the API level. TOWER targets the insurance, financial services and healthcare industries, as well as government agencies. According to TOWER, it targets implementations ranging from 50+ users to thousands of users. Its average deal size, for software only, is about $1.5 million.

Kofax Image Products Inc.

Founded in 1985, Kofax was one of the pioneers in developing component imaging software for production applications. Kofax was acquired by Switzerland’s DICOM Group in January 2000 and now operates as a division of that company. Kofax develops and markets application software and image processing products for the electronic data and document capture market. Like ActionPoint, Kofax has a strong partnership focus with leading imaging and document management software vendors. The firm’s Ascent Capture software has been integrated with leading document imaging and document management systems, such as IBM’s Content Manager, Documentum’s 4i platform and Documentum 5, Hummingbird’s DM suite, and Open Text’s Livelink.

Technology Alternatives

It has been widely believed that the digital technologies used for storing images of documents on optical disk or magnetic media would inevitably replace the old analog technology of microfilm (or microforms). Often viewed as competing technologies, the fact is that digital imaging and microfilm complement each other, and a combination of the two may make sense in many applications.

Image-based technologies, although clearly superior for retrieval, distribution, copying, and transmission, require a commitment to maintenance. As the underlying technologies evolve and change, the media and systems used to retrieve the images are no longer available; on some media types, there is even a risk that images can disappear with associated data. Microfilm is still viable in applications where documents are infrequently retrieved. Capture and storage costs are low, and microfilm is widely recognized as admissible in court and approved by regulatory bodies. Microfilm has benefits for long-term archival purposes: it is fairly compact, it does not deteriorate quickly, and it cannot be overwritten. Microfilm has an estimated life of up to 500 years, if properly stored. Also, since microfilm is human readable, it should be possible to view it even if the original retrieval and display system is no longer usable. There is, therefore, a role for both images and microfilm, where images act as a fast medium in an electronic age, while film acts as an archival medium for storage of both data and images.

Insight

Most of what traditionally characterized core imaging functionality has been subsumed by other document processing technologies and merged into the computing infrastructure. Images are treated as just another object to be managed within the context of a business transaction. While business information is still
Document Imaging: Perspective

contained in paper form and organizations still conduct paper-based transactions, pure document imaging has become a niche technology for vertical markets. Enterprises that need to manage static documents, particularly high volumes of images, will find a small but capable group of vendors with proven technology and market longevity to address their requirements.