

Digital Images Legal Guidelines

This document contains Section 4 of the Ohio Revised Digital Imaging Guidelines for State of Ohio Executive Agencies and Local Governments.

All four sections can be seen at www.ohiojunction.net/erc/imagingrevision/revisedimaging2003.html

4.0 Archiving and Long-term Maintenance

4.1 Provide specific plans for creating and sustaining digital images that will be retained more than 10 years.

Public officials are responsible by law for ensuring that their records are protected and accessible for the time period stipulated in the records retention schedule. This responsibility applies regardless of the storage media of the records. It is the responsibility of the agency to decide on which medium to maintain their records. Should an agency decide to destroy original records once imaged, the agency must ensure that their imaged records are accessible for the time period stipulated on their records retention schedule.

If digital images need to be maintained for long periods of time, i.e., more than ten years, it will be necessary to take several steps to ensure accessibility. These steps include requirements for master image capture (scanning specifications, file formats, metadata); system information and maintenance (system documentation, copying and refreshing media); and sustainability.

4.2 Scanning Specifications and File Formats

For images that need to be retained ten years or more, it is recommended that originals be scanned at a minimum of 300 dpi and saved as a TIFF (Tagged Image File Format) Group 4 or higher. Currently this would include TIFF 4, TIFF 5, and TIFF 6. Best practice indicates that it is preferably to use the most recent version of the file format (currently TIFF 6, soon TIFF 7). (Other file formats are currently under development that may be appropriate in the future for the creation of a Digital Master. These may include PDF-A and JPEG2000.)

This TIFF will serve as a "master image" or "archival copy" that is similar to a microfilm master negative. The master image should capture as much information as possible from the original in order to serve as a long-term, high-quality digital version. Derivative images for use on the web or within the application itself should be made from the master image. Master images can be bitonal, grayscale or color. Quality control applied to master images should be inten-

sive. This is especially true if an agency is retaining ONLY a digital image (no paper or microfilm copies) of a record with permanent retention.

The TIFF images should be stored uncompressed, or the compression used should be lossless. If it is absolutely necessary to compress the master image files, current industry standards recommend the following:

TIFF (Tagged Image File Format) with CCITT Fax 4

Compression: Better suited to bitonal text documents. This format can provide a high level of detail combined with a smaller file size. May be used as a master image file format.

TIFF (Tagged Image File Format) with LZW

Compression: A 24-bit, lossless compression format, commonly used by Adobe Photoshop and other image editing software. Used to store color and grayscale files. May be used as a master image file format.

The TIFF images should be stored in a secure and stable environment, preferably offline. Other derivative images can be created from the TIFF files to enable web access.

4.3 Digital Image Metadata

Metadata is simply defined as data about data. More specifically, metadata consists of a standardized structured format and controlled vocabulary that allow for the precise description of record content, location, and value. Metadata often includes, but is not limited to, attributes like file type, file name, creator name, date of creation, and use restrictions. Metadata capture, whether automatic or manual, is a process built into the actual information system.

Note that TIFF file headers can contain system-generated metadata. You should be aware of which metadata elements your scanning software is able to use and which viewer you plan to use. Please refer to the National Information Standards Organization's Draft Data Dictionary - Technical Metadata for Digital Still Images. Although this is a draft standard, it may be considered the most current Best Practice for this topic.

4.4 System Information and Maintenance

On a system level, documentation is information about planning, development, specifications, implementation, modification, and maintenance of system components (hardware, software, networks, etc.). System documentation includes such things as policies, procedures, data models, user manuals, and program codes. Documentation capture is not a system process.

Please refer to the Ohio Electronic Records Committee Trustworthy Information Systems Handbook Section 9 for more complete information on system documentation.

4.5 Copying and Refreshing Media

Media copying and refreshing are essential for all digital media to avoid degradation and to facilitate longer-term preservation strategies. This involves periodically copying data onto identical media to address media degradation and impermanence, and periodically reformatting the data from an obsolete storage device to a newly emerging one, in some cases bypassing the intermediate generation that is mature but at risk of becoming obsolete

Media copying and refreshing should take place as follows:

- Within the minimum time specified by the supplier for the media's viability under prevailing environmental conditions.
- When new storage devices are installed.
- When an audit discloses significant temporary or read errors in the resource.

4.6 Sustainability

Hardware, software, and file formats could be operational for ten years or more, but technology will often be superseded within two to three years. If the system stores records with retention periods exceeding the life span of the hardware and software, it is essential that the application or system administrator plan for data sustainability. A digital sustainability strategy documents how an organization will transfer long-term or permanent records from one generation of hardware and software to another generation while maintaining system functionality and data. The strategy should be written and available with current system documentation and should be updated when technology changes. It is important to budget for these costs.

4.7 Reformatting and Migration

Copying and refreshing will not ensure that the system remains accessible. New software, platforms, and file formats will need to be utilized in order to facilitate long-term accessibility and reliability of the records. Further steps will need to be taken that may include the following:

- Reformatting of existing file formats to appropriate newly emerging formats.
- Migrating one component of the system, such as the database that provides indexing information to the system, to a new hardware and/or software platform.
- Migrating the whole system from one hardware and/or software platform to another.
- Documenting the changes made to the hardware, software, and file formats. Include changes that could affect data viability such as moving to a new file format or moving indexing information to a new database system.

Reformatting and migration should take place when the existing file formats, software, platforms or systems are no longer viable, usually due to obsolescence or the necessity of enhanced system performance.

Reformatting and migration will be much easier if the technical specifications of the system and the metadata relating to the digital images are created appropriately. For systems, this requires creating and maintaining System Documentation.

4.8 Scheduling

Agencies must submit a revised records retention schedule for records that they are reformatting, i.e., changing medium from paper to digital image. Agencies that are considering imaging records with permanent retention periods should contact the State Archives for a system and records analysis to determine if maintaining the records in an eye readable format (paper, microfilm, etc.) may also be necessary.

##