E-Records Guidance and Best Practices

I. Electronic Records consist of:
   A. Images created from scanned paper documents
   B. “Born digital” materials that are created electronically such as word processing documents, spreadsheets, and presentations as well as emails, databases and business applications.

II. Retention of Electronic Records
   A. Electronic records have the same retention periods as their paper equivalents.
      1. The Citywide Retention schedule and the Department’s own retention schedule apply to all paper and electronic records.
   B. Official Record Copy
      1. Departments must decide whether the electronic version or the paper version will be considered the official record copy. Things to consider in this decision are:
         a. If retention is greater than 10 years, then the official record copy is the paper or microfilm.
         b. If retention is less than 10 years, then the official record copy can be the paper or the electronic file
      2. Departments must notify the Records Department of their decision on which version – paper or electronic – is their official record copy so that it can be reflected on the Retention Schedule.
      3. Decisions on whether paper or electronic records are considered the official record copy have implications for document destruction. See section V, part G.

III. What Is Needed for Control of Electronic Records?
   A. Controlling electronic records requires a system — sometimes called a document management system or a content management system — that consists of:
      1. Storage for the images or electronic documents
      2. A database that contains identifying information about the stored documents that aids in their retrieval and control.
      3. A network that enables users to put information into the system, query the database and retrieve the stored documents.
   B. See Illustration: How it works

IV. Imaging Considerations
   A. When does imaging make sense?
      1. There is a high volume of paper documents
      2. Documents are accessed frequently
      3. Many people need access to the documents
      4. Several people need access to the same documents at the same time — as part of a work process, for example
      5. Document searches are directed, that is, people know what they are looking for and know what it will look like when found
      6. Documents must be shared across departmental boundaries
      7. There is time pressure to find desired documents
8. There are penalties for not finding the documents promptly
9. There is a need to share documents with the public

B. When is imaging cost effective?
1. When it will make a process faster. For example, if workers have access to images they
don’t need to wait for paper files to reach them. This means that work tasks can be
done concurrently (at the same time) instead of sequentially (one at a time, depending
on who has the paper file). In this case, images can eliminate time spent waiting for
paper files to pass from one person to the next.
2. When it eliminates the need for duplicate files. One shared collection of images allows
different departments to access the same documents whenever they need to – even
after hours in the case of emergencies. A shared imaging collection also eliminates the
need to keep duplicate files in each department, saving time and effort.
3. When it gives the public easy access to documents. In some cases, allowing the public
to access document images can eliminate photocopying and postage or delivery
charges.

C. When does imaging not make sense?
1. Purely to save space. The cost to scan and index paper documents may be more than
the savings realized from freeing up floor space. This is especially true for documents
that are not retrieved often after they are created. Most records more than 4 years old
are retrieved fewer than 4 times per month. If space is a problem, consult the
retention schedule and see if older files are ready for destruction or if they can be sent
to the Records Center.
2. When documents have a short retention period, for example, less than two years. The
cost to scan and index doesn’t make sense if the documents will be disposed of after a
relatively short time.
3. When documents have a very long retention period, for example, 10 years or more.
The reason is that electronic files, including images, rely on application software and
operating systems to enable their use. Over a long span of time, it is more likely that
the technology needed to search for and view the images will become obsolete, making
it impossible to find and see them in the future.
   a. One possibility for this situation is to scan and microfilm the documents at the
same time. This provides images for today’s work and microfilm preservation of
the documents for retrieval in the future.

V. Scanning Decisions
A. Scanning is the process of converting paper documents into digital images that can be
displayed on a computer screen. Scanning is also called conversion (because you are
converting paper into images). Scanned images must also be indexed in order to become
searchable.

B. What strategy is best for scanning? How much should be scanned? There are four
possibilities:
1. **Scan documents from a particular date forward.** This is called day-forward conversion. It means that new work documents coming into the department will be scanned as they are received.
   a. The advantage is that the cost to scan and index will be spent on the most recent work.
   b. The disadvantage of this strategy is that it will take some time to build up a collection of useful images.
2. **Decide on a particular time period of older files** that will be converted to images. The time period depends on how often files are retrieved over time. For example, invoices are generally most active for about six months after they are sent, so scanning a year’s worth of invoices will be enough.
   a. The advantage is that the files most likely to be needed will be images.
   b. The disadvantage is that older files will have to be found manually.
3. **Scan the entire back file up to the present.** This generally is **not** recommended because:
   a. Imaging old files that are accessed infrequently is usually not a good investment.
   b. On the other hand, in situations where any file has the same chance of being retrieved regardless of its age – such as pension files or medical files – the decision may be to image the entire back file or to go day forward.
4. **Scan on demand.** In this situation, files are scanned as they are retrieved.
   a. The advantage is that the most active files will be available as images.
   b. The disadvantage is that it will be difficult to tell what is available as an image and what is not.

C. **Estimating guidelines: How many images do you have?**
1. Estimate the number of images by figuring that:
   a. One inch of files contains 200 pages
   b. One 25 inch filing drawer holds 5,000 pages
   c. One four-drawer upright filing cabinet holds 20,000 pages
2. Estimate the percentage of pages that are double-sided and add this to the estimate. For example, if measuring the files shows that they contain 100,000 pages, and 30% of pages are double sided, then the number of images needed is 130,000.
3. Estimate the amount of indexing required by figuring the number of pages per document. For example, if there are 5 pages to the document, and 100,000 pages in the collection, then there will be 20,000 documents to be indexed.

D. **How will scanning be accomplished?**
1. **Do it yourself on-site**
   a. This alternative works best for day-forward or on-demand scanning strategies. Issues to consider for this alternative include:
      1. Additional staff to prepare documents for scanning, operate scanning equipment, index documents, and perform quality control
      2. The number of scanners needed to keep pace with daily scanning volumes
3. PCs for performing document indexing and quality control operations
4. Space for preparing document batches and for holding documents after scanning
   b. For large scale backfile conversions – more than 1,000,000 pages – consider using an outside service bureau to do the work.

2. Use an outside service bureau
   a. This alternative works best for large scale backfile conversions where millions of pages must be scanned and indexed.
   b. Typical service bureaus charges:
      1. For project set-up: $150 per hour
      2. For document preparation - $20 to $40 per hour or .03 cents per page.
      3. Per scanned image, including quality control – 8 cents to 10 cents per page
      4. Indexing – 1 cent per keystroke. Also estimate 3 cents per folder to insert barcode sheets to facilitate automatic indexing.
      5. For pick-up and delivery of documents - $50 per truck trip
      6. For disks or tapes on which images and indexing are stored - $15 each, but amount varies according to medium chosen.
   c. Contract considerations:
      1. How does vendor handle documents picked up and returned so that all are accounted for?
      2. What remedies are there for lost boxes or batches of documents?
      3. What service levels are expected?
      4. What turnaround time is required?
      5. How are problems with image quality or index quality remedied?
      6. The City retains ownership of the documents before and after scanning
         a. Should paper files be re-assembled after scanning? For example, should case files be put back together into their original folders after scanning?
         b. The City retains the right to inspect how data and images are captured, indexed, stored, and protected against loss.
   d. Local service bureaus:
      Meta Source Imaging
      Bristol, PA
      Phone: 215-788-8885
      Fax: 215-788-8835
      www.metasource.com

      Conversion Technologies Intl.
      7000 Atrium Way, Suite 2,
      Mt. Laurel, NJ 08054
      Phone: 856-722-5588
E. **Should paper documents be kept after scanning?**
   1. **Documents that require long-term retention** (more than 10 years) should be kept in paper or microfilm format even if they have been scanned.
      a. Long retention periods increase the probability of technology obsolescence, so images scanned today may not be readable 10 years from now because computer operating systems and application software will change.
   2. **For images that are actively used right after scanning**, figure out how long it would likely take to find an error and hold the original documents that long after scanning.
      a. For example, if images are active for six months after scanning, hold the originals for at least six months. If a bad image is found during this time, the paper would be available for rescanning.
   3. **For images that receive 100% quality control**, it is still advisable to hold the paper originals for a minimal time period because it is still possible for scan errors to occur and go undetected.
      a. Asking for this level of quality from a service bureau is very expensive.

F. **How should paper documents be kept after scanning?**
   1. It will be important to organize and inventory paper records so that an original document can be easily located and retrieved if an image is later found to be of poor quality. At a minimum:
      a. Document batches should be kept together by batch number (which is generated by the imaging system) and date scanned.
      b. The document batches should be placed in boxes in order by batch number and by date.
      c. Each box should be entered into Versatile with the record series code for the scanned records, a description of the records scanned, the range of batch numbers in the box and the scanning date(s).
         i. Versatile will generate a box number so that the box can be tracked and called back if the original document needs to be retrieved.
      d. If possible, the imaging system should cross-reference the box number in which the original document is stored.

G. **What is needed for document destruction after scanning?**
   1. If the Department has decided that the electronic record is the official record, then no approval is needed to destroy the scanned paper.
2. If the Department has decided that the paper is the official record, then no approval is needed to destroy electronic files.
3. Destruction of the official record, whether paper or electronic, will still require the approval of the Records Department.

VI. Indexing

A. **Indexing** is the process of capturing identifying information about documents to facilitate search and retrieval.
   1. Identifying information about documents is sometimes called meta data or attributes.

B. **Capturing the right indexing attributes must take the needs of all searchers into consideration.**
   1. **Primary users** are workers directly involved in the business process. These searchers may need only minimal indexing to find the documents they need. For example:
      a. Document type: Invoice
      b. Document number: 123456
      c. Date: July 31, 2007
   2. **Secondary users** are those outside the department who need to see documents for other purposes. For example:
      a. Law department staff involved in disputes or legal claims may need to use documents as evidence. They are more likely to need additional search fields such as customer name, address, or other fields.
      b. The public may want to access records pertaining to their own accounts based on name, address or other criteria.
      c. The City Controller’s Office commonly needs to see documents as part of audit activities.
      d. Organizations that supply funding for City programs through grants may also have the right to inspect records
   3. **Records retention users** – All departments need to assure that records are kept as long as needed, then disposed of in the due course of business. For this, they must be able to search by record series code and review date.

C. **Minimum standard indexing required for electronic records:**
   1. Unique document identifier (assigned by the system)
   2. Document type or document title
   3. Date created
   4. Customer, author, sender or recipient
   5. Record Series Code
   6. Disposition/ review date (calculated by the system)

D. **Indexing best practice**
   1. All images should be indexed, at least at the folder level.
      a. It should not be necessary to search an entire CD or tape item by item to find the desired image or document.
2. All images should include indexing that enables records retention and disposition to be done

VII. Planning for an Imaging System
A. Technology Considerations
1. In addition to scanning and indexing, imaging requires, at minimum, a database, a storage repository and a network.
2. Departments should confer with their IT representative regarding their requirements to assure that any system purchased or developed conforms to City IT policy.
   a. Note that it may be possible to add images to an existing database.
      1. The City uses ApplicationExtender in order to add images to existing databases.
   b. In other circumstances, you may need to consider a commercial imaging product.
      2. Many such products exist, including LaserFiche, IBM FileNet, Oracle RM, and others.

B. Storage Media
1. Images may be stored on centralized network storage servers for quick access by users or on CDs.
   a. For images stored on CDs, there should be a directory that tells which images are on which CDs. This eases retrieval, especially for older images.

C. System Administration
1. Access permissions
   a. The department must consider the confidentiality of the images and plan access privileges accordingly.
2. Functionality privileges
   a. The department must consider what rights users will have, for example:
      i. Systems administrator – has all functionality privileges, including the ability to add new users to the system, change passwords, change indexing information, and so on.
      ii. Department user – has ability to search for and retrieve images. May also have ability to change or edit indexing information for images.
      iii. Secondary user and public – has ability to search for and view images only.
3. Backup of images and indexing data
   a. The department will need to schedule with IT to assure that the storage repository containing the images is backed up on a regular basis so that images can be recovered in the event of system failure.
      i. Best practice depends on how frequently new images are added.
b. The department will need to schedule with IT to assure that the database containing indexing information is backed up regularly so that it can be recovered in the event of system failure.
   i. Best practice depends on how frequently indexing information is added or changed.