Guidelines

Digital Imaging Systems

Approved and Recommended by the State Records Commission
February 22, 2001

Office of the Secretary of State
INTRODUCTION

Organizations use digital imaging systems to create efficiencies and improve their effectiveness. Imaging gives them the ability to capture, store, retrieve, and share an enormous number of records over their networks. Users can typically find a document on an imaging system faster than they can find the paper or microfilm version. They can easily share documents using workflow software or e-mail. Organizations may even find a decreased need for filing cabinets and box storage space. Though this last point is predominantly emphasized, the true benefit of digital imaging comes from the first two – on-line access to documents and sharing of valuable information.

Is Imaging Right for the Project

The decision whether to implement an imaging system is complex. Many issues must be considered. Primarily, what is the desired outcome? How will imaging solve user problems? Will it meet real needs? How will it integrate into the existing computer environment? Are there sufficient financial resources to support the system over time? The decision to implement an imaging system must be based on the business need. The key to the successful design, integration, and implementation of a digital imaging system is proper analysis. The Secretary of State’s Office has created these guidelines to assist agencies and local governments with this analysis. They are intended to offer practical advice as you tackle the three major phases of system development: project planning; technology assessment and selection; and system implementation. Each phase is addressed in a separate section of the guidelines. By using the guidelines, you will gain a better understanding of the opportunities, challenges, scope, and scale of a digital imaging project.
**PROJECT PLANNING**

An examination of the following issues and analyses is among the first steps to take when considering digital imaging as an option for your organization.

**Legal Issues**

The laws governing the creation and use of records may affect your decisions. Statutes and administrative regulations may define how records are created and what media on which they may be stored. You need to develop strategies to meet your legal requirements including compliance with your approved records disposition schedule. We strongly recommend consultation with your organization’s legal counsel well before implementation of an imaging system.

**Feasibility Study**

Before the decision to purchase an imaging system is made, a thorough feasibility study should be conducted. At a minimum, the following analyses should be included:

1. **Needs assessment.** Conduct a needs assessment to determine what, if any, benefits you will gain by using an imaging system. An understanding of the total business process is critical to the project’s success. A good starting point is to define and analyze the existing document workflow. Look for any problems and look for opportunities for improvements particularly in the areas of document storage, retrieval, and access. The knowledge gained in the needs assessment phase will prove invaluable in the system design and product selection phase.

2. **Customer satisfaction.** You need to consider how both your internal and external customers will or will not benefit from the new system. How much resistance to change is acceptable? How will satisfaction be measured?

3. **Alternatives assessment.** Microfilming should be among the alternatives you research. Microfilm and digital imaging each yield different, yet excellent benefits. For example, imaging is more cost effective than microfilm when retrieval rates are high. However, microfilm is ideal for long-term records retention and preservation. When comparing alternatives, be certain to consider on-going costs to determine the cost of ownership over time.

4. **Economic feasibility and cost/benefit analyses.** You should determine whether expected cost savings, productivity improvements, and other benefits outweigh the initial purchases and on-going expenses including hardware, software, network, office space, and staffing considerations.

5. **Technological feasibility analysis.** The first step is to verify that reliable hardware, software, and storage media can be acquired or developed and will integrate with your existing systems. The second step is to verify that you will
be able to migrate to newer technology as necessary. In other words, the rapid rate of technology change must be considered and built into your plan. (More information on migration may be found on page 19.) The use of open systems and non-proprietary solutions will help to ensure the continued viability of the imaging system. Additionally, the use of open systems will increase the numbers of IT professionals and consultants capable of performing work on the proposed imaging system.

6. **Operational feasibility analysis.** The goal is to discover the level of the willingness and the ability of management and staff to operate, use, and support an imaging system.

**Budget Issues**

The total cost of an imaging system includes not only the initial purchase of hardware, software, and technical support, but also on-going costs such as the following:

- **Storage media** – including optical disks, hard drives, network servers, etc.
- **Maintenance contracts** – this includes preventative maintenance, hardware repairs, software upgrades, and telephone technical support (typically this expense runs between 15% to 20% of the original purchase price per year).
- **On-site technical support.**
- **Labor.**

You must also consider the periodic costs associated with ever-changing technology such as the following:

- **Refreshing media** – every three to five years depending on the media (more information on refreshing media may be found on page 19).
- **Technology upgrades** – as necessary.
- **Replacement of obsolete hardware** – every three to five years is a good rule of thumb.
- **Migration** – as necessary.
- **System documentation** – reflects changes as a result of upgrades and migration.
- **Training** – new staff and retraining staff on new systems.

There are other costs you should consider. The consequences of market change may be difficult to quantify. Nonetheless, market changes are inevitable. Newer technology drives out older technology. How many PC’s in your office still have a 5-¼” floppy disk drive? The same fate will happen to one or more component of your imaging system. You must budget for change.

Before selecting a vendor, you should research the company’s stability and reputation. You want to ensure that the company has provided others with excellent
products and services. You also want to ensure, as much as you can, that the company will remain in business as long as you operate your imaging system.

Unfortunately, system failure is always a possibility. The system may be down for several hours or even several days. We recommend that you have a contingency plan in place. The plan should provide answers to these questions: How fast can the vendor be on-site for equipment repair? What can users do to get by while the system is down? What will support staff do while the system is down?

Records Management Issues

Many records management issues must be considered when planning an imaging system. For instance, every record series has a disposition schedule stating how long records must be retained. A records management plan must be in place and followed scrupulously to ensure that information is kept and remains accessible as long as its retention schedule states is necessary.

Procedures manuals and other system documentation should become a part of your organization’s approved records disposition schedule. Addressing these concerns at the design stage and putting in place the proper procedures from the beginning will ensure routinely managed retention and destruction of records on the system.
DIGITAL IMAGING TECHNOLOGIES

One of the greatest challenges facing managers and administrators today is the continued viability of their systems and the preservation of their records. To meet this challenge, they must employ strategies that transcend rapid technology changes. The success of such strategies depends upon several factors including open systems architecture, component upgrades, stable storage environment, and accurate data transfer. The goal of this section is to give agencies and local governments the information tools they need to design systems to ensure long-term functionality.

What is Imaging

Imaging is the process of converting human readable media, such as paper or microfilm, into information that can be stored and retrieved electronically. Basically, an image is a digitized picture of a document, drawing, or photograph. Another way to think of it is, an image is like a photocopy that you view on your computer.

An imaging system’s components are interdependent. Hardware, software, compression techniques, file formats, and media all have to work together. Each component does this using interfaces and drivers. While these topics are not always fascinating reading, each is an essential element. You need a basic understanding of what each does to make an informed decision.

Using open systems architecture as much as possible is the best method to ensure system viability over time. A system is a combination of components working together. The system’s architecture is its design or configuration. An open architecture means the system’s components use standards or specifications that have been made public by their designers. The use of standards and publicly available specifications mitigates the impact of incompatibility. Technical staff can upgrade hardware and software with minimal impact on the overall system and without significant risk of data loss. Thus, the integrity of the system is more likely ensured by using open systems architecture.

Recommendation: Use open systems architecture or require vendors of proprietary systems to provide a method to migrate to non-proprietary configurations.

Hardware

Imaging hardware performs three basic jobs. It captures, stores, and retrieves digital images. The results of the needs assessment and workflow analysis will guide your hardware decision-making process. When configuring the system, plan for your peak workload. In other words, if time is a factor, configure the system’s capacity to handle the highest volume you will need to push through it on a given day. Also, plan for predictable expansion. In other words, if you know your workload increases 10% per year, configure the system so that it can meet your needs for the near future (three to five years). The essential thing to remember is purchase what is needed, not what is perceived to be the latest and greatest. Bear in mind how much funding is
available for implementing a new system and how much will be available in the future. The more complex the system, the more costly it will be to implement and to maintain.

**Capture**

Typically, when people think of a digital imaging system what immediately comes to mind is scanning. Within the scanning process are several components and considerations. Among these are scanner types, feed-method and throughput, and resolution. Also included in this section is a discussion on hardware interfaces and software drivers as both are an integral part of the capture process.

1. **Scanner types.** There are two broad categories of scanners – document and graphics. As the name implies, document scanners are designed to capture typical black-and-white or grayscale documents. Graphics scanners, conversely, are designed to capture color.

2. **Feed method and throughput.** The feed method is literally the way paper or microfilm is moved across the scanner. When selecting the appropriate feed method, an important consideration is the condition of the documents. Frail documents should not be run through an automatic document feeder. Conversely, high-volume documents, such as checks, should not be placed individually on the flatbed. Throughput is the rated speed at which the scanner can process pages. For example, a scanner rated at 60 ppm should be able to process 60 pages per minute. Vendor claims of throughput are useful for comparison; however, throughput on real-world documents will vary somewhat. In fact, once you begin production, you are likely to find the scanner’s throughput to be less than the its rated speed.

- Flatbed scanners consist of a platen glass on which you manually position books, magazines, and other documents that you want to scan.
- Automatic document feeders (ADF) allow stacks of pages to be fed through the scanner.
- Simplex captures one side of a page at a time.
- Duplex captures both sides of a page in a single pass.
- High-speed can scan several hundred pages or more per hour.
- Film readers are very specialized and require attachments to scan microfilm and microfiche.
- Fax machines can be connected to the imaging systems so that faxed images can be imported into your indexing and retrieval software.

**Recommendations:**

- Choose a scanner appropriate for the condition of your documents.
- Plan the scanning process for your peak load and plan for expansion as needed.
3. **Scanning resolution.** Scanning resolution is measured in dots-per-inch. This is literally a measurement of the dots per square inch of an imaged document. If you were to look at an image under a microscope, you would see that each character is made up of many, many dots. Scanners are capable of supporting resolutions from 72 to 8800 dpi. For example, if an image is captured at 600 dpi, it has 360,000 dots per square inch. Typical office documents generally can be accurately captured using a scanning resolution of 200 dpi. More detailed documents, such as maps or drawings may require a higher resolution. Also a higher resolution, generally 300 dpi, is necessary to produce a relatively legible text file. There is a tradeoff between image quality (legibility) and storage. The higher the scanning resolution, the larger the image file.

**Recommendations:**
- Scan office documents no lower than 200 dpi.
- For data capture using OCR/ICR scan no lower than 300 dpi. (More information on OCR and ICR may be found on page 12.)

**Storage**

Storage systems can be broken into three categories: on-line; near-line; and off-line. Which type of storage system you select will depend upon how fast and how often you must access the images.

1. **On-line.** On-line storage means the image file can be displayed in a matter of seconds. Either the file is stored on a hard-disk or the media is in a system drive. RAID (which is discussed in the Media section of this document) is considered an on-line system.

2. **Near-line.** Near-line storage means the image file can be retrieved by the system and loaded into a drive to be read. Optical jukeboxes (including CD jukeboxes) are considered near-line systems.

3. **Off-line.** Off-line storage means the image file resides on media (such as microfilm, optical disks, data tapes, etc.) that requires human intervention to load into a system drive.

**Recommendation:** Consider retrieval-time requirements when selecting storage options. Depending on your access requirements, you may want to consider a hybrid system – on-line access for active records, near-line for semi-active, and off-line for inactive.

**Interfaces and Drivers**

Interfaces connect devices, such as scanners and other peripherals, to each other and to the system software. In scanning, the interface controls how fast data is moved from the scanner to the software. The common interfaces are video, serial, parallel, and SCSI. (SCSI, pronounced *Scuzzy*, stands for Small Computer Systems Interface.) Video interfaces tend to be used in proprietary systems. Through
customization, video interfaces can enhance scanning performance. SCSI is an industry standard parallel interface. It can support very high transfer rates. Serial and non-SCSI parallel interfaces that cannot support high data transfer rates tend to be used with low-end scanners. These scanners typically have a lower throughput than those using video or SCSI interfaces.

Recommendation: Use the interface that is appropriate for your system. Fully document all customization. If proprietary interfaces are employed, budget for programming when the need to migrate occurs.

Software drivers are another means of communication between devices and software programs. Each device and each program has a set of commands that it understands. Drivers interpret the commands from the program to the device and from the device to the program. The standard software drivers that digital imaging systems use are TWAIN (Technology Without an Interesting Name) and ISIS (Images and Scanners Interface Standard). TWAIN drivers support basic functions. ISIS drivers can be customized to take advantage of all the scanners features. Systems using TWAIN tend to experience slower throughput than those using ISIS.

Recommendation: Use the interface that is appropriate for your system. Fully document all customization. If proprietary interfaces are employed, budget for programming when the need to migrate occurs.

Retrieval

Once the image is stored, it must be retrieved for viewing and/or printing. The monitor and printer must be able to render the image as closely as possible. The performance of these devices has seen dramatic improvements and the costs have dropped considerably in recent years.

1. Monitors. The resolution at which your monitor can display an image is measured in picture elements or pixels. A pixel is a single point in a graphic image. The relationship between scanning dpi and display and print resolution is not one-to-one. Display quality largely depends on the monitor’s resolution, that is, how many pixels it can display. The higher the display resolution, the more colors the monitor can display. Bear in mind, however, the display color may not be important for the task. A monochrome monitor may be appropriate for systems that do not use color. On the other hand, users tend to prefer color monitors to monochrome and user satisfaction is a critical consideration. Another important consideration is the viewable display size. Image viewing is not easily done on small monitors. Eyestrain and frustration set in early. If users will retrieve and view images often throughout the day, a large monitor is justifiable. If they will retrieve images sporadically, a mid-size monitor may be acceptable.
Recommendations:
- Whether monochrome or color, select a monitor capable of high-resolution display.
- Select at least a 17” monitor for infrequent retrievals.
- Select at least a 19” monitor for frequent retrievals.
- Select a 21” monitor for continuous retrievals.

2. Video/Graphics Adapter. The video adapter is a card that plugs into the computer. Since monitors are analog devices, they require an adapter to translate the digital information into an analog signal. You can spend hundreds of dollars on a high quality monitor but it will only generate an image as good as the video adapter.

Recommendation: Review monitor specifications and system compatibility carefully.

3. Printers. Two broad categories of printers are laser and ink jet. Laser printers fuse toner onto paper to produce images. Ink jet printers use wet ink. The ink has to be placed in specified locations which means the jets have to pass over the page multiple times to produce the image. Laser printers generally produce higher resolution images faster than ink jet printers. No matter what type of printer is used, the print process can be extremely slow because uncompressed image files are so large. A print accelerator board can speed up the process dramatically. The accelerator board is installed in the printer. The board instructs the imaging system to send the image file to it in a compressed format. The board then decompresses the image and prepares it for printing.

Recommendation: Consider print resolution and throughput requirements when choosing a printer.

Software
Two broad software categories are operating systems (OS) and applications. The computer’s OS performs basic input and output functions. Microsoft Windows™ and Apple Macintosh™ are two well-known operating systems. The OS provides a platform on top of which applications, such as digital imaging systems, can run.

There are innumerable software application choices. Application programs must be written specifically to run on a particular OS. Which OS is used is dependent on the system hardware which includes both network servers and user workstations. This can be an issue when future migrations occur.
Application software must perform certain tasks such as the following:

- **Scanning.** This function is required to operate the scanner. It translates the image captured by the scanner into a graphics file format capable of being stored.
- **Indexing.** This information is necessary to identify the image. Indexing information is stored in the database.
- **Database Management System (DBMS).** The DBMS catalogs the location of the stored image on the storage media and adds that information to the index information created by the indexing application. Whatever DBMS is procured, it should understand Structured Query Language (SQL). The prominent DBMSs on the market today can do so.
- **Image Retrieval.** This application must be capable of retrieving the image by querying the database for its location, calling the image to a monitor or a printer, and converting the file into a format that device can use.

*Recommendation:* Choose the OS and applications that satisfy the requirements of the system. Bear in mind that your existing network infrastructure may influence the decision-making process.

### Compression

Image files are quite a bit larger than text files. For instance, a two-paged document created and stored as a word processing file might be 20 kilobytes. The same document, printed, scanned, and stored uncompressed might be as large as one megabyte. Because image files are so large, they are usually compressed. That way they do not require as much storage space on the computer system and they can be transmitted over the network faster. However, they then must be decompressed for viewing and/or printing. The compression technique you use depends on whether the image is going to capture color or not. Although proprietary techniques exist, their use is not recommended for retention or preservation as they will severely limit your ability to migrate to newer technologies.

1. **International Telecommunications Union (ITU, formerly the International Telegraph and Telephone Consultative Committee - CCITT) Group III and Group IV.** These techniques are used to compress images of black-and-white documents. Each serves a different purpose. Group III is specified for analog transmissions (standard facsimile machines) and Group IV is specified for digital transmissions (computer networks, digital facsimile machines, etc.). While Group III compression ratios are generally up to 10:1, Group IV typically ranges from 15:1 up to 20:1. Therefore, Group IV compression creates smaller files.

*Recommendation:* Use CCITT Group IV compression. Use Group III as needed.
2. **Joint Photographic Experts Group (JPEG).** JPEG is designed for compressing either full-color or grayscale photographs. Compression ratios can range up to 100:1 using JPEG. Within JPEG, there are two broad categories of compression: *lossy* and *lossless*.

   - *Lossy* compression determines whether a color is useful (different enough from other colors to be visually perceived). If the color is useful, it is kept. If it is not useful, it is dropped or lost. Therefore, the decompressed image is not an exact replica of the original.
   - *Lossless* compression does not drop colors; however, the ratios are not as high as with lossy compression.

   **Recommendation:** Carefully consider how the images will be used. If computer analysis is anticipated, lossy compression will not render 100% accurate results.

3. **Proprietary formats.** Proprietary compression techniques may offer benefits such as higher compression ratios. Thus, the image files may require less storage space and can travel through the network more quickly.

   **Recommendation:** Use proprietary compression techniques for daily use only. Require additional non-proprietary compression for retention, preservation, and back-up copies.

### File Formats

The file format tells the computer how the information within the file is stored.

1. **Tagged Image File Format.** TIFF is the most commonly used format for digital imaging systems. The file contains a tag or header that stores information (metadata) about the file. The header contains information such as how the image is compressed. It is important to understand that not all TIFFs are alike. A TIFF from one vendor’s system may not be compatible with another vendor’s system. Some systems will allow TIFFs to represent single-page or multi-pages documents. For example, 12345.tif could be a one-page document or it could be a 120-page document. In such systems, the operator chooses the setting (single-page v. multi-page) before scanning. File errors can be a major problem with multi-page TIFFs. If a multi-page TIFF file is corrupt, every page in the document is rendered inaccessible. However, if a document is made up of single-page TIFFs and one of the TIFF pages is corrupt, then only that page is inaccessible. Users can still access the other TIFF pages. Therefore, multi-page TIFFs are not recommended for retention, preservation, or back-up.

   **Recommendation:** Use single-page TIFFs for retention, preservation, and back-up.
2. JPEG. Although JPEG is a compression technique, files created using it have a .jpg file extension.

*Recommendation:* Carefully review the system’s file format. Ask the vendor to provide proof that its files can be transferred to another vendor’s system. Budget for programming fees when migration is necessitated.

3. Proprietary formats. Vendors may develop proprietary formats for a variety of reasons and with various benefits.

*Recommendation:* Use proprietary formats for daily use only. Require additional non-proprietary formats for retention, preservation, and back-up copies.

4. Text files. Text files are created from image files by using Optical Character Recognition (OCR) or Intelligent Character Recognition (ICR) software. OCR creates computer readable text from a digital image. ICR is often used for forms recognition. The system is set up to recognize key fields as containing specific information. ICR converts that information into text that can be imported into a database.

**Media**

When designing a storage solution as part of a digital imaging system, you must consider several factors including retrieval-time requirements, records retention and preservation, and costs over time.

**Fixed Storage Media**

Fixed storage is so named because the media stays within the computer or drive. The imaging system uses the computer’s hard drive, or magnetic disk, throughout the conversion, storage, and retrieval processes. The hard drive’s performance, capacity, and reliability are critical elements to consider when selecting a hard drive. From time-to-time, hard drives will fail. Therefore, back-ups are critical. A back-up contains copies of the information on the hard drive.

Random access memory (RAM) is also important. Imaging programs use RAM to process information and data. Imaging programs tend to be more intense users of a computer’s memory, therefore, the more RAM, the better. In the computing world RAM is expressed in terms of megabytes such as 64 MB or 128 MB.

Another fixed storage option is RAID (redundant array of independent disks). There are several levels of RAID. Each has different options and benefits. In general, RAID can be used to increase on-line capacity and speed retrieval while providing a good back-up strategy.
Recommendations:

- Select a hard drive that meets or exceeds the vendor recommendations.
- Select RAM that meets or exceeds the vendor recommendations.
- Consider RAID to increase on-line capacity and increase retrieval speeds.
- Adhere to routine back-up schedules.

Removable Storage Media

Removable media, as its name indicates, can be removed from the imaging system. Three general categories of removable media are optical disks, magnetic tape, and microfilm. Factors to consider when selecting removable media include capacity, life-expectancy, and cost over time. Two important cautions about media life-expectancy or shelf life are in order. First, the terms refers to how long the physical media will last in perfect environmental conditions. When you read that a medium has an expected shelf-life of X-number of years, be aware that the prediction is based upon laboratory-controlled conditions. While microfilm has been in business use since the mid-1920’s, optical media is still relatively new. The actual life-expectancy of a given medium has not been proven yet. Second, shelf-life does not refer to how long the information on the media will remain viable. Thus, error testing and migration strategies are essential to ensure the viability of records whose retention periods are longer than five years.

1. Optical disks. Lasers are used to write data onto an optical disk. Optical disks, typically, have high storage capacity. There are three major categories of optical disks used in imaging systems:

- WORM (Write Once, Read Many) optical disks have large storage capacities – generally in measured in terms of gigabytes (thousands of megabytes). Once information is written to the disk, it cannot be erased. Data is written to the disk sequentially. To improve efficiency, writing can be done intermittently or in batches. Because data is written sequentially, it can be read back from the disk faster than the same data on a CD-ROM. The indexing information the software uses to retrieve the image can be deleted. This makes it virtually impossible for the average user to retrieve the image. However, the image is still on the WORM disk. Sophisticated computer sleuths can retrieve images whose indexes have been deleted. Shelf life of WORM media is typically up to 25 years.

- Magneto Optical (MO), erasable, or rewritable optical disks also have large storage capacities, however, images can be deleted from the disks. Since MOs allow multiple write sessions and erasures, the media tends to degrade quickly if many write/rewrite sessions are performed. Shelf life of MO media is typically up to seven years.

- CD-Recordable (CD-R) and CD-Rewritable (CD-RW) are arguably the best known type of optical disk. A CD-R becomes a CD-ROM once data is written to it. Therefore, like WORM disks, data cannot be erased from CD-R disks. Typically, CD-Rs hold about 640 to 700 megabytes. Data is written to the disk spirally and is done in one session. In other words, if
you only write 16 megabytes to the CD with a capacity of 700 megabytes, then you have 684 megabytes of wasted space on that CD. CD-Rewritables are slightly different. You can do multiple writing sessions and erase data. Since CD-RWs allow multiple write sessions and erasures, the media tends to degrade quickly if many write/rewrite sessions are performed. Shelf life of CD media is rated up to 30 years under laboratory-controlled conditions. Although CD-RW shelf life can be 30 years, it is typically less if many write/rewrite sessions are done.

- Digital Versatile Disks (DVD) are much like CDs. They come in both recordable and rewritable formats. DVDs can store up to 12.4 gigabytes of data. That is about 20 times that of a CD. Shelf life is the same as CDs. However, as of this writing, we are unaware of any imaging systems that use DVDs.

2. **Tape.** Magnetic tape is useful for infrequent retrievals and for back-ups. Tape is generally slower than optical media is. Tapes can hold anywhere from one to ten gigabytes of data. The shelf life for magnetic tape is up to seven years.

3. **Microfilm.** Like tape, microfilm can be useful for infrequent retrieval and back-ups. Some imaging systems can digitize existing microfilm and others can create microfilm from images stored on the system or on optical disk. Shelf life for microfilm is rated from 100 to 200 years or more.

**Cost**

Another inverse relationship to consider exists between storage capacity and cost per megabyte. For example, the faster the medium, the fewer records it can hold, and the more it costs per megabyte to store those records. Conversely, the more records a medium can hold, the lower the cost per megabyte, and the slower your access to the records is.

**Recommendations:**

- Permanent records should not be stored on rewritable media.
- For retention and preservation, store permanent records on microfilm
- Once the records have met their retention period, the media should be physically destroyed. Therefore, records with similar retention periods should be stored on the same disk, tape, or reel.
- If a record or record series is identified to be a part of litigation and that record or record series resides on a disk, tape, or reel that is scheduled for destruction, a strategy should be in place to copy that record or records series from the disk, tape, or reel to another. This is necessary to ensure that records identified for litigation are not inadvertently destroyed.
**SYSTEM IMPLEMENTATION**

Implementing an imaging system requires careful planning of your resources and of the imaging system. You should have in place a plan and budget for updating and operating the imaging system. You should identify legal issues regarding access to and integrity of your records in an imaging system. Additionally, you should have a plan in place for records storage and migration.

**Staffing**

Successful implementation of an imaging system requires the cooperation of many different internal and external resources. Internally, staff must be hired to operate and maintain an imaging system. The system administrator must have knowledge ranging from system design and maintenance to daily operations. Imaging operators must have a general knowledge of computers and of imaging systems. They must also have a good working knowledge of the organization and the records it creates. Processing staff will be needed to prepare the documents for scanning. Document preparation consists of organizing the documents in a logical order, removing staples or paper clips, repairing tears, flattening wrinkles and creases, etc. This process is necessary to ensure pages do not get caught in the scanner.

The imaging system staff must have a good working relationship with the organization’s technical support staff. The support of the technical staff is essential to the successful operation and maintenance of your imaging system. Technical staff can advise on compatibility issues, data integrity, and system upgrades.

Externally, the system vendor must be able to provide training, maintenance, and ongoing support in operating your imaging system. Both the imaging staff and the technical support staff must have a good working relationship with the vendor’s technical support staff to ensure the smooth operation of your system.

**Legal Issues Revisited**

The laws governing the creation and use of records may affect your decisions. Statutes and administrative regulations may define how records are created and what media on which they may be stored. You need to develop strategies to meet your legal requirements including compliance with your approved records disposition schedule. We strongly recommend consultation with your organization’s legal counsel well before implementation of an imaging system.

Whether imaged or not, the ability to use records as evidence for legal, audit, and other purposes depends on establishing their authenticity and reliability. First, your organization must be able to prove that a recordkeeping system is used as part of the normal course of business. This can be accomplished by documenting the specifications of the imaging system, training staff in the operation of the system, ensuring the integrity of the records, and conducting random audits. Second, state agencies and local governments must provide access to public records as mandated in Missouri's Sunshine Law.
1. **Imaging system documentation.** You should maintain a record of information about your system. This documentation should become a part of your approved records disposition schedule. This document should contain, at a minimum, the following information:

- All system equipment specifications.
- A description of all hardware and software upgrades to the system including date of maintenance and version of software.
- Contact information for manufacturers and vendors.
- Technical operation manuals.
- User operation manuals.
- All policies and procedures related to access to and security of the records.

2. **Training.** You must be able to show that the imaging staff have been trained in operating the system. Also, you must be able to prove that they follow the normal recordkeeping practices established by your organization. Your human resources department may be able to help you develop training records and guidelines. Such records should be included on your approved records disposition schedule.

3. **Integrity.** Integrity of electronic records refers to both the physical and intellectual integrity of the information. Maintaining the physical integrity is concerned with two issues:

- *The actual condition of the media storage device* – questions such as has the media deteriorated, been scratched, or has it been exposed to extreme temperatures need to be assessed.
- *The reliability of the record after compression or migration* - after such events, has the appearance of the document been altered in a way that could be misleading?

The intellectual integrity of a record is based upon the authenticity or truthfulness of the information within the record. A system should be in place for electronic records that validates access procedures and documents modifications to the records over time.

To establish the integrity of the records, standards should be in place that ensure:

- The identity of a record’s creator is verified.
- Permission to read, write and delete files is appropriately restricted.
- Periodic system audits are conducted.
- Data transmission includes data error checking and correction.
- Data are regularly backed up.
4. **Auditing.** You must conduct periodic and random audits of the imaging system to ensure that the system is operating within the established records management guidelines and that the data remains viable. Prior to beginning an imaging project, management should establish both acceptable error limits and procedures for correcting systems which do not meet those limits. Documentation of audits should be kept with the imaging system documentation and should become a part of your approved records disposition schedule.

5. **Access.** Unless otherwise closed by statute, the Missouri Sunshine Law requires the public records of state and local governments to be made available to the public upon request. To ensure the records are easily accessible throughout their retention period for internal, as well as public use, the recordkeeping system must:

- Provide for clear identification of the record.
- Permit easy and timely retrieval of individual records and record series.
- Retain the records in a usable format.

**Indexing and Labeling**

Records constitute a corporate resource. To assure that the resources remain accessible, an indexing database which facilitates efficient retrieval, ease of use, and up-to-date information about the images stored in the system should be developed. To ensure that the value of the information in the record is maintained and can be retrieved from an electronic record, three elements must be present:

- **Content** refers to the subject matter of the record, while
- **Structure** focuses on use of fonts, headings, spacing, etc., as part of the meaning of the content, and
- **Context** refers to the relation of one record to other records.

The content, structure, and context a system captures in the index and label component is commonly referred to as metadata. Metadata is essentially “data about data,” it describes an information resource. It is important to capture the image metadata to allow future users to discover the context in which the record was produced and to permit the owner of the record to manage it.

There are several national and international organizations working toward establishing indexing standards. There are many metadata schemes, including but not limited to, the Dublin Core (DC), Encoded Archival Description (EAD), Text Encoding and Interchange (TEI), and Machine Readable Cataloging (MARC), which are being used at various institutions today. If you choose not to adopt one of these standards, then at a minimum, the metadata you capture should include:

- **Title** – Of the document
- **Creator/Author** – Including other contributors
- **Date** – Data of Document Creation/creation of image
Unique identifier – Such as a coding system for different document types,
Format - To include operating system, software configurations, and versions thereof
Definition - A statement that clearly represents the concept and essential nature of the record
Rights/Security – Indicate if special authority is needed to access the information, and who has that authority
Datatype - Indicates if it is a written document, photograph, etc.
Keywords – use a controlled vocabulary thesaurus if possible. Several industries have published their thesaurus on the World Wide Web.
Comment – For additional information

Whatever media you chose for your imaging system, make sure that you clearly label all disks, tapes, or other removable storage containers, with enough information that the contents of the storage instrument can be determined years later, by different staff. At a minimum, the label should include:

- Identifiers - including creator, date created, division or agency where created
- Hardware, operating system, and software required to access a document
- Encoding standard and version
- Level of security or restricted access
- Retention dates of the information on the media

If the disk or other format is too small to include all of the information on the label, then establish a coding system that can be linked back to an index which holds all of the vital information. Documentation relating to the coding system and index must be maintained for as long as it relates to any labeled storage medium which utilizes that coding system, and should become a part of your organization’s approved records disposition schedule.

Quality Control

It is essential to the success of the imaging project that the image produced is the best image possible. Imaging operators should test the system at least at the beginning of each series of similar documents by scanning a test target. Test targets can be obtained from the Association for Information and Image Management. (See appendix) The image operator should also monitor the scanned images. If the quality is poor, then the system operator must rescan the documents or pages and possibly adjust the scanning parameters. The frequency of quality control monitoring should be determined by the type of the documents, but should not be less than once during each session.

Once the images have been captured and placed on the storage medium you will need to monitor the maintenance of the storage medium to ensure the records are accessible and their integrity intact. The type of storage medium you choose will determine how you will set up your quality control guidelines for storage. You should check with your vendor for recommendations for winding magnetic tape. A
schedule and percentage base for random sampling of disks for signs of deterioration and corrupt files should be in place.

**Security Copies**

Producing a security copy of the output serves to protect your organization in the event the working copy becomes damaged, lost, or destroyed. We recommend you store security copies in an off-site location. The security copy is your record copy and must be included in your approved records disposition schedule. The on-site copy is a convenience copy.

**Storage**

You should adhere to vendor’s recommendations for specific environmental conditions in which the media should be stored. These recommendations should include information relating to ideal temperature, humidity, and storage orientation. Both the working and security storage locations should adhere to these recommendations. If the recommended storage conditions cannot be met, you should ensure that the environmental conditions of the storage location are stable.

In addition to the guidelines for the physical storage of the media, it is vital that guidelines be in place for the storage and maintenance of the records on the media. As a rule electronic media are less stable than paper. In addition, unlike paper or microfilm, digital images are not eye-readable without the assistance of computer hardware and software. Because digital storage media is not permanent, and rapid changes in computer technology are constant, a conversion strategy for retaining and retrieving stored information should include refresh and migration guidelines.

1. **Refresh.** Refresh involves copying information from one storage medium onto a newer more stable storage medium. For example, transferring the information stored on a three-year-old WORM disk to a new WORM disk. It is recommended that the media should be refreshed every three to five years, depending on the media type.

2. **Migration.** Migration is the process of transferring digital information from one storage format to another, or from one generation of hardware and software to the next. For example, transferring data from 5¼ inch floppy disks to a CD-ROM, or upgrading from imaging software version 3.0 to version 4.0. Currently, migration is the best practical means for retaining and retrieving data over time. Migrations have to be carefully planned, executed, and audited to ensure against data loss. Though migration is a time consuming and expensive process, with proper guidelines in place, the costs can be minimized.

An additional factor which must be considered if a migration strategy is adopted is copyright laws. Currently, the same copyright laws apply to both printed materials and electronic media. Since most software is copyrighted, make sure you negotiate with vendors for the rights to the data you have created and to have the ability to migrate necessary software components to be able to access your data.
Another option for avoiding copyright infringement is to use open systems architecture.

**Retention Schedules**

The fact that you are storing information in an electronic format does not impact the decision as to how long you will retain that record. It is the information contained on the medium, whether paper or electronic, which must be considered. As a governmental body you are required by law to retain and to make available, certain types of records for a determined period of time. Each record series should have a defined retention schedule. If you are uncertain of the retention schedules for your agency contact your records custodian or the Records Management division of the Secretary of State’s Office.

For the purpose of this document retention periods will be defined as:

- Short-Term – creation to five years
- Medium-Term – six to nine years
- Long-Term – ten years and longer

Guidelines for each period will need to be in place. You will need to determine how often you will refresh and/or migrate the records, if different quality control measures should be used for different retention periods, and if you will store all the records in one location or separate locations based on the retention periods.

After a record has met its retention period, then your agency should determine the disposition of the record. If you will be destroying the records it is important that the plan for destruction is in place before you implement the imaging system. Depending on the type of storage system you choose, this may mean destroying disks or magnetic tapes. By placing like records on the same storage disk you will be able to destroy records with the same retention periods. Otherwise, you would have to migrate records that are not eligible for destruction to another medium first. You must develop standards that ensure the record is completely destroyed once the retention period has been met.

**Preservation**

Preservation in the digital world means ensuring continuing access to high-quality, eye-readable original source documents. Permanent records will require a well developed migration strategy and the most diligent efforts to keep them accessible. As much contextual information as possible must be captured to ensure the historical meaning of the image is not distorted.

At this time we recommend microfilm for permanent retention. With the proper equipment, digital images can be scanned to and created from microfilm. Microfilm’s longevity is proven, it does not require software to read, and it is a cost-effective alternative to refreshing optical or magnetic media.
**Risk Management**

As part of your initial planning phase a comprehensive disaster management plan should be developed. The plan should include standard back-up and recovery procedures, as well as quality control and storage procedures such as those mentioned previously. In the case of a natural disaster, having off-site copies of records may be the only answer for recovering data. A test of both the prevention and data recovery guidelines should be conducted on a regular basis.
**FINAL RECOMMENDATIONS**

The Imaging Guidelines provide an introduction to some of the many issues that you will face when planning, selecting, and implementing an imaging system. The guidelines are intended to be used as a starting point for the decision making process. Because of the number of issues involved, no single set of guidelines will be able to answer all of your questions. Regardless of the issues that you will face, project planning is the key to the successful selection and implementation of an imaging system. If you have additional questions, please contact the Secretary of State’s Office, Records Management Division.
Suggested Reading

The Records Management Division reviewed numerous resources while preparing these guidelines. We recommend these sources for your information gathering pursuits. We suggest that you use these sources for evaluation. However, we do not endorse any products, vendors, or claims.

A note about the citations: dates appearing after Internet addresses reflect the date the material was accessed – not the date of the material.

Project Planning

“Electronic Imaging Regulations Draft (Version 1).” Nebraska Secretary of State.


“Electronic Records Draft Guidelines.” Mississippi Department of Archive & History.


“Guidelines for the Use of Digital Imaging Technologies for Long-Term Government Records in Alabama.” Alabama Department of Archives and History.

“Guidelines for Determining if a Stand-Alone Imaging System is the Best Choice for You.” New York State Archives and Records Administration.


Digital Imaging Technologies


Georgis, Steven. “The Advantages of Helical-Scan Over Linear Recording Technology.” Mammoth Tape Technology, 

“Hardware or Software Based RAID, Which Solution is Best for You?” Adaptec, Inc. 

Hodge, Gail and Carroll, Bonnie C. Digital Electronic Archiving: The State of the Art and 


http://www.state.sc.us/scdah/113_odp.htm (30 December 1999).

“Informix and RAID, a Discussion Document.” Baydel. 


“Linking Applications and Images, What is TWAIN?” Hewlett-Packard Company. 

“Mechanism of Magneto Optical (MO) Disk.” Teijin Recording Media. 


Messmer, Hans-Peter. The Indispensable PC Hardware Book, Second Edition. Harlow, 

“Primer on DLTtape Technology.” *DLTtape Technology, Quantum Corporation.*

“RAID Definitions.” *Adjile Systems.*


Sandgren, Mattias and Markus Westergren. “SCSI and IDE, Overview and Comparison.” *Institute of Computer Science, Umeå University.*


**System Implementation**


AIIM TR21-1991. “Recommendations for the Identifying Information to be Placed on Write-Once-Read-Many (WORM) and Rewritable Optical Disk (OD) Cartridge Label(s) and Optical Disk Cartridge Packaging (Shipping Containers).” *Association for Information and Image Management.* (1991).


