Document Imaging: Perspective

Summary

Document imaging refers to the capability to take a paper document and transform it into a digital image for storage and retrieval. It is a document technology used mainly in insurance, financial services, healthcare, legal, and government.

Note

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

Table of Contents

List Of Tables

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.

Over the years, document imaging systems evolved from stand-alone, turnkey systems to image management software that runs on industry-standard platforms, typically using a client/server architecture with industry-standard personal computers (PCs) connected via a local-area network (LAN) as the image workstations. This traditional client/server approach has since been supplanted by Web-centric architectures. Most vendors now market solutions built upon a three-tier architecture with access to image repositories provided through thin clients (Web browsers).

Moreover, 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.

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 eiStream WMS, Inc. (formerly Eastman Software), FileNET, IBM, and Tower Technology. Others, such as ActionPoint (formerly Input Software) and Kofax, supply the front-end capture software or subsystems that can link to the back-end repositories.

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.

Document Imaging: Perspective

Because image software applications have varying requirements, these programs often come with a means to customize the basic image management runtime components. Customization capabilities can 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).

In the client/server approach, scanners are often attached to a dedicated scanning workstation or server. In some cases, scanning workstations operate offline; the images may be cached on magnetic disk and then later transmitted to online optical disk storage (optical disks are specialized storage devices for images), or the images are scanned to offline optical disks. Later, the optical disks are brought online.

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. Scanners 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.

Image capture functionality was originally a core part of the image processing and management software developed by the vendor, but gradually the market moved towards 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 ActionPoint 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. Traditional image capture software vendors— like ActionPoint and Kofax— have added forms-processing capabilities to their capture products, while forms-processing vendors— like Captiva— have expanded their suites to include image capture and image enhancement functionality.

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.

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.

Data Compression

Document Imaging: Perspective

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 towards 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 existing databases from other applications by pointers or common fields.

Imaging systems may not need to provide index management services if an existing 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 existing 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 existing system allowing the integration of images with accounting or other computer records is fairly easy.

Image Storage

Two factors influence the choice of storage media for imaging systems: how long the documents need to be stored and how often the images are retrieved. For documents that must be stored over long periods of time, microfilm or optical storage is preferred over magnetic storage. Magnetic storage is useful in applications that require a high retrieval rate but relatively low volume. Optical storage is recommended for high-volume, high-retrieval rate environments. Optical storage is used mostly in document imaging and Integrated Document Archival and Retrieval Systems (IDARS), which include computer output to laser

disk (COLD) and Report Distribution Systems applications, nearly all of which have at least two of the following needs in common:

- Moderately fast retrieval times needed (10 to 20 seconds, faster with caching).
- Capacities in excess of 100GB.
- Write-once capability, which is needed or mandated by laws or regulations.

Several technologies are used for optical storage; they are read-only (CD-ROM and DVD), write-once, and rewritable/multifunction. The choice of rewritable versus write-once (WORM) optical disks depends on the application's requirements for ensuring the preservation of the original document over time. As a rewritable storage medium, optical storage has fallen behind magnetic disk technologies, such as RAID, in the race for price/performance. Rewritable and write-once 5.25" optical storage are used to store large numbers of scanned documents, Enterprise Report Management/COLD output, and archival data. They are mostly used in jukeboxes where the data can be retrieved in about ten seconds.

Image Distribution

Document imaging systems require some means of distributing the images, whether they are retrieved in order 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. The Internet and corporate intranets and extranets have increasingly become the mechanism for distributing images and work items throughout an enterprise and beyond.

Traditional 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. An example is claims processing. In this environment, the claims are received with associated support documents. They can be distributed to clerical staff based on criteria such as value of claims, customer identification, complexity, and completeness of supporting documentation. Once they have gone through a check for completeness and, if necessary, letters have been written to request more documentation, they can be sent to an initial approval process. Based on value, they can be sent for supervisory approval(s) if necessary. After approval, the claims can be filed and checks issued. Alternatively, specific claims may be routed to exception queues based on customer-determined criteria. At each stage, the document is tracked for timeliness of response, and any work created can be easily added to the claim file.

Imaging System Segmentation

Document imaging systems can be classified into production or office applications:

- Production Imaging— Production imaging systems, by far the most common types marketed, are designed for high-volume, transaction-oriented applications. They are used for applications such as mailroom automation, check processing, correspondence processing, etc. Typically, these systems have dedicated servers for document capture, indexing, and output and archiving. Production imaging systems include FileNET's Panagon Image Services, eiStream WMS's (formerly Eastman Software) Enterprise Imaging and Workflow, IBM Content Manager, and Tower Technology's Tower IDM.
- Office Imaging— Office or desktop imaging systems let casual users view, share, and manipulate a document image, usually on an ad hoc basis. This type of functionality is increasingly being built into

operating systems software (like Windows). Office imaging systems tend to be distributed throughout the organization and are typically departmental implementations.

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, there has been continued movement away from the traditional client/server architecture 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 client/server and or Web 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 and IDARS 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 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. Typically, it 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.
- 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.
- Health care— medical records, patient accounting, admitting, HMO enrollment, physician credentials.

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 headcount 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 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 1: Document Imaging and Related Standards

Standard	Description
ISIS	The ISIS standard, a software interface created by Pixel Translations (now part of ActionPoint), has become the most popular standard for scanners designed for document imaging. It was designed to provide a common interface that integrators could use to support and enable image scanning, compression, and display for any imaging scanner with no code change.
ODMA	The Open Document Management API (ODMA) is a client-side standard that enables desktop applications to talk to a single document management system. The ODMA specification was completed in 1994 and ported to 32-bit Windows environments in 1995. In late 1996, the Open Document Management API task force produced new query extensions consisting of 23 APIs that give developers a way to provide stronger interoperability between desktop applications such as word processors, spreadsheets, etc., and document management systems. ODMA 2.0 was approved in 1997. The ODMA API is platform independent.
OLE	Microsoft's Object Linking and Embedding (OLE) is a set of services that enable applications to interact and interoperate. OLE has become a de facto standard for linking objects. Images are treated as just another object and can be easily integrated with other desktop applications via OLE.
TWAIN	Sponsored by the TWAIN Working Group (www.twain.org), TWAIN is an application program interface that enables users to acquire images from a variety of sources, including CDs, hard disks, frame grabbers, and electronic databases. Scanner manufacturers no longer need to develop proprietary scanner drivers, and software developers do not need to develop interfaces for each scanner model the application supports. Version 1.9 of TWAIN was ratified in January 2000.

Document Imaging: Perspective

Price vs. Performance

The cost of imaging systems varies greatly depending on the configuration (e.g., the number of servers, scanning/capture workstations, and image workstations required), the number of concurrent users, and the complexity of the application (i.e., how much customization and ongoing administration is needed). Typical costs range from US$50,000 to millions of dollars. In the early days of imaging, systems were quite expensive— often reaching into millions of dollars. In addition, costs for back-file conversions and ongoing maintenance and support are often more than the initial system. Systems were hard to cost-justify for all but mission-critical applications. As vendors moved away from proprietary platforms and turnkey systems to distributed, client/server platforms and now Web-based architectures with thin clients, the costs began to come down. In addition, imaging technology itself has become a commodity, further driving costs down. Viewers are often included free of charge, and the cost of hardware devices, such as scanners, has decreased. Imaging became more affordable, hastening its adoption for office and ad hoc-type applications.

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 30,000 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 (e.g., 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 ActionPoint have carved out successful niches in the capture/transform subsystem market.

ActionPoint

ActionPoint, a public company formerly known as InputSoftware, develops and markets software for capturing documents (paper, faxes, or microfiche) and transforming them into digital content. Its core product, Input*Accel*, provides enterprise capable document capture and processing functionality. Core markets for ActionPoint include government, financial services, insurance, and manufacturing. The firm's customer base includes 300+ Global 2000 companies. In addition, ActionPoint has a strong partnership program; Input*Accel* is used by Documentum, Open Text, and IBM as the capture module for their document management solutions.

eiStream WMS, Inc.

Eastman Software, which was spun off from Eastman Kodak in September 2000, is now operating as eiStream WMS, Inc. eiStream WMS, Inc., is among the leading suppliers of production imaging and workflow software, and targets applications in financial services, insurance, and government. WMS Enterprise Edition: Imaging and Workflow has typically been implemented to support back-office applications with anywhere from 50 to 500 seats.

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.

FileNET Corp.

Since its inception in the early 1980s, FileNET has been known primarily for its integrated document and output management (IDOM) 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; Panagon eProcess Services and WorkFlo Services for workflow and process automation capabilities; 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; and Panagon Report Manager for accessing COLD data. FileNET's strategic direction for 2001 and beyond is to diversify its product line, targeting multiple markets with multiple products. In July 2001, FileNET announced a new product offering, Brightspire, an e-business framework for building collaborative commerce applications. In addition, FileNET is building vertically focused e-Business applications— the Acenza line— on top of the Panagon platform. Acenza extends business processes and associated content across the Web. The Acenza family currently consists of:

- Acenza Claims— designed for Web-enabled customer service and claims processing applications. Targeted to the Property and Casualty insurance market.
- Acenza Payables— for streamlining the processing and approval of vendor invoices and associated transactions.
- Acenza— Mortgage Lending— for streamlining loan origination and processing.

IBM

Since the 1980s, IBM has been a leading production imaging and workflow vendor with one of the largest installed bases (4,000+ 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 new product that combined the features and functionality of IBM's Digital Library media asset management system and its ImagePlus VisualInfo production imaging solution into a single digital content management repository running on DB2 or Oracle in Windows or AIX server environments. IBM positions the Content Manager brand as its enterprise infrastructure for e-content applications within the banking and finance, insurance, retail, telecommunications, and government markets. Content Manager handles content types ranging from scanned images and business documents to streaming audio and video. In addition, computer-generated output is an option by integrating Content Manager OnDemand. For data and document archiving applications, IBM provides the complementary CommonStore offering, which is available for SAP, Lotus Domino, and Microsoft Exchange applications. With the release of Content Manager v.7.1 in January 2001, IBM added basic document management functionality, including version control and support for ODMA. This move gives IBM the basic capabilities required to position Content Manager as a complete integrated document management (IDM) system for customers that want a one-stop solution from a single vendor.

TOWER Technology, Inc.

Founded in 1987, TOWER Technology has a rich background in document imaging systems integration and software development. The firm markets the TOWER IDM suite, a set of software modules for production imaging, document management, enterprise report management, and workflow (case management). It specializes in distributed imaging and workflow applications, as well as in production color imaging, for large-scale, enterprise deployment. 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 developed its own imaging, workflow, and enterprise report management technology, but partners with Staffware, Plexus, and Fujitsu for additional workflow choices and with 80-20 Software for the collaborative document management component of the TOWER IDM suite.

TOWER targets the insurance, financial services, healthcare, and utilities industries as well as government agencies. The market leader in Australia (with about 70 percent market share) and the United Kingdom, TOWER has been most successful in banking/financial services and in government installations, particularly in the United Kingdom and Europe. TOWER's strategic direction going forward is focused on supporting customer-facing e-business applications. Through the TOWER Document Portal and e-Process Objects middleware, TOWER provides a framework for integrating Web-based transactions and traditional business processes to enable customer self-service. IDM WebCapture captures transactional content associated with e-business transactions and stores it in the IDM Content Repository.

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 component software applications, software development tools, and hardware accelerator boards for building document imaging solutions. The firm's toolkits and accelerator boards are targeted at system integrators and independent software vendors (ISVs) that develop customized systems, while its component image capture and storage software line is marketed to value-added resellers that build document imaging solutions from best-of-breed software. 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, Hummingbird's DOCS Open, 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 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.