

Procedures for the Transfer of Permanent Electronic Records: Records in Portable Document Format (PDF)

1.0 Preface

As part of the Kentucky Department for Libraries and Archives' (KDLA) electronic records initiative, and in cooperation with other state agencies, KDLA is issuing guidance for transferring permanent electronic records to KDLA. This guidance expands currently acceptable formats to enable the transfer of records in Portable Document Format (PDF) to KDLA.

This guidance provides transfer requirements for:

- all records in PDF (sections 3.1 and 3.2),
- records converted to PDF from their native electronic formats (e.g., office automation products) (section 3.3),
- records converted to PDF from scanned paper or image formats such as TIFF (section 3.4).

Additionally, sections 3.5 and 4.0 provide transfer requirements including transfer documentation and related information on how to transfer PDF records to KDLA.

1.1 Effective Dates

This guidance applies to all records in PDF format that have been appraised or scheduled for permanent retention and transfer to KDLA on or after March 3, 2005.

KDLA recognizes that legacy records and records whose disposition is changed from temporary to permanent may present unique circumstances for agencies. Any agency that has permanent records in PDF format that do not meet the requirements in this guidance should contact KDLA to determine the most appropriate medium and format for transfer (see section 6.0).

2.0 Background

PDF is a priority electronic records format identified by KDLA and partner agencies as part of the Electronic Records Management (ERM) initiative. A major goal of this initiative is to provide the tools for agencies to access electronic records for as long as required and to transfer permanent electronic records to KDLA for preservation and future use by government and citizens.

725 KAR 1:040 (2)(2) requires agencies to transfer state publications to KDLA in PDF format (in lieu of paper copies). PDF is also the state's Enterprise IT Architecture Standard for non-alterable electronic documents. (See **Enterprise IT Architecture Standard: 2330 Electronic Documents – Non-Alterable** located at: <http://gotsource.ky.gov/dsweb/Get/Document-9351/2330+++Electronic+Documents++Non-Alterable.doc>)

KDLA will issue other transfer guidance for other electronic records formats under this initiative. These guidance documents and additional information about the ERM E-Gov initiative can be found on KDLA's web site.

3.0 Transfer Requirements for PDF Records

KDLA will accept transfers of PDF records that have been scheduled as permanent records on an approved records retention schedule. Any agency that has permanent records in PDF format that do not meet the requirements in this guidance, should contact KDLA for further instructions (see section 6.0).

To facilitate preservation processing and future access to these records, agencies must comply with the following *minimum* requirements:

3.1 PDF File Specification for All PDF Records

3.1.1 PDF records must comply with PDF versions 1.0 through 1.4 (i.e., all existing PDF versions as of the effective date of this guidance) and meet the requirements outlined in sections 3.2 through 3.4.

3.1.2 KDLA periodically will update the list of acceptable PDF versions as required.

3.2 General Requirements for All PDF records

3.2.1 Security Requirements

3.2.1.1 PDF records must not contain security settings (e.g., self-sign security, user passwords, and/or permissions) that prevent KDLA from opening, viewing or printing the record.

3.2.1.2 In addition, PDF records must have all security settings deactivated (e.g., encryption, master passwords, and/or permissions) prior to transfer to KDLA. Deactivating security settings ensures KDLA's ability to support long term migration and preservation of the records.

Note: Agencies may use appropriate security settings (as identified in the Enterprise IT Architecture Standards and Policies) up to the time of transfer to KDLA.

3.2.2 Review of Special Features

Because of the complexities associated with certain PDF features, KDLA will review PDF records containing special features on a case-by-case basis when the records are scheduled. Examples of special features include but are not limited to: digital signatures; links to other documents, files or sites; embedded files (including multimedia objects); form data; comments and/or annotations.

3.3 Requirements for Records Converted to PDF from Their Native Electronic Formats (e.g., office automation products)

3.3.1 Electronic records that have been converted to PDF from their native electronic formats must include embedded fonts to guarantee the visual reproduction of all text as created. All fonts embedded in PDF records must be publicly identified as legally embeddable (i.e., font license permits embedding) in a file for unlimited, universal viewing and printing.

3.3.1.1 PDF records that reference fonts other than the "base 14 fonts"¹ must have those fonts referenced in the record (i.e., as a minimum, subsets of all referenced fonts) embedded within the PDF file.

3.3.1.2 PDF records created must have all fonts referenced in the record, including the "base 14 fonts," embedded within the PDF file. This requirement is met by having, as a minimum, subsets of all referenced fonts embedded within the PDF file.

3.4 Requirements for Scanned Paper or Image Formats Converted to PDF

3.4.1 Scanned images of textual paper records converted to PDF must adhere to the requirements in the *Procedures for the Transfer of Scanned Images of Textual Records (Scanned Images Transfer Guidance)*.

Any agency that has PDF records that have not been scanned according to the minimum image quality specifications in the transfer guidance, should contact KDLA for further instructions (see section 6.0).

3.4.2 PDF records that contain embedded searchable text based on Optical Character Recognition (OCR) must be identical in content and appearance to the source document. KDLA understands that the ability to embed OCR'd text in PDF records enhances access to the records. While KDLA will accept PDF records with uncorrected OCR'd text, it will not accept PDF records resulting from OCR processes that either alter the content or degrade the quality of the original bit-mapped image.

3.4.2.1 KDLA will accept PDF records that have been OCR'd using processes that do not alter the original bit-mapped image. An example of

¹ The base 14 fonts are: Courier (Regular, Bold, Italic, and Bold Italic), Arial MT (Regular, Bold, Oblique, and Bold Oblique), Times New Roman PS MT (Roman, Bold, Italic, and Bold Italic), Symbol, and ZapfDingbats.

an output process that accomplishes this requirement is Searchable Image - Exact.

3.4.2.2 KDLA will not accept PDF records that have been OCR'd using processes that substitute OCR'd text for the original scanned text within the bit-mapped image. Such OCR processes may involve loss of data through imprecise interpretation of scanned characters. Examples of output processes that use this prohibited technique include Formatted Text and Graphics and PDF Normal.

3.4.2.3 KDLA will not accept PDF records that have been OCR'd using processes that use lossy compression to reduce file size (e.g., JPEG). Such OCR processes degrade the quality of the original image and may make such images unsuitable for archival preservation. An example of an output process that uses this lossy compression technique for color and grayscale images is Searchable Image - Compact.

3.5 Transfer Documentation

This guideline supplements transfer documentation requirements already in place to ensure that transfers of records in PDF format are clearly identified and described. Agencies must submit a signed transmittal form that can be found on the KDLA website (<http://www.kdla.ky.gov/recmanagement/schedules/kyrectransmittal.pdf>).

3.5.1 For each transfer, agencies must supply documentation that identifies the software used to create the PDF records (if available) and the version(s), and the operating system (if available) and version(s).

3.5.2 Agencies must provide all external finding aids for the transferred PDF records (e.g., indexes; descriptive, administrative, or technical metadata; and/or databases of OCR'd text) in formats approved by KDLA, with the appropriate documentation.

3.5.3 When an agency has developed standards or guidelines to assist in formatting, validating, or accessing PDF records (including recommended software or quality settings, and/or guidelines for embedding metadata within PDF records), a copy of these standards or guidelines must be included with the transfer.

3.5.4 PDF records converted from scanned images also must adhere to the transfer documentation requirements in section 3.3 of the Scanned Images Transfer Guidance.

4.0 TRANSFER MECHANISMS

4.1 Agencies may transfer PDF records using any of the approved media or methods listed in section 4.0 of the *Guidelines for Transferring Records to KDLA*.

4.2 PDF records must not be compressed (e.g., Winzip, PKZIP) or aggregated (e.g., TAR) for purposes of transfer unless KDLA has approved the transfer in compressed or aggregated form in advance. In such cases, KDLA may require the agency to provide the software to decompress the records.

5.0 LEVELS OF ACCESS

KDLA will provide access to the creating agency and to all researchers requesting PDF records accessioned from state agencies, subject to review of content for Open Records Act exemptions as is feasible. While compliance with these requirements will improve future access to accessioned PDF records, KDLA's ability to provide access to certain records will vary according to their hardware and software dependencies. At the present time, KDLA provides users with a copy of fully releasable electronic record files on any of the media currently approved by KDLA. For PDF records transferred to KDLA the user will be responsible for obtaining the necessary hardware and software to view the records.