

DELAWARE PUBLIC ARCHIVES
POLICY STATEMENT AND GUIDELINES

**GUIDELINES FOR UTILIZING PAPER AND COMPUTER OUTPUT IMAGING
CONVERSION SERVICES**

STATEMENT OF PURPOSE

The Delaware Public Archives (DPA) is issuing these " Guidelines for Utilizing Paper and Computer Output Imaging Conversion Services" for use by all agencies in state and local government in Delaware. The purpose of these guidelines is to provide agencies with a set of uniform best practices for the definition of requirements and the selection of imaging conversion services for converting paper records to digital image format and computer output to digital format or microfiche. The purpose of these guidelines is to help ensure that the resulting digital images or microfiche will be accurate, complete, accessible, reliable, trustworthy and, therefore, legally acceptable.

BACKGROUND

An increasing number of agencies in the State of Delaware have a need to convert existing paper to digital images and computer output reports and documents to digital or microfiche formats in order to achieve more efficient storage, retrieval and management of business records and to facilitate improvements in business process productivity and citizen and supplier service quality and responsiveness.

With the goal of supporting these needs, the DPA is issuing these guidelines to assist agencies in evaluating needs, defining their requirements, and communicating the requirements to a DPA qualified vendor to provide image conversion services and computer report conversion services that meet their requirements.

The Guidelines consist of the following sections:

- I. Definitions
- II. Paper to Image Conversion
- III. Computer Report to Digital Conversion
- IV. Computer Report to Microfiche Conversion

I. DEFINITIONS

COLD/ERM – Computer Output to Laser Disk/Enterprise (or Electronic) Report Management: The conversion of computer-generated reports in “as-if-printed” textual/graphic format (including automatic extraction of index data) for storage on and retrieval from digital media. The computer-generated reports are typically formatted ASCII text, but some may include graphics elements or forms overlays.

COM (Computer Output Microfilm). COM recorders write the print stream of electronic records, which is human readable, onto silver halide film by using a laser to expose the film to page images that are developed by wet chemistry. Most COM records produce 4” X 6” sheets of flat film otherwise known as microfiche.

Image: an image is a page-side of a hardcopy record that has been processed through a digitizing image scanner. A scanned image is digitally represented as rows and columns of bits which symbolize black and white “dots” of the page. The presentation of image as bits of black and white “dots” on a viewing screen or printer reproduces a human readable facsimile of the page.

International Telecommunication Union (ITU) Group 4: The International Telecommunications Union or ITU (previously known as the Consultative Committee for International Telegraph and Telephone - CCITT) is a United Nations standards organization that has developed protocols for the transmission of bi-tonal (black/white) images over telephone lines and data communication links. One of these protocols or standards for encoding bi-tonal or one-bit-per-pixel images is called Group 4, which is a “lossless” compression method developed specifically for scanned, bit-map images. Group 4 compression is the standard compression technique used for high-end fax machines and is the international standard for compression of bi-tonal scanned images in document and record management systems.

JPEG – Joint Photographic Experts Group. This compression standard is the product of an expert group organized by the International Standards Organization to develop a standard for compressing either full-color or gray-scale digital images and rendering these images for human viewing rather than by computer viewing.

TIFF – Tagged Image File Format: A proprietary image file format that is owned and maintained by Adobe, Inc. TIFF was originally developed by the Aldus Corporation for storing black and white images created by scanners and desktop publishing applications. It has gone through several revisions and extensions since its release in 1986, the most recent being TIFF 6.0, which was released in September 1995. The TIFF format is widely implemented internationally and consequently can be considered a de facto standard for scanned images.

PDF - Portable Document Format: File format for electronic document exchange that preserves document integrity so files can be viewed and printed on a variety of platforms. Owned and maintained by Adobe Inc. and adopted by governments and enterprises worldwide.

II. PAPER TO IMAGE CONVERSION

Approximately 150-200 Delaware government agencies, departments and groups have indicated an interest in converting ongoing and backfile paper records to digital images. The range of documentation types range from checks and small paper forms to very large engineering drawings. The following are guidelines designed to reflect available standards and to provide best practices for converting paper-based documents and records to digital images, including associated index metadata information.

SOURCE DOCUMENT FORMAT AND VOLUMES

One of the first steps in the process is to conduct an inventory of the records that are to be imaged. Being able to provide the imaging services vendor with this information is important in order to receive an appropriate “volume-related” price per image quote and to estimate the time required to perform the conversion.

The format (size, color, etc.) of the source paper records is important since odd sizes of documents and documents with color or gray areas may require different equipment and processes and also require special pricing.

If the conversion relates to an existing backfile of records, then the headings entitled “Docs/Day” and “Pages/Day” would be for the total document and page count of the backfile. The percentage of double-sided documents is also important to note since a different type of scanner and process may be required and each side would be counted as one image scanned for pricing purposes. Finally, providing a general quality indicator – poor, good, very good – (related to the overall readability and condition of the source document) is important to judge the level of document preparation and quality control that may need to be performed.

Table 1. Paper Formats and Volumes to be Imaged

Source Paper Format	Docs/Day	Pages/Day	% Double - sided	Source Quality
Checks and small forms				
½ letter (5x8.5 in.)				
Letter (8.5x11 in.)				
Legal (8.5x14 in.)				
B-size (11x17 in.)				
C-size drawing (17x22 in.)				
D-size drawing (22x34 in.)				
E-size drawing (34x44 in.)				
J-size drawing (34xunlimited length in inches)				
Gray scale documents and photographs				
Color documents				
Color photographs				
Books (other non-autofeed documents)				
Photos				
Other – please specify:				

DIGITAL IMAGE SCANNING

For hard copy documents (paper or micrographics), an image scanning device and associated software converts the document into digital bits that can be accurately stored, viewed and reproduced. The degree of accuracy and quality with which a hard copy document can be digitized is directly related to the resolution and threshold (contrast/brightness) settings of the scanner.

The level of scanning reproduction accuracy required will differ depending on the needs of the application. Documents that have a high degree of consistency regarding black to white contrast may only require a sampling to assure that an accurate and complete reproduction has been achieved. A higher and/or more intense level of quality assurance may be required where the black to white contrast varies considerably or in cases where a complex and/or fine-lined engineering or graphics record is being scanned. Therefore, the quality assurance procedures for validating the accuracy and completeness of reproduction are dependent on the type(s) of documents being captured.

Resolution

The legibility of scanned images depends largely upon the resolution at which the scanning occurs. Digital resolution is typically referred to as “dots per inch” (dpi) or “bits per inch” (bpi) and specifies the number of “bits” or picture elements (pixels) generated by a scanner in one horizontal and vertical inch of a document. The greater the number of dots per inch the greater is the likelihood that fine detail in small size type fonts and lines in drawings can be accurately captured and legibly reproduced.

The resolution at which pages are scanned also influences the file size of the resulting compressed image. For instance, a page scanned at 200 dpi and compressed at a 10:1 ratio will result in an image file size of 46,750 bytes. The same page scanned at 300 dpi and compressed at 10:1 will result in an image file size of 105,187.5 bytes or over twice as large as with 200 dpi. While the number one priority is to ensure that the scanning resolution is sufficiently high so that the resulting images are readily legible, the resulting file size may also be a consideration.

AIIM TR26-1993, “Resolution as it Relates to Photographic and Electronic Imaging,” suggests using a Quality Index (QI) as a reliable tool for predicting the legibility of scanned images based on various font size and required scanning resolutions. The Quality Index (see table below) is expressed numerically as:

- 3.6 – Marginal quality
- 5.0 – Quite legible with some loss of details or serifs
- 8.0 - Excellent with details resolved

For instance, the table below, which is extracted from AIIM TR26-1993, Annex C. Sample Quality Index Calculation (with examples), indicates that documents with smaller than 8 point type should probably be scanned at 300 dpi in order to achieve a “quite legible” QI of 5.0. However, the only sure method to determine the appropriate image scanning resolution for achieving acceptable legibility is to conduct one or more scanning tests of a representative sample of the documents to be scanned at various resolutions.

		Font size																		
		4			6			8			10			12						
height of small "e" in <i>Times Roman</i> font	(inches)	0.0275"			0.040"			0.055"			0.068"			0.078"						
	*metric	0.7 mm			1.0 mm			1.3 mm			1.7 mm			1.9 mm						
Quality Index value		3	3.6	5	8	3	3.6	5	8	3	3.6	5	8	3	3.6	5	8			
	required Digital Resolution (dpi) as predicted by above equation	218 dpi	261 dpi	363 dpi	561 dpi	150 dpi	180 dpi	250 dpi	400 dpi	110 dpi	130 dpi	182 dpi	280 dpi	88 dpi	105 dpi	147 dpi	235 dpi	77 dpi	92 dpi	128 dpi

*Arithmetic conversion from inches

The following guidelines should be followed for ensuring accurate (legible) reproduction of original documents:

- The Agency should submit – to one or more qualified conversion services vendors – a representative sample of documents to be test-scanned at 200 dpi and 300 dpi. The results of the test-scanning will be evaluated with the objective of identifying the resolution that provides an acceptable level of legibility, both as viewed on a higher resolution display screen (1280 x 1024 bits of resolution) or reproduced on at least a 300 dpi laser printer. For complex engineering drawings or very small font sizes (4 point type or smaller), it may be necessary to conduct test scanning at a resolution as high as 400 dpi.
- The minimum resolution for image scanning should be 200 dpi.
- A resolution of 300 dpi may be required if there are lower case letters smaller than 8 point type or if there are very fine lines in graphic or engineering drawings.

The best method to determine the appropriate resolution (or resolutions) for the records to be scanned is to submit a representative sample to the one or more qualified image conversion service vendors and then check the quality of the resulting scanned images.

Compression

Compression techniques are used in all document imaging systems to reduce a document image to a more manageable size for storage and transmission. The de facto industry standard compression scheme used by most all vendors is ITU Group 4, which is the worldwide standard used for compression and decompression in facsimile machine transmissions. These compression schemes can reduce most originally scanned letter-size documents by a factor of 10 to 15 times or from an original size of 467,500

bytes to approximately 30,000 to 50,000 bytes.

For documents with different shadings of gray (e.g., photographs) and for color documents, the compression standard called JPEG (Joint Photographic Expert Group) should be employed.

IMAGE QUALITY TEST SAMPLE

Best practice indicates that in order to achieve the best image and index quality, a sample of the range of documents to be imaged should be provided to the qualified image conversion vendor(s) for use in testing the document preparation, image scanning resolution and thresholds, and indexing methods. The sample of documents should be representative of what will be imaged during the full production operation. It also is helpful to include a sample of the “worst case” set of documents so that quality tests can be conducted regarding appropriate image scanning settings and methods to achieve the highest quality image.

The quality sampling should be done to ensure that the appropriate resolution has been selected, and to determine what special handling may be required, if any, for preparing the paper documents for scanning. Also, a test of the indexing for the selected documents should be done as a means of determining what the best methods and process (automated, manual or some combination) are that achieve the highest quality at a reasonable price.

INDEXING

Proper indexing of documents is critical for purposes of accurate and complete retrieval and for providing other means of controlling or distributing records. When determining the index attributes to associate with the scanned documents, the following potential needs should be considered:

- Provide ready access to a record or logically associated set of records (e.g. the complete contents of a case file).
- Identify documents by type (e.g., application, contract, invoice, etc.) and possibly by group (e.g., medical documents, correspondence, etc.) to allow for direct access to a document or logically related group of documents.
- For case files, e.g., contract files, human resource files, etc., select a common identifier that will logically group all documents related to the case file together.
- Where the document management system is to be linked to a legacy system, an attribute may need to be defined (which could be a case file attribute, a document group or an individual document attribute) that links to an attribute in the legacy database.
- A date that establishes the start date for the retention period. For “event-based” retention periods, consider leaving a “blank” date field that could be filled in – either programmatically or manually – at the time the event occurs.
- An attribute that may support establishing access security at the document level, such as for an individual or a group of medical records within a case file.
- An attribute that may help facilitate automated routing of the documents to a particular workflow or process step.

When establishing the document index, it generally is best, from a data entry cost, index maintenance, and storage space perspective, to limit the number of attributes to only those required to satisfy the considerations outlined above. Additional attributes may be helpful in providing a broader access capability to the documents, such as using both an employee ID number and last/first name as alternate search criteria.

A high accuracy level is required for the index information in order to ensure that accurate retrieval of the

records can be performed and that all logically associated records are properly linked. Best practice suggests that automatically reading index information from bar codes or using manual, double-blind keying of index fields is the most accurate method to capture the appropriate index information. Use of various index field verification methods, such as validation against a database, can help assure a more accurate indexing process.

The options to be considered for indexing include the following:

- Single manual keying from image (basic method 95-98% accuracy level).
- Double-blind manual keying where two different operators independently key the same index data, which is then compared for consistency and corrected if not consistent (typically achieves approximately 99.5% accuracy).
- Reading from bar codes that are affixed as a label, pre-printed on the form/page, or read from cover sheets (typically achieves approximately 99.99% accuracy – assuming the bar code itself is accurate). The highest bar code accuracy is achieved when it is possible to use an existing computer database or other digital source to pre-print the bar code on the page to be scanned, or on labels or cover sheets.
- Zoned optical character recognition (OCR/ICR) – where one or more specific areas of a page can be identified as containing the imaged information that is to be converted to text for use as an index attribute. If OCR/ICR is to be used for automatically capturing index information, then this process should be thoroughly tested prior to production to clearly understand the level of accuracy that can be achieved – which also will determine the level of quality control, correction capability and quality assurance that needs to be established.
- Use of edit checks to ensure that index field formats, values, etc. are correct.
- Database lookups (against an Agency-provided database) as a means of validating specific index values (e.g., social security number, personnel number, contract number, etc.).

QUALITY CONTROL AND ASSURANCE

The requirement for accurate electronic records is clearly stated in the Delaware Public Archives, Model Guidelines for Electronic Records:

7. Records which are created by the electronic records system must meet accepted definitions of accurate, understandable and meaningful records.

Summary: Electronic records must be:

1. accurate, in that there is a quality control check to ensure correct data.

Also, there is Federal case law that stipulates the importance of having an accurate and reliable process for capturing electronic records:

[T]he foundation for admission of computerized records consists of **showing the input procedures used, the tests for accuracy and reliability** and the fact that an established business relies on the computerized records in the ordinary course of carrying on its activities. The ... opposing party then has the opportunity to cross-examine concerning company practices **with respect to the input and as to the accuracy of the computer** as a memory bank and retriever of information ... [T]he court must “be satisfied with all reasonable certainty **that both the machine and those who supply the information have performed their functions with utmost accuracy.**”¹[bolded emphasis added]

¹ United States vs. Russo, 480 F.2d 1228, 1239 (6th Cir. 1973) (quoting United States v. De Georgia, 420 F.2d

In order to assure the highest quality of image scanning and indexing, the following areas should be given due consideration and agreed upon with the selected, qualified image conversion services vendor prior to the start of conversion:

Document Preparation

Document preparation should include steps that make the individual pages easy to scan and that produce high quality images. Document preparation includes removing paper clips and staples and repairing any pages (tears, folds, hard creases, etc.) that may be difficult to feed or not accurately read. It may also include making a copy of the page in order to produce a higher quality source document for scanning (such as in cases where the source document is tattered, of very poor contrast, e.g., when it has been copied multiple times or is a carbon copy). The Agency should assure itself that, at a minimum, the image conversion services vendor follows the technical report: ANSI/AIIM TR15-1997, Planning Considerations Addressing Preparation of Documents for Image Capture Systems.

Individual records or case files may then be accumulated into batches for scanning or scanned as individual case files or documents.

Image Scanning

The objective of image scanning is to capture the most accurate and complete digitized facsimile of the document as possible. The image conversion vendor should provide the following process(es) and methods for image scanning:

- Periodic testing of the scanner's proper functioning. The testing should at a minimum adhere to *ANSI/AIIM MS44-1988 [R1993], Recommended Practice for Quality Control of Image Scanners*, which includes a standard scan target page for checking the image quality for bi-tonal scanners
- Cursory operator review of individual images as they are being scanned
- Ability to stop and restart the scanning of a batch or document by the scanning operator
- Single or double-sided scanning
- Option for automatic elimination of blank pages, particularly when scanning double-sided pages that may contain a large number of "blank" back sides
- Ability to implement a digital scanning resolution that will ensure QI values of 8 (excellent with all detail resolved) for documents containing six point type font or higher.

Image Enhancement

Today's scanning technology can apply image enhancement techniques that improve the readability of scanned images. These automatic image enhancement techniques are an industry accepted method for enhancing images and are considered to meet legal requirements since it does not alter the actual content of the document or record.

Quality Control

It is important for the Agency and the image conversion vendor to agree on the quality level goals for images and index information prior to start of image scanning production. Conducting image and index quality testing prior to production scanning (see Image Quality Test Sample above) and setting goals and standards for image and index quality is the best practice for achieving highest quality (see Image Quality Sampling above).

The image conversion vendor should provide a thorough quality control process that is sufficient to meet the quality levels required by the Agency and the documents being imaged. The quality control levels applied should vary based on the training and experience of the individual operator and the level of ongoing quality being achieved (if quality levels fall or where, during operator training, a higher sampling of images and index information for quality control should be instituted). There are various points in the process where quality control should be applied:

- *Document preparation* – a percentage of documents/files prepared should be audited or reviewed, usually based on the training and experience level of each person performing the process.
- *Image scanning* – the scanner operator should perform a visual check of the images in real-time as they are being scanned to identify any major scanning problems. A separate image or image quality control process should also be employed, generally at the time of indexing. A page count from document preparation that is verified at the time of image scanning can help ensure that the complete document was scanned.
- *Indexing* – quality control can be done as part of a double-blind key entry process, or by conducting a visual review of the index information against the scanned images.

Rescan of Images

If the quality control determines certain images to be unacceptable, the image conversion vendor should provide the capability to rescan the images from the original source document.

Quality Assurance

No matter how stringently quality controls are applied by the imaging conversion vendor, quality assurance (after the scanning process is fully completed) should be performed on a sample of images and associated metadata to ensure that all scanning, indexing and quality control processes are being performed accurately and consistently. This Quality Assurance should be performed by personnel from the Agency requesting the image conversion, not by the image conversion vendor.

The quality assurance process generally entails reviewing a sample of the images and index information at the time it is returned by the image conversion vendor (or possibly after it has been loaded into the target document/content/record management system). The sample will vary depending on the stability of the process at the image conversion vendor's site and the level of errors that are being found in quality assurance. Generally a sampling between 3% and 10% of total images scanned is an acceptable range, with the lower sample size conducted when the image quality is generally high and the process is relatively stable.

DOCUMENTATION OF IMAGE CONVERSION SERVICE PROCESSES

The Agency should obtain assurance from the image conversion vendor that they have up-to-date and complete documentation of their conversion process, methods and steps. Accurate and up-to-date documentation of responsibilities and methodologies is a requirement stated in the Delaware Public Archives, Model Guidelines for Electronic Records:

2. Electronic records systems must have accurately documented policies, assigned responsibilities, and formal methodologies for their management.

TRANSFER MEDIA AND FORMAT FOR IMAGES AND INDEX DATA

The media and format to be used for delivery of scanned images and associated index information should be determined by the Agency and specified to the image service provider. The scanned images or digital

computer output provided by the conversion services vendor will normally be loaded into the document/content/records management system or computer output storage and retrieval system used by the Agency. Therefore, the media and the format of the images and index information will be dependent on the requirements of the document/content/records management system employed by the requesting Agency.

Transfer Media or Methods

Examples of transfer media or methods are:

- a. Compact Disc (CD)
- b. Digital Versatile Disc (DVD)
- c. DLT magnetic tape cartridge
- d. Download (e.g., via direct link or Secure File Transfer Protocol [SFTP] via the Internet) of images and index information to a designated Agency location or directory. The Agency should ensure that the SFTP communications connection with the imaging services vendor has the appropriate level of security to maintain the integrity and the necessary confidentiality of the records.
- e. Other - any unique format or media requirements should be specifically defined and described to the conversion services vendor.

If a copy of the media used for transferring scanned images is to be provided to DPA for archival or preservation purposes, the required formats are those currently designated as acceptable transfer media:

- a. Compact Disc – Read Only Memory (CD-ROM), ANSI/NISO/ISO 9660-1990.

Output Format

The output format of the scanned images will be TIFF Version 6 compressed using ITU Group 4 compression. However, the format of the file(s) that are used by the conversion services vendor to deliver the images and index data to the Agency will depend largely on the type of document imaging, document management or content management systems employed by the Agency. (e.g. FileMagic/Fortis, OTG Software (Legato), DocStar, FileNET, etc.). The Agency should obtain the appropriate formats from the document/content/records management system supplier and communicate those to the image conversion service vendor.

LOCATION FOR PERFORMING IMAGE CONVERSION SERVICES

Consideration should be given to where the image conversion services will be performed. Generally an imaging services vendor can provide the lowest cost by imaging at one of their service center sites. However, there may be considerations where the imaging services vendor could be requested to come to the site of the requesting Agency to perform the imaging conversion services, such as:

- There is a relatively large backfile conversion to be performed and it would be very inconvenient for the documents to be sent off-site during the conversion, (although most image conversion service vendors provide “on-request” retrieval of records at their site with fax or courier delivery).
- When the records to be converted are of such a sensitive nature (e.g., medical or other personnel records) that it would be highly preferable for the source records to remain in the physical control of the Agency.

TRANSPORT/RECEIPT AND RETURN OF SOURCE DOCUMENTS

Consideration should be given to mode of transportation used to transport the source records from the Agency

to and from the image conversion service site in order to ensure that the records are moved in an expedient manner while fully protecting their completeness, integrity and confidentiality.

PHYSICAL SECURITY AND MAINTENANCE OF SOURCE RECORDS

The Agency should ensure that the image conversion services vendor provides the appropriate levels of security at their conversion site (or on the premises of the Agency) that will protect the integrity and confidentiality of the source documents and maintain their completeness while undergoing conversion.

THIRD PARTY SERVICE EMPLOYEES

Agencies should ensure that the employees of third party service providers who transport source documents, inspect source documents, image or index from source documents or images thereof, and review images of source documents for quality control are bonded and have a state and federal background check to ensure that confidential information is not compromised. While this was a requirement for establishing qualified vendors, each Agency should assure itself that these qualifications are still valid. The security requirements of some agencies may entail additional security arrangements.

III. COMPUTER OUTPUT (COLD/ERM) SERVICE REQUIREMENTS

Many Delaware government agencies, departments and groups have paper or computer output microfiche reports that they are interested in converting to digital format for storage, retrieval and retention. This section provides guidelines for determining requirements and working with the conversion services vendor for converting computer output reports to a digital format, typically referred to as computer output to laser disk (COLD) or enterprise report management (ERM).

COMPUTER OUTPUT OR REPORTS, FORMATS AND VOLUMES

The specific computer reports or documents to be converted should be listed, along with the source format of the report, or the volume per period (daily, weekly, monthly, annually, etc.). This information will be required for the conversion services vendor to provide a quote and to understand the overall conversion volume requirements, the types of computer output formats that are to be converted and the anticipated periodic volume of the output. The following table provides a format for documenting and conveying the possible formats and volumes that would be submitted to the conversion services vendor for processing.

Table 2. Computer Output or Reports, Formats and Volumes

Report Title (Filename)	Source Computer Output Format*	Pages/Period	Period**
EXAMPLE			
General Ledger (GL013103)	Line Data (preformatted ASCII with line feeds)	5,000	monthly

* Source Computer Output Formats may include: Line Data (pre-formatted ASCII with line feeds); IBM Advanced Function Printing (AFP); Xerox DJDE/Metacode; PCL; PostScript; Multi-font; Office Productivity Documents, e.g., MS Word; etc.

** e.g., daily, weekly, monthly, annually, etc.

INDEX DEFINITION FOR EXTRACTION

Index metadata must be defined to be able to search for the report once it is converted to digitally accessible format. Once the index attributes for each report are defined, the conversion service vendor will perform a one-time setup of the report that identifies the index metadata to be automatically extracted. Agencies should give consideration to the following guidelines when defining index metadata:

- When a report is relatively small (1,000 pages or less), indexing the report only by the report title is generally sufficient. Once the report is located and retrieved from the storage media, a full-text search capability can be employed to find the appropriate page(s).
- If the report is large (1,000 pages or larger) or if the report contains discrete document-based information (e.g., invoices, statement of benefits, 1099s, etc.) then it is best to define specific index attributes at the page or document level that will allow searching and locating a page or

specific set of pages.

Attributes such as invoice number, employee number or SSN can be used to more easily locate a specific page or document. Also, the “print date” of each page or document should be captured so that date or date-range searches can be performed as required.

- Also, if the report or document information is confidential, the report name or other specific attributes, e.g., number or type of report or specific page-level attributes can be used to restrict access to the report.
- Other index attributes may be defined that will further ease the searching and access control of the reports or selected pages of a report.

REPORT SETUP

Once the reports have been identified and the corresponding index information for each report has been defined, a one-time report setup is performed to allow the report conversion software to automatically identify the report and to also automatically extract the prescribed index attributes from the report. While the conversion services vendor can perform this setup, the Agency personnel could also perform this task with proper training by the conversion services vendor or the vendor of the target COLD/ERM system that will be used to search for and display the reports.

QUALITY CONTROL

The Agency should ensure that the vendor is regularly performing quality checks on the converted digital report output to ensure that:

- All reports submitted to the conversion service provider were indeed converted.
- That the microfiche label is being accurately produced with the correct index information.
- That any index frame being produced is complete and accurate for the images located on the microfiche.
- That the “as-if-printed” formatting of the report pages is being done correctly.

QUALITY ASSURANCE

The Agency should perform quality assurance after receipt of the converted reports to make sure that:

- All reports submitted to the conversion service provider were indeed converted.
- Conduct a sample testing of the converted reports to ensure that all pages of a report are retrievable and that the appropriate index information was accurately and completely extracted and associated with the converted report or documents.

REPORT PAGE CONVERSION OPTIONS

Generally, the report pages are converted and stored in the format that the Agency submits them to the conversion services vendor, e.g., ASCII text or ASCII text with appropriately linked or overlays of graphics. However, report pages can be converted from the submitted format to other formats that may be easier to store and retrieve and also may be required to meet certain compliance or recordkeeping requirements.

The typical converted formats are:

- PDF for easier viewing via a Web browser or to provide a more technology neutral format for longer term preservation.

- XML – for certain reports where the identification of selected fields may provide value.

If the converted reports are to be retained for archival purposes, it is recommended that the file format be the “as-if-printed” form of the report formatted as either plain ASCII text (with all line feeds, etc. converted) or PDF.

STORAGE AND RETRIEVAL OF DELIVERED COMPUTER OUTPUT

Once the reports are converted and the index information has been extracted, the reports will be put in an appropriate format for retrieval. The Agency should determine if they have any preference regarding the type of media that is used for storage of the converted reports. The following media alternatives are generally provided by a conversion services vendor:

- Self-contained CD or DVD with report pages, index data and retrieval, viewing, printing, faxing software contained on the storage media.

- A format that can be used for loading delivered computer reports to a magnetic disk directory where pre-loaded software will then be used to provide retrieval, printing and faxing.

- Specify which search functions are provided using the extracted index attributes as a means of locating a report or specific report pages.
 - Search using one or more specified index keys
 - Full text or string search
 - Boolean search (AND, OR, etc.)

- Determine what output functions are required for retrieval and reproduction of reports or report pages.
 - Print the full report or selected pages
 - Fax the full report or selected pages
 - E-mail the full report or selected pages

DELIVERY OR TRANSFER

Define the means that should be used to deliver the converted reports to the Agency (this is somewhat dependent on the choices for retrieval made above).

Examples of transfer media or methods are:

- a. Compact Disc (CD)
- b. Digital Versatile Disc (DVD)
- c. DLT magnetic tape cartridge
- d. Download (e.g., direct link or Secure File Transfer Protocol [SFTP] via Internet) of images and index information to a specified Agency location or directory. The Agency should ensure that the SFTP communications connection with the imaging services vendor has the appropriate level of security to maintain the integrity and the necessary confidentiality of the records.

e. Other - any unique format or media requirements should be specifically defined and described to the conversion services vendor.

If a copy of the media used for transferring scanned images is to be provided to DPA for archival or preservation purposes, the required formats are those currently designated as acceptable transfer media:

a. Compact Disc – Read Only Memory (CD-ROM), ANSI/NISO/ISO 9660-1990.

TRANSPORT/RECEIPT OF COMPUTER REPORTS

The modes of transportation provided by the conversion services vendor for moving the digital computer report data from the Agency to the conversion service site must ensure that the records are moved in an expedient manner while fully protecting their completeness, integrity and confidentiality.

PHYSICAL SECURITY AND MAINTENANCE OF SOURCE RECORDS

The conversion services vendor should provide the means and method of security at the conversion service site to protect the integrity and confidentiality of the computer output data while at the conversion site.

THIRD PARTY SERVICE EMPLOYEES

Agencies should ensure that the employees of third party service providers who transport source documents, inspect source documents, image or index from source documents or images thereof, and review images of source documents for quality control are bonded and have a state and federal background check to ensure that confidential information is not compromised. While this was a requirement for establishing qualified vendors, each Agency should assure itself that these qualifications are still valid. The security requirements of some agencies may entail additional security arrangements.

IV. COMPUTER OUTPUT TO MICROFICHE

Delaware government agencies, departments and groups may choose to have computer page formatted records converted to micrographic images (COM), rather than digitized images. This section provides guidelines for determining requirements and working with the conversion services vendor for converting computer page formatted records to computer output microfiche, which is generally known as COM.

OVERVIEW

Microfiche are sheets of film. They are usually four (4) inches by six (6) inches and contain images set in a grid pattern in rows and columns, which reflect the reduction ratio used. A 42X reduction ratio, which is preferred, produces a fiche with 98 frames or images. An eye-legible title along the top of the microfiche permits instant identification without the need for magnification. Standard microfiche are produced according to strictly controlled layouts which permit the retrieval of any image via a coordinate representing the intersection of the row and column where it is located. At least one index frame is usually included on the microfiche to provide access to its content without the need for an external index. Microfiche should be stored in plastic acid-free paper in accordance with *ANSI/NAPM IT9.2-1991 and ANSI/NAPM IT9.11-1993*.

Converting computer output to microfiche involves the transfer of the output pages to microfiche at a 42X reduction. There are two different processes for doing this. One involves only ASCII line printer output and uses an alphanumeric COM quality test form slide. ANSI/AIIM MS1-1996, *Standard Recommended Practice for Alphanumeric Computer-Output Microforms –Operational Procedures for Inspection and Quality Control* provides guidelines and recommended practices for this process. The second process also involves ASCII line printer output along with bit map images (logos) and vector graphics. This second process uses software generated COM quality test forms to establish a quality base line and for periodic testing of COM image quality. ANSI/AIIM MS62-1999, *Recommended Practice for COM Recording Systems Having an Internal Electronic Forms Generating System – Operational Practices for Inspection and Quality Control* describes standard methods that should be followed when using this COM technology. The selected conversion services vendor should adhere to the standards described in these two processes, depending on the type of computer output that is being converted.

Only wet chemistry film processing is acceptable. The images should be written on silver halide on 105 mm (fiche), for the master and first-generation copy. The film used must comply with ANSI/AIIM MS14-1988 (R1996), *Specifications for 16 mm and 35 mm Roll Microfilm*.

Film processing is done either off-line or in-line. Off-line processing is a separate operation in which the film processors are not connected directly to the COM recorder. Most conversion services vendors use in-line film processing, whereby the microfiche are processed as an integrated part of the COM recorder. Typically, in-line film processing does not require continuous replenishment or replacement of chemicals. ANSI/AIIM MS1-1996, *Standard Recommended Practice for Alphanumeric Computer-Output Microforms –Operational Procedures for Inspection and Quality Control* provides guidelines and recommended practices for this off-line and in-line processing. Regardless of the type of film processing, the Methylene Blue test must be conducted in accordance with ANSI/NAPM IT9.17 – 1993, *Determination of Residual Thiosulfate and other Related Residual Chemicals in Processed Photographic Materials – Methods Using Iodine-Amylose, Methylene Blue, and Silver Sulfide*.

COMPUTER OUTPUT OR REPORTS, FORMATS AND VOLUMES

The specific computer reports or documents to be converted should be listed, along with the source format of the report and the volume per period (daily, weekly, monthly, annually, etc.). This information will be

required for the conversion services vendor to provide a quote and to understand the overall conversion volume requirements, the types of computer output formats that are to be converted and the anticipated periodic volume of the output. The following table provides a format for documenting and conveying the possible formats and volumes that would be submitted to the conversion services vendor for processing.

Table 3. Computer Output or Reports, Formats and Volumes

Report Title (Filename)	Source Computer Output Format*	Pages/Period	Period**
EXAMPLE			
General Ledger (GL013103)	Line Data (preformatted ASCII with line feeds)	5,000	monthly

* Source Computer Output Formats may include: Line Data (pre-formatted ASCII with line feeds); IBM Advanced Function Printing (AFP); Xerox DJDE/Metacode; PCL; PostScript; Multi-font; Office Productivity Documents, e.g., MS Word; etc.

** e.g., daily, weekly, monthly, annually, etc.

RESOLUTION

The digital resolution of the computer output files being used as input to the COM process do not translate directly into optical resolution and reduction because alphanumeric-only COM recorders cannot generate patterns equivalent to ISO Test Chart No. 2.

The Agency should work with the COM conversion services vendor to use quality test forms to compare characters in the test target with those in computer output to establish a relative quality level. Figure 5 in ANSI/AIIM MS1-1996, *Standard Recommended Practice for Alphanumeric Computer-Output Microforms –Operational Procedures for Inspection and Quality Control* is a Quality Index graph that should be used to determine the appropriate resolution pattern, which then can be used to estimate the number of dots per inch the image generator should produce. Annex A of AIIM TR26-1993, *Resolution as it Relates to Photographic and Electronic Imaging* has a table for conversion of resolution patterns into dpi requirements.

INDEXING

Microfiche must be appropriately indexed to facilitate timely and accurate retrieval. The label at the top of each fiche must be human readable and, at a minimum, identify the subject matter (case file number, social security number, invoice number, and the like) and the date of the report. One frame, usually the first frame, will contain indexing information, such as the date range, number range, and other relevant information that will facilitate manual searching for a specific frame. Generally, this information must be manually keyed in. In some instances agencies will attach bar codes (code 3 of 9 with ASCII text) to documents that contain indexing information about the file name or identifier that can be captured and used to automatically generate index for retrieval purposes.

REQUIRED MINIMUM MICROFILMING STANDARDS

The DPA “Required Minimum Microfilming Standards for Public Records (July 2001)” defines three categories of microfilming: (1) Preservation microfilming, which includes archival documents already in its custody or that are transferred to its custody (county court records); (2) Archival microfilming encompassing records that are scheduled for permanent retention by DPA; (3) Administrative microfilming that includes records that are not scheduled for permanent retention by DPA but that an agency may be required to retain for longer periods of time. DPA has issued microfilming standards for public records that incorporate by reference all relevant ANSI/AIIM microfilm standards and technical reports. In addition, these standards identify certain features that all microforms made for archival or administrative purposes should include. However, these features impose certain requirements that cannot be achieved with COM (e.g., ISO Test No. 2). Consequently, vendors should be required to specify which features of the DPA Minimum standards will not be implemented and offer an alternative solution.

QUALITY CONTROL

The Agency should ensure that the selected conversion service vendor has set up the COM system and process according to ANSI/AIIM MS1-1996, *Standard Recommended Practice for Alphanumeric Computer-Output Microforms –Operational Procedures for Inspection and Quality Control*. Further, the Agency should ensure that the documented processes of the conversion service vendor include conducting regular and ongoing quality control of resulting microfiche to ensure that the quality and accuracy of the process is being maintained.

QUALITY ASSURANCE

The Agency should, at a minimum, review the first and last sheets of microfiche for each converted report to ensure that they are legible and that they meet the guidelines for microfilm quality as set forth in the *DPA Required Minimum Microfilming Standards for Public Records (July 2001)*.

To ensure that the quality level expected by the Agency has been achieved, the following quality assurance steps should be performed:

- Review the first and last sheets of microfiche for each converted report to ensure they are legible.
- Ensure that all reports submitted to the conversion service provider were indeed converted.
- Conduct a sample testing of the converted reports to ensure that all pages of the report are retrievable and that the appropriate index information was accurately and completely extracted and associated with the converted report or documents.
- Ensure that the microfiche output by the conversion services vendor is consistently meeting the guidelines for microfilm quality as set forth in the *DPA Required Minimum Microfilming Standards for Public Records (July 2001)*.

PHYSICAL SECURITY AND MAINTENANCE OF SOURCE RECORDS

The conversion services vendor should provide the means and method of security at the conversion service site to protect the integrity and confidentiality of the computer output data while at the conversion site.

THIRD PARTY SERVICE EMPLOYEES

Agencies should ensure that the employees of third party service providers who transport source documents, inspect source documents, image or index from source documents or images thereof, and review images of

source documents for quality control are bonded and have a state and federal background check to ensure that confidential information is not compromised. While this was a requirement for establishing qualified vendors, each Agency should assure itself that these qualifications are still valid. The security requirements of some agencies may entail additional security arrangements.

Effective December 1, 2003