

New York State Archives  
**Imaging Production Guidelines**  
2006

Do not use these guidelines, alone, as a specification. Instead, complete a File Information Form for each records series to be scanned and add the following essential items to the guidelines:

- Name of organization
- Name of records series
- Total number of images
- Preparation requirements
- Condition of records series (sizes, type, whether single- or double-sided)
- Color and age of paper
- Type of fasteners
- Percentage of documents flat, rolled, or folded
- Type and color of imprint
- Preparation requirements
- Required retrieval aids (indexing, filenaming, etc.)
- Required storage media numbering
- Number of duplicates required
- Storage media types required (CD-R, DVD-R, computer tape, etc.)
- Delivery information (date, destination, etc.)
- Technical enhancements to the guidelines needed to accommodate the needs of the records series

Packaged together, the appropriate imaging guidelines, the File Information Forms, and the supplemental requirements outlined above form an imaging project specification. A specification describes the essential technical requirements for items, materials, services, and procedures. Incorporate the specification into the standard contractual documents used by your organization, and ensure that your legal counsel has reviewed these documents.

**SCOPE:** These guidelines list the minimal standards for producing and inspecting digital images of records. The term “large-scale architectural, engineering, and topographical drawings” refers to hard-copy documents in excess of 11" X 14" in size and consisting predominantly of lines, or to similarly formatted electronic equivalents. The term “office documents” refers to standard 8½ X 11", 8½ X 14", or similarly sized paper documents consisting predominantly of textual data, or to similarly formatted electronic equivalents. The term “small-format pictorial documents” refers to documents of a size similar to office documents but which consist primarily of photographs, drawings, or other visual information. Where applicable, these guidelines follow national digital imaging standards and industry practices. All references to industry standards (ANSI, AIIM, etc.) are to the latest revision thereof.

**1. IMAGE FORMAT:** Requirements for image formats shall depend on the particular purpose of the image.

**1.1 Master Images:** Record copies of images shall meet these guidelines:

**1.1.1 For black and white textual or line-art documents:**

Format: 1-bit TIFF (latest version)  
Tonal depth: Bitonal  
Compression: Uncompressed  
Spatial resolution: At least 200 dpi (unenhanced true scan)

**1.1.2 For black and white photographs or other pictorial documents:**

Format: 8-bit TIFF (latest version)  
Tonal depth: Grayscale  
Compression: Uncompressed  
Spatial resolution: At least 200 dpi (unenhanced true scan)

**1.1.3 For documents in which color is essential:**

Format: 16, 24, or 36-bit TIFF (latest version)  
Tonal depth: Color  
Compression: Uncompressed  
Spatial resolution: At least 200 dpi (unenhanced true scan)

**1.1.4 For backup images of the above, as applicable:**

Format: As applicable above  
Tonal depth: As applicable above  
Compression: Latest ITU standard compression preferred  
Spatial resolution: As applicable above

**1.2 Access Images:** Use copies of images, if they must differ technically from the master copies, shall meet one of the following guidelines:

**1.2.1 General Requirement (Alternative 1):**

Format: TIFF (latest version)  
Tonal depth: Bitonal, grayscale, or color, as appropriate  
Compression: Allowed (latest ITU standard compression preferred)  
Spatial resolution: At least 200 dpi

**1.2.2 General Requirement (Alternative 2):**

Format: JPEG or JPEG2000 (latest version)  
Tonal depth: Grayscale or color, as appropriate  
Compression: Allowed as needed, JPEG2000 lossless compression preferred  
Spatial resolution: At least 200 dpi

### 1.2.3 General Requirement (Alternative 3):

Format:	PDF/A with images using lossless compression preferred
Tonal depth:	Bitonal, grayscale, or color, as appropriate
Compression:	Allowed as needed
Spatial resolution:	At least 200 dpi

2. **COMPRESSION:** Maintain record copies uncompressed to ensure easy accessibility to the image over time. Compress access or use copies using non-proprietary, lossless compression algorithms. Document that lossless compression is being used to ensure there is not a mixture of lossy and lossless images in the files.
3. **SCALING:** Scale access images so most documents fit within the typical computer screen or window for the given application. For instance, a particular application may require documents be scaled to half their size or less to comfortably fit a screen.
4. **IMAGE HEADERS:** Images cannot have proprietary headers.
5. **IMAGE ORIENTATION:** Upright (maintain portrait or landscape orientation as appropriate).
6. **STORAGE MEDIA:** Regardless of the media, master copies of images must be accessible for the records' entire retention period. The media shall depend on the purpose of the images stored.
  - 6.1 **Master or Backup Images:** Store these images on computer tape, CD-Rs, DVD-Rs, or other digital storage formats, as appropriate.
    - 6.1.1 The manufacture date of removable storage media shall be less than a year before first use.
    - 6.1.2 Removable storage media shall have a pre-write shelf life of at least five years and a minimum post-write life of twenty years, based on accelerated aging test results that report on specific disc areas, such as those found in ANSI IT 9.21 CD-ROM.
    - 6.1.3 Computer tape is strongly preferred when using digital images as the storage format for permanent records.
    - 6.1.4 Discs used as storage media must comply with the applicable ISO standards, which specify how this type of media disc must store information and which allow the interchange of discs within different systems, such as ISO/IEC 13490-1:1995 ("Information technology—Volume and file structure of read-only and write-once compact disk media for information interchange") and ISO/IEC

20563:2001 (“Information technology—80 mm and 120 mm DVD-recordable disk”).

**6.1.5** The media should be examined at least twice a year to check for signs of degradation.

**6.2 Access Images:** Store these images on magnetic non-removable equipment (server or hard drive), CD-Rs, DVD-Rs, or other digital storage formats, as appropriate. CD-RWs or DVD-RWs are not allowable.

**7. MULTIPLE-PAGE IDENTIFICATION:** All images that comprise a single document shall be accessible and presentable in their original order and be clearly associated with each other as parts of a single document.

**8. VENDOR QUALITY CONTROL AFTER SCANNING:**

**8.1 Inspection** of the images by the vendor for quality shall verify the following:

- a. Correct image filename (unique identifier)
- b. Correct file format for each image type (master and access)
- c. Image scanned at appropriate unenhanced dpi for each image type
- d. Image oriented properly, whether landscape or portrait
- e. Image is correct size (in pixels along both dimensions)
- f. Image is not skewed
- g. Image is not rotated or flipped
- h. Image is neither too light nor too dark
- i. Appropriate contrast exists within the image
- j. No distortion of the image
- k. No extraneous materials (fingers, or fasteners) obscure the image
- l. No noise or other problems in image file
- m. Appropriate indexing terms are associated with the scanned image
- n. Monitor images are verified under controlled viewing conditions
- o. Image viewer used to view and evaluate the images is indicated
- p. Dpi verified by an independent software program

**8.2 Correction of unacceptable images** shall consist of the following:

- a. Correcting image filename
- b. Deskewing, rotating, or flipping the image to correct its orientation
- c. Adjusting brightness, contrast, or tone through rescanning
- d. Cropping that does not remove any information in the document
- e. Rescanning, followed by a re-inspection of the new image
- f. Updating index database to correct errors

**8.3 Unacceptable modifications to the images** include the following:

- a. Sharpening the image
- b. Retouching or despeckling
- c. Dithering or quantization
- d. Removing information from the images
- e. Adding information to the images
- f. Burning annotations or “sticky notes” onto the image file itself

**9. RESOLUTION:** Image sharpness shall be equivalent to the dots per inch (dpi) required for the original image type as explained above (“1. Image Format”). Use commercially produced resolution targets, such as those outlined in ANSI/AIIM TR38-1996, “Identification of Test Images for Document Imaging Applications,” and follow techniques in ANSI/AIIM MS44-1988 (“Recommended Practice for Quality Control of Image Scanners”) to verify scanner performance. Provide evidence of adherence to these standards at the close of the project.

**10. DOCUMENTATION TO SCAN WITH THE RECORDS:** Scan the following documentation with the records, ensuring that this documentation is clearly associated with this specific set of records:

- a. State Archives Records Inventory Data Worksheet
- b. File Information Form
- c. Organization name and address
- d. Contractor’s name and address, and dates of scanning
- e. Resolution Target that complies with ANSI/AIIM TR38-1996 (“Identification of Test Images for Document Imaging Applications”) or other industry-standard resolution target
- f. Indices, finding aids, and other metadata associated with the documents (which differ from the index database for the images), if supplied by the customer

**11. RETRIEVAL AIDS:** Indexing shall comply with specific requirements of the customer, but shall at a minimum include the following:

**11.1. Unique Identifier for Images:** Each image shall have a unique identifier, preferably sequential, which can be numeric, alphanumeric, or alphabetic as required by the customer. Each filename shall be unique across all separate external media, not merely within a single disc or tape. If required, file the documents in appropriate electronic folders.

**11.2. Indexing Data Fields:** The index of images shall consist of a limited number of field names to ensure adequate access to the records. Whenever possible, the field data shall consist of objective indexing terms (such as personal names, file numbers, dates, etc.) rather than subjective data (such as subject terms).

- 11.3. Optical or Intelligent Character Recognition:** If required, the vendor shall conduct optical or intelligent character recognition (OCR or ICR) to convert digital images into electronic text. The vendor shall certify the conversion to be at least 99.9% accurate as measured by character count, which is equivalent to 10 to 20 errors per page on a standard 2000-character page when formatting is taken into account. The converted text shall be logically associated with the respective digital image or document. This percentage
- 11.4. Indexing Database:** The indexing database (including, if applicable, OCR'd text) shall store the required index data in ASCII or Unicode, and shall associate each record within the database with the respective digital image or document.
- 11.5. Index Accuracy:** The vendor shall verify the index via dual data entry, data entry operator verification immediately subsequent to data entry, or other means, as appropriate or required, to ensure accuracy.
- 12. PACKAGING:** Optical media (CD-Rs, DVD-Rs, etc.), if used, shall be stored in unbroken jewel cases and shall rest on the inner spindle without pressure that could produce damage during removal or re-emplacement. Tapes and other media, if used, shall be in individual containers of the appropriate size for the particular media. The customer may accept or require alternate packaging as suitable. The vendor shall deliver separated sets of master and duplicate copies of media to the customer in boxes, with the media fitting firmly but not tightly.
- 13. PACKAGE MARKING:** At a minimum, the following data should be machine printed on the label on each jewel case or other storage container of both the originals and the backups:
- Organization Name
  - Records Series Title and Date
  - Range of Records (if appropriate)
  - Package or Media Number
- 14. QUALITY OF WORK:** Scanning shall capture each digital image of a document page so that every line and character on the document appears and is legible in the image. Removable media shall be free of scratches, cracks, finger marks, warping, or any other defect that might adversely affect quality or usability.
- 15. CONTRACTOR INSPECTION:** The vendor shall inspect each individual image, disc, tape, or other storage medium for compliance with the requirements herein, including resolution, image quality, accuracy of the index, and general workmanship. The vendor shall include an inspection report or certification covering each disc, tape, or other storage medium included in each shipment.

- 16. CUSTOMER QUALITY CONTROL:** The customer shall have the right, after inspection, to reject any images determined not to meet the requirements of these guidelines. In such cases, the contractor must rescan at its expense.
- 17. REJECTION OF BACKUP MEDIA:** When the customer or its inspection agent (if applicable) rejects an entire disc, tape, or other storage medium, the customer or vendor may deface the rejects by cracking, punching, or shredding. The customer may retain rejected media at its discretion.
- 18. VENDOR FACILITY INSPECTION:** The customer reserves the right to inspect and approve the vendor's work site before and at any time during the performance of a contract to ensure the vendor's production and quality control capabilities. An inspection of the vendor's facilities prior to production is always preferred.
- 19. FILE INTEGRITY:** Unless otherwise specified elsewhere in the contract, the vendor shall maintain the original documents in their existing file order before, during, and after scanning. The vendor shall return file material to the original storage containers in the same order that existed before scanning, except that the vendor shall maintain any corrections to file order made during the preparation for scanning. The vendor shall not restore any fasteners (staples, clips, or tape) removed during document preparation.

# Appendix A

## Explanations and Notes on Imaging Production Guidelines

Below is a series of questions and answers that clarify these guidelines.

### 1. Image Format

#### **Why did the Archives choose the TIFF format over other file formats for images as a master copy?**

TIFF is an open standard, meaning its source code is freely available to the public. Since the source code is available, there is a greater likelihood of current or future access to the image because users will always have the opportunity to write new code for software to view the image. In addition, TIFF is an international standard that is used for master copies of digital images by many cultural institutions, including the Library of Congress, for their collections of digital images. Finally, TIFF has been an international standard since 1992, so archivists consider this a stable format for long-term retention.

#### **Why are there separate requirements for master and access images?**

Master images must be high-quality images to ensure the best legibility of the file and to allow for the printing of sharp images. Access images generally need to be smaller to allow for quicker access, such as over the Internet, so they do not need to adhere to the same quality requirements as master images.

#### **Why are there different file formats for the access copies?**

##### *JPEG (Joint Photographic Experts Group)*

Since this file format (which is officially called JFIF) can use both lossy and lossless compression, it is essential that the vendor document which compression process is being used. JPEG2000 provides lossless compression of images. The Library of Congress and other cultural institutions now tend to prefer JPEG2000 for service images.

##### *PDF (portable document format)*

Portable Document Format (PDF) is a de facto standard for secure and reliable distribution of electronic documents and forms. PDF is a universal file format that preserves the text, fonts, images, graphics, and layout of any source document, regardless of the application and platform used to create it. An open file format specification, PDF is available to anyone who wants to develop tools to create, view, or manipulate PDF documents. The latest format is a current international standard for

archival records, the PDF/A, ISO 19005-1. When creating PDF files, remember to avoid images that might use lossy compression embedded into the image itself.

## **2. Compression**

### **Is compression allowed?**

Compression algorithms can help reduce the size of any image file greatly. For example, a typical uncompressed image averages about one megabyte per page, while a Group IV compressed image with 200 DPI resolution will be about 50 to 70 thousand kilobytes. However, since lossy compression causes some information loss, it is inappropriate for the master or record copy of a digital image, even if the data loss is not visible to the naked eye. For this reason, the Archives recommends the use of uncompressed images for the master images and the latest standard lossless compression for the access images.

## **3. Scaling**

### **Why is scaling important?**

It is important for the organization and vendor to consider the issue of scaling from the outset. When creating specifications for access images, you must take into account that most documents need to fit within the typical computer screen or within a particular window for the given computer application. For instance, an image with dimensions of 640 X 480 pixels might be appropriate for systems using low-resolution monitors, but a particular application may require documents be scaled to half their size or less to function well over the Internet.

## **6. Storage Media**

### **Why are rewritable forms of media not suitable for storage?**

Rewritable forms of media (such as CD-RWs and DVD-RWs) are not acceptable because people can easily alter the images on the media. In addition, due to their chemical composition, they are more susceptible to fluctuations in the environment than read-only discs. The information on read-only discs (such as CD-Rs and DVD-Rs) cannot be changed.

### **What is the Archives' recommendation on media longevity?**

The issue of the longevity of optical media, even when stored under optimal conditions, is still under debate. Most forms of media have very short life spans (about two to five years), so finding information about a CD or DVD that guarantees a post-write shelf life of twenty years does not necessarily mean that the disc will last that long. Because

media longevity can be unpredictable and the loss of information instantaneous and catastrophic, the best option is to use brand name media that have proved reliable in the past rather than discount CDs and DVDs.

### **Why is computer tape preferred by the Archives?**

When optical media and computer tapes are stored under similar conditions, the tapes tend to have a longer shelf life than the optical media. However, since no computer storage medium is permanent, the data stored on any medium must be refreshed periodically.

## **9. Resolution**

### **Why must scanner performance be checked if someone has set the scanner for the appropriate dots per inch?**

Checking the performance of the scanner is required because quite often problems arise in image files because of the performance of the equipment. Setting the dpi merely ensures that a certain number of dots per inch are captured; it does not ensure that the scanner has correctly captured the quality of the original document. You must keep in mind that scanners have a limited effective lifetime.

### **Can I use a microfilm resolution target for imaging or COM?**

You must use a standard imaging resolution target rather than a microfilm target even for Computer Output Microfilm or (COM). You cannot monitor film image quality by putting a micrographic control target through a scanner and then writing it to film, since the scanner will not be able to capture enough visual information.

## **11. Retrieval Aids**

### **What is OCR?**

OCR stands for optical character recognition, and it is the recognition of machine printed characters by a computer and the conversion of the characters in those images into electronic text.

### **What is ICR?**

ICR, or intelligent character recognition, is the recognition of printed and handwritten (but not cursive) characters by a computer that uses context to determine the most likely character and the subsequent conversion of images of those characters into electronic text.

### **Why does the index data have to be put into ASCII or Unicode?**

Unlike other proprietary word processing programs, such as Microsoft Word, ASCII and Unicode documents can be opened and viewed by a number of different computer programs regardless of what operating system is in use.

### **Are there other ways to save indexed data?**

You can also use XML as a data format for your index.

## **13. Package Marking**

### **Can I put a label directly on a CD or DVD?**

Do not put adhesive labels on either side of optical media (CDs and DVDs). The glue can corrode the polymer layer of a disc, destroying the data.

## **17. Rejection of Backup Media**

### **Why do rejects have to be destroyed?**

Vendors must destroy rejects or give them to the customer because information in these records may be sensitive or confidential.

## **18. Vendor Facility Inspection**

### **Why is a visit to the vendor's facility recommended?**

An inspection of the vendor's facilities is always recommended because it gives you an accurate idea of the work environment where your records will be stored and handled.

## **19. File Integrity**

### **Why should the vendor not restore the fasteners to the original records?**

Adding staples will increase the work the vendor does for you, and hence the cost. Also, since most fasteners are made of metal and tend to rust, it is better to remove fasteners entirely for the longevity of the records.

## Appendix B

### File Information Form

## Appendix C

### Glossary of Imaging Terms

**accelerated aging.** the process of speeding up the deterioration of a material to estimate the rate at which deterioration would take place under normal long-term conditions

**access image.** the version or copy of the master image that is used by individuals in lieu of the master copy

**accessibility.** the ability to retrieve and use an electronic image or document

**AIIM International.** previously known as the Association for Information and Image Management, an international group specializing in the capture, management, storage, preservation, and delivery of enterprise content to support business processes

**algorithm.** the mathematical formula used to conduct an operation upon electronic records (such as compression or encryption)

**American National Standards Institute (ANSI).** a non-governmental organization that develops and publishes standards for voluntary use in the United States and that represents the US in international standards activities

**ANSI.** see “American National Standards Institute (ANSI)”

**American Standard Code for Information Interchange (ASCII).** a very common computer code used to preserve and present the Latin alphabet, punctuation marks, some symbols, and simple text formatting (including spaces and carriage returns)

**annotation.** a note added to or associated with a digital image to provide additional information, such as an explanation of existing defects in the original image; also called “sticky note”

**ASCII.** see “American Standard Code for Information Interchange (ASCII)”

**back up.** the process of copying an electronic record to ensure its information will not be lost

**backup.** a copy of an electronic record maintained to protect the information from loss

**barcode.** a coding system composed of vertical lines set in patterns that, when read by an optical reader, can be converted into electronic text

**barcode recognition (BCR).** the recognition of barcodes by a computer and the conversion of those into electronic text to create an index for the documents

**BCR.** see “barcode recognition (BCR)”

**bitmap.** a representation of an image captured as a set of pixels arrayed in a rectangular series of horizontal lines

**bitonal.** purely black and white

**CD.** see “compact disc (CD)”

**CDIA.** see certified document imaging architect

**CD-R.** see "compact disc, recordable (CD-R)"

**CD-ROM.** see "compact disc, read-only memory (CD-ROM)"

**CD-RW.** see “compact disc, rewritable (CD-RW)”

**certification target.** a document captured in the microfilming or imaging process in which the machine operator and the records custodian document that the images were captured during the normal course of business

**certified document imaging architect.** a person who possesses the expertise to successfully plan, specify and design an imaging solution, and who demonstrates that knowledge by passing the CDIA examination

**character.** an alphabetic, numeric, or other symbol used to convey information in textual form

**COM.** see “computer-output microfilm (COM)”

**compact disc (CD).** a family of standard flat, circular digital storage media formats that are written to and read by a laser, that are 120 mm (4.72 inches) in diameter, 1.2 mm (0.05 inches) thick, and made of clear polycarbonate plastic, and that store approximately 650 megabytes of data

**compact disc, recordable (CD-R).** a version of the compact disc technology that a user can copy to only once

**compact disc, read-only memory (CD-ROM).** a version of the compact disc technology that is used exclusively as a multi-copy publication format

**compact disc, rewritable (CD-RW).** a version of the compact disc technology that allows information to be rewritten on the same disc numerous times and which is therefore less secured than a CD-R as a form of storage

**compression.** a computer process using algorithms that reduces the size of electronic documents so they occupy less digital storage space

**computer-output microfilm (COM).** the process of converting electronic data directly to microfilm or microfiche

**crop.** an image processing method of removing the region near the edge of the image, but keeping a central area

**data field.** a location in a database that stores one type of data (such as an address field, a city field, a state field, etc.)

**declaration by records custodian.** a document, usually included as a target, in which the custodian of the records verifies the records were scanned or microfilmed in the normal course of business

**deskew.** to use computer software to make upright a scanned but off-balance image

**despeckle.** to erase extraneous dots in a digital image produced during the scanning process

**digital.** in binary code (as is used by computers)

**digital versatile disc (DVD).** a family of flat, circular digital storage media formats that are written to and read by a laser, that are 120 mm (4.72 inches) in diameter, 1.2 mm (0.05 inches) thick, and made of clear polycarbonate plastic, and that store at least 4.7 gigabytes of data

**digital versatile disc, recordable (DVD-R).** a version of the digital versatile disc technology that a user can copy to only once

**digital versatile disc, rewritable (DVD-RW).** a version of the digital versatile disc technology that allows information to be rewritten on the same disc numerous times and which is therefore less secured than a DVD-R as a form of storage

**digital versatile disc, read-only memory (DVD-ROM).** a version of the digital versatile disc technology that is used exclusively as a multi-copy publication format

**digitize.** to convert an image or signal into binary data

**dither.** to add more colors or shades of gray to a digital image to reduce its jagged appearance; the approximate opposite of “quantize”

**dots per inch (dpi).** a measurement of resolution of a digital image or of the resolution a computer device can input or output (for example, the number of pixels a printer can print or a monitor can display)

**dpi.** see “dots per inch (dpi)”

**DVD.** see “digital versatile disc (DVD)”

**DVD-R.** see “digital versatile disc, recordable (DVD-R)”

**DVD-ROM.** see “digital versatile disc, read-only memory (DVD-ROM)”

**DVD-RW.** see “digital versatile disc, rewritable (DVD-RW)”

**file integrity.** the state of a set of records being an accurate representation of the original set of records and in the same order originally maintained

**finding aid.** a tool (such as a series description, catalog, or index) designed to help users find information within archival records

**GIF.** see “graphics interchange format (GIF)”

**graphics interchange format (GIF).** a lossless file format for displaying bitmapped images that uses a proprietary compression format

**grayscale.** the range of shades of black an image has; the color of a digital image captured or preserved in shades of black (rather than color or pure black and white)

**guidelines.** criteria that attempt to ensure the best quality product is produced by adhering to a specified process

**header.** identifying information contained within the properties of an image; see also “metadata”

**ICR.** see “intelligent character recognition (ICR)”

**image.** a graphic representation of an object (such as a document)

**image capture.** see “imaging”

**image header.** see “header”

**imaging.** the process of electronically capturing the visual appearance of (usually) paper documents; informally called "scanning"

**index.** an information guide that identifies the location of specific pieces of information within a document or a set of documents (for example, an index to a set of minutes could list topics and when they were discussed, or an index to personnel files could list the names of people included)

**intelligent character recognition (ICR).** the recognition of printed and handprinted (but not cursive) characters by a computer that uses context to determine the likely character and the subsequent conversion of images of those characters into electronic text; see also “optical character recognition (OCR)”

**International Telecommunication Union—Telecommunication sector (ITU-T).** a standards-making body and the standards it produces that are used to ensure accurate compression of digital information (such as ITU Group 3, 4, and 5)

**ISO (International Organization for Standardization).** a worldwide federation, founded in 1947, of national standards bodies from some 100 countries, one from each country

**ITU.** see “International Telecommunication Union—Telecommunication sector (ITU-T)”

**JBIG.** see “joint bi-level image group (JBIG)”

**joint bi-level image group (JBIG).** a lossless image compression format used frequently for black and white images

**joint photographic experts group (JPEG).** an image compression standard sanctioned by the International Standards Organization and International Telecommunication Union and used for photographic images rather than text images

**JPEG.** see “joint photographic experts group (JPEG)”

**JPEG 2000.** an image compression standard that supports both lossy and lossless compression and that was created by the Joint Photographic Experts Group committee with the intention of superseding their original JPEG standard

**landscape.** the orientation of an image that is wider than it is tall

**lossless compression.** a compression method that retains every bit of data that was in the original digital file (such as ITU Group 4)

**lossy compression.** a compression method that reduces a digital file by permanently eliminating certain information (such as JPEG)

**master image.** the version of the imaged original that serves as the official copy (see also “record copy”)

**metadata.** “data about data”; information contained within an electronic document that identifies and describes it to the operating system or user

**migration.** the periodic transfer of data from one electronic system to another to retain the integrity of the data and to allow users to continue to use the data in the face of changing technology; sometimes called "data migration"

**non-proprietary.** open format; not owned and controlled by a single company and, therefore, usable in various computing environments

**OMR.** see “Optical Mark Sense Recognition (OMR)”

**open format.** a format that is in the public domain and, therefore, readable in a number of computing environments

**optical mark sense recognition (OMR).** the recognition of marks on a form (such as filled-in bubbles) by a computer and conversion of that information into electronic text; also called “optical mark sense recognition”

**optical character recognition (OCR).** the recognition of printed characters by a computer and conversion of images of those characters into electronic text

**optical disc.** a storage medium for electronic records that is written to and read by laser, such as CDs and DVDs

**original.** the final version of a document, as opposed to copies or duplicates made of it

**PDF.** see “portable document format (PDF)”

**pel.** picture element; see also “pixel”

**pixel.** picture element; a tiny dot in a digital image composed of a set of such dots

**PNG.** see “portable network graphic (PNG)”

**portable document format (PDF).** a file format that has captured all the elements of a printed document as a multi-page electronic image, along with the possibility of copying electronic text from the document

**portrait.** the orientation of an image that is taller than it is wide

**proprietary.** owned and controlled by a single company and, therefore, usually only readable in a certain software and hardware environment and not necessarily exportable to another environment

**proprietary header.** information stored in the header of a digital image that may render an otherwise open-format image unreadable except in a certain computing environment

**quantizate.** to reduce the number of colors or shades of gray in a digital image to reduce file size; the opposite of “dither”

**record copy.** an original record or a copy of an original record that is used to meet the minimum retention period for that record; also called "official copy"

**records series.** a group of related records (such as minutes of a board, payrolls, and purchase orders) that are normally used and filed as a unit and that normally have the same retention requirements

**removable media.** computer storage devices that can be separated from a computer device, such as CDs, computer tape, or thumb and jump drives

**render.** the process of displaying an image

**resolution.** the measure of the quality of a digital image, usually expressed in dots per inch (DPI); see also “spatial resolution”

**resolution target.** a technical image that measures the resolution captured by a scanner or a microfilm camera; see also “resolution”

**ROM disc.** “a read-only memory disc”; a flat, circular computer storage medium that is manufactured in large quantities and which cannot be rewritten to by the user

**scale.** to contract or enlarge a bitmapped image to achieve a required output size

**scanner.** a machine that captures eye-readable images as digital image files

**scanning.** see “imaging”

**spatial resolution.** a measure of the detail captured in a digital image (represented by dots per inch)

**specification.** a written document that details the technical requirements for a particular imaging or micrographics project

**sticky note.** see “annotation”

**tagged image file format (TIFF).** a standard, lossless file format for capturing color and grayscale images; sometimes called “TIF”

**target.** information imaged or microfilmed along with a set of documents to supply bibliographic or technical data

**TIFF.** see “Tagged Image File Format (TIFF)”

**tonal depth.** a measure of the color quality captured in a digital image (such as black and white, grayscale, 8-bit color, and true 24-bit color)

**true scan.** a legal term meaning a digital image is an exact copy of the original; there are no enhancements of any kind so it is admissible as evidence in court.

**uncompressed.** (said of a computer file) maintained in its original size with all of its original digital data; not reduced in size through an algorithmic process for the purposes of storage and space savings

**Unicode.** a 16-bit standard for representing characters in digital code and designed to include all known characters in all modern and ancient writing systems; see also “ASCII”

**use copy.** a version or copy of the master image that is used by individuals in lieu of the master copy

**zoom.** to make an image appear larger (zoom in) or smaller (zoom out) by redisplaying the image at different resolutions

# File Information for Microfilming Paper Records

**SCOPE:** The information on this form, combined with the applicable *Microfilm Production Guidelines*, constitutes a technical specification for producing high quality microfilm. Governments/agencies contracting with a microfilm service bureau should attach whatever additional contracting documents your local government requires. A typical bid offering will, therefore, consist of at least four (4) items:

1. this File Information Form
2. the Microfilm Production Guidelines
3. a SARA Microfilm Agreement
4. locally-required contracting materials

Instructions for completing this form are on a separate sheet, titled *Instructions for Completing File Information Form*.

1. NAME OF LOCAL GOVERNMENT \_\_\_\_\_

2. PREPARED BY (*Print name*) \_\_\_\_\_

TELEPHONE NUMBER (*Include Area Code*) \_\_\_\_\_

3. NAME OF RECORD SERIES \_\_\_\_\_

4. \_\_\_\_\_ Years    ◀ The retention schedule for these records is:

5.  Active     Inactive    ◀ Is this record series ACTIVE or INACTIVE?

6. Order of Filming: The records in this series will be microfilmed in this existing order:

A. **Alphabetically**     by personal name (*last name first*)

by corporate name (*business name*)

by another alphabetical order, (*describe*) \_\_\_\_\_

B. **Numerically**     by parcel number

by case number

by personal identification number

by another numerical order, (*describe*) \_\_\_\_\_

C. **Chronologically**     by year/month/day (yyyy/mm/dd)

by month/day/year (mm/dd/yy)

by another chronological order, (*describe*) \_\_\_\_\_

D.  by a COMBINATION of the above, or other factors. Describe the hierarchy of the records:

---

---

---

E.  in RANDOM order, requiring a separate index for document retrieval; this index is usually computer-based ("Computer-Assisted Retrieval, CAR"). Describe the desired indexing system:

---

---

---

7. \_\_\_\_\_ ◀ Total number of DOCUMENTS  
\_\_\_\_\_ % ◀ PERCENTAGE of documents that have information on both sides of the sheet
8. \_\_\_\_\_ ◀ Total number of IMAGES
- 8a. \_\_\_\_\_ ◀ Total number of ROLLS
- 16mm microfilm -- office documents (divide total images by 2500)
  - 35mm microfilm -- bound documents (divide total images by 1000)
  - 35mm microfilm -- drawings (divide total images by 500)
9. \_\_\_\_\_ ◀ Number of images by which this record series INCREASES by each year

10. Condition of Documents:

A. Size

\_\_\_\_\_ Length      \_\_\_\_\_ Width      ◀ LARGEST Document

\_\_\_\_\_ Length      \_\_\_\_\_ Width      ◀ SMALLEST Document

B. Age/Brittleness/Fasteners

\_\_\_\_\_ Years      ◀ AGE of paper (*oldest*)

Yes     No      ◀ Edges of paper are FLAKING

\_\_\_\_\_ %      ◀ Percentage of the documents which have FASTENERS  
(*staples, paper clips, etc.*)

C. Enclosure

Yes     No      ◀ Are documents in folders?

Yes     No      ◀ Are documents bound? *If Yes,*

Yes     No      ◀ Can the books be disbound

Type of binding:

sewn (only)     sewn and glued     drill post     comb

glued (only)     ring     clamp

other \_\_\_\_\_

D. Lay of the paper

- \_\_\_\_\_ % ◀ Percentage of the documents which are FLAT
- \_\_\_\_\_ % ◀ Percentage of the documents which are ROLLED
- \_\_\_\_\_ % ◀ Percentage of the documents which are FOLDED

E. Color of the paper

- \_\_\_\_\_ % ◀ Percentage of documents on WHITE PAPER
- \_\_\_\_\_ % ◀ Percentage of the documents on \_\_\_\_\_ paper  
(indicate color)
- \_\_\_\_\_ % ◀ Percentage of the documents on \_\_\_\_\_ paper  
(indicate color)
- \_\_\_\_\_ % ◀ Percentage of the documents on \_\_\_\_\_ paper  
(indicate color)
- \_\_\_\_\_ % ◀ Percentage of the documents on \_\_\_\_\_ paper  
(indicate color)
- \_\_\_\_\_ % ◀ Percentage of the documents on \_\_\_\_\_ paper  
(indicate color)
- \_\_\_\_\_ % ◀ Percentage of documents that are BLUEPRINT or PHOTOSTAT

F. Imprint (More than one entry may be checked)

- handwritten only       printed/typed only       handwritten and printed
- ink       pencil       carbon paper

G. Color of imprint \_\_\_\_\_

11. Document Preparation for Microfilming

- A.  the Government/Agency       the microfilming vendor      ◀ FILE ORGANIZATION shall be performed by (check one)

---

- B.  the Government/Agency       the microfilming vendor      ◀ PURGING shall be performed by (check one)

---

- C.  the Government/Agency       the microfilming vendor      ◀ PREPPING the documents shall be performed by (check one)

12. Records Usage

- A. \_\_\_\_\_ ◀ Estimated ANNUAL number of people who use this record series
- B. \_\_\_\_\_ ◀ Monthly number of "look-ups" (reference or retrieval)
- C. From the date of document creation, most retrievals occur within
  - one week       one month       one year
  - 2 - 6 years       continued steady usage
- D. The persons who **physically handle** the documents are:
  - employees only
    - this department only       several departments
  - general public
  - other (describe) \_\_\_\_\_

E. The persons who **make use** of the documents are:

- employees only  
 this department only     several departments  
 general public  
 other (*describe*) \_\_\_\_\_

F. \_\_\_\_\_ minutes    ◀ How much time is normally necessary to retrieve a document?

G. \_\_\_\_\_ minutes    ◀ no more than

H. Does retrieving documents entail any special problems? (*If Yes, describe*) \_\_\_\_\_  
\_\_\_\_\_

I.  Yes     No    ◀ Does one document have to be compared to another?

*If Yes, are the two documents*

- from the SAME record series?  
 from DIFFERENT record series?

J.  Yes     No    ◀ Are the documents PHOTOCOPIED?

K. \_\_\_\_\_    ◀ Approximately how many photocopies are produced on a monthly basis?

### 13. Microfilm Format

A. The first-generation silver-gelatin (camera) film shall be (*check one*):

- 16mm roll microfilm  
 35mm roll microfilm

**Warning: First generation silver-gelatin microfilm (camera film) shall not be jacketed**

B.  Yes     No    ◀ Is a second-generation silver gelation negative required?

C.  Yes     No    ◀ Is a second-generation diazo (duplicate) required?

*If Yes, the duplicate shall be:*

- 16mm or 35mm roll microfilm

\_\_\_\_\_    ◀ *Number of duplicates required*

- loaded into 105mm x 148mm microfiche jackets

\_\_\_\_\_    ◀ *Number of duplicates required*

- loaded into 105mm x 148mm microfiche jackets AND duplicated to unitized microfiche

\_\_\_\_\_    ◀ *Number of duplicates required*

14. **Delivery Date:** All paper records and completed microfilm shall be returned to the contracting local government not later than \_\_\_\_\_

#### Additional Information:

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

**DATE THIS FORM COMPLETED:**

\_\_\_\_\_