Document Storage Workgroup Agenda
August 6, 2015
9:00-11:00

George Edgecomb Courthouse
Judicial Conference Room, 6th floor
800 East Twiggs Street, Tampa, 33602

    I.   **Welcome** – *Steve Shaw, Document Storage Workgroup Chair*

    II.   **Review Proposed Standards for Electronic Document Submissions** – *Group Discussion*

    III.   **Other Items/Wrap Up** – *Group Discussion*

    VII.   **Adjourn**

Proposed: Standards for Electronic Document Submissions 07/31/2015

1. **Document Size and Content Standards for Original Filing Documents**
   a. Limit document file size to 50MB. (This increases the portal submission size from 25MB to 50MB.)
   b. Embedded images should not exceed a resolution of 600 dpi.
   c. Scanning of documents is discouraged. If necessary, the document should be scanned in monochrome (black and white) mode with resolution of 300 dpi.

2. **Document Format-PDF/A**
   a. Documents submitted to the E-Portal must be in PDF/A format. The filer may use either the 1a or the 1b level of conformance as described below. Although this standard generically refers to the functionality of PDF/A, the functionality may differ depending on whether 1a or 1b file type is used.
   b. Justification: Within the PDF standard, only PDF/A is designed for long term archiving that preserves the document format and appearance.
      i. PDF/A ISO 19005 provides specifications for the creation, viewing, and printing of digital documents used for long-term preservation. PDF/A preserves and protects final documents of record as self-contained files. References to external content are not allowed because the referenced items might not exist in years to come. The proposed Florida courts PDF standard is PDF/A-1a or PDF/A-1b which are based on the PDF v1.4 specification. This format is:
         1. Self-contained—everything needed to render or print a PDF/A file is contained within the file.
         2. Self-documenting—PDF/A promotes the use of metadata, enhancing the document by providing information about the document itself. For example, it provides recommendations for documenting file attributes such as the file identifier, file provenance (i.e., the history and genealogy of the document or file), and font metadata. When metadata is used, PDF/A requires the use of the Adobe Extensible Metadata Platform (XMP) for embedding the data in the files.
         3. Device-independent—PDF/A requires device-independent components, such as specific RGB or CMYK color profiles, so that the visual appearance can be reliably and consistently rendered and printed without regard to the hardware or software platform used.
         4. Encryption Prohibited. A compliant PDF/A file must be open and available to anyone or any software that processes the file. User IDs and passwords cannot be embedded.
   c. **ADA Compliance**
      i. **PDF/A.** PDF/A allows for but does not guarantee ADA compliance. Full ADA compliance is only accomplished when the author creates the document to be ADA compliant in its original document authoring tool and adds all necessary tags to graphic objects, the necessary structure to embedded tables, and otherwise makes the document screen reader friendly. ADA compliance is the responsibility of the creator of the document.
      ii. **PDF/A-1a** (accessible). Documents submitted to the E-Portal should be in PDF/A-1a format. This requires the author to create the document as PDF/A-1a from the original document authoring tool.
         1. Within the PDF/A group of standards, PDF/A-1a inherently provides searchability and basic structure, and is generally screen reader friendly. When documents are saved to PDF/A-1a from a document authoring tool, they are ADA compliant to the degree that the author created the necessary tags, etc.
         2. PDF/A-1a provides document structure called tags, which provide an underlying structure for the content within the document and facilitate searching, repurposing

of content, and accessibility for people with disabilities such as blindness. This higher-level specification applies to "digitally born" documents captured directly from applications like Microsoft Word, which create document structure during the authoring process.

    iii. **PDF/A-1b** (basic). When the filer is unable to create a digital document via a document authoring tool, the document may be scanned and the filer should use the PDF/A-1b format.

        1. Documents that are created using the PDF/A-1b format and made searchable through OCR technology may not produce 100% text recognition due to the imperfect nature of the OCR process.

        2. PDF/A-1b provides a minimal compliance to ensure that the rendered visual appearance of a conforming document can be preserved and reproduced over the long-term.

        3. While PDF/A-1b ensures that the document text and additional content can be displayed properly, it does not guarantee the document screen reader content to be 100% correct or ADA compliant. This is because the OCR text recognition process is not 100% perfect and does not always supply the tags and document structure that screen readers rely on.

    iv. **Hybrid Documents.** Hybrid documents are documents that are composed of both PDF/A-1a and PDF/A-1b components. In essence, these are documents that were created as PDF/A-1a from a document authoring tool and have been edited by inserting a second document that has been rendered as PDF/A-1b. Hybrid documents are discouraged. Florida Courts are working to facilitate the means by which all documents can be rendered as PDF/A-1a directly from a document authoring tool.

d. **PDF/A Document Intelligence Elements Supported**

    i. **Bookmarks:** Bookmarks are links referring to specific pages of a PDF document. Bookmarks usually indicate the document structure, but may also include sensitive words and phrases.

    ii. **Internal Links**: Internal Links lead to specific targets within a document. If you remove the internal link, the link text or object remains but the PDF no longer performs the action to take the user to the linked content.

    iii. Related section from *Standards for Electronic Access to the Courts:*

        3.1.16. Embedded Hyperlinks
        Hyperlinks embedded within a submission should refer only to information within the same document, or to external documents or information sources that are reasonably believed to be trustworthy and stable over long periods of time. Hyperlinks should not be used to refer to external documents or information sources likely to change.

e. **PDF/A Document Intelligence Elements Prohibited**

    i. **Comments:** Comments usually refer to text markups and show the opinions of the author or reviewers. Comments include notes, markups, drawing objects as well as file and sound attachments.

    ii. **Attachments:** Document-level attachments are files (both PDF and other formats) added to a PDF document as reference or related data. *Embedded files are prohibited under PDF/A 1a and 1b.*

    iii. **Deleted Items:** Deleted Items refer to pages, text, images, and cropped items that are deleted using the PDF editor deletion utility. These items are not always actually removed from the PDF; they are just hidden. They could be restored by curious parties, restoring content that is not part of the official document.

    iv. **Destinations:** Destinations refers to an icon in a PDF file associated with a link to an external file. When the icon is clicked, the file is opened.

v. **Embedded External Links:** *Embedded external links are not permitted in PDF/A.* This does not prohibit the inclusion of unlinked URLs within a document, e.g., a law firm's web address in the letterhead.
vi. **Embedded Audio and Video:** *Embedded audio and video are not permitted in PDF/A.*
vii. **Annotations:** Annotations can be easily hidden and moved to change the meaning of the document.
viii. **Embedded thumbnails:** Thumbnails are miniature PDF pages often embedded in a PDF document.
ix. **Form fields and actions:** Form fields in a PDF document are dynamic controls that contain interactive features. If this item is removed, all form fields are flattened and cannot be edited, filled in, or signed any more. *Form fields and actions are prohibited under PDF/A.*
x. **JavaScript:** JavaScripts are active components that perform commands or actions under specified conditions. *The use of JavaScript is prohibited under PDF/A.*
xi. **Embedded non-display data:** This includes internal objects not considered to be metadata which are viewable with PDF tools.
xii. **Digital Signature:** A digital signature field that is generated by a PDF editor can be removed like any other form field by the person who signed it or otherwise by someone that has access to the private key**.** *Digital signatures become invalidated by time stamping the filing such as when a document is received by the E-Portal.*

f. **Items that Require Education and Development of Best Practices to be Included in the Standards**
   i. **Redaction:** Redacting a PDF document does not redact the content inside an already embedded index. To maintain document security, you should choose to remove embedded non-display data to delete the index before redacting. After redacting, you can re-index the document and the removed texts will not enter the new index. *Rule 2.425 requires the filer to minimize the sensitive data included in a filing. PDF redaction is more complex than TIFF redaction and the information is likely vulnerable when performed by the average filer.*
   ii. **Document sanitizing.** Document sanitizing involves the removal of certain metadata, comments, markup, etc. PDF editors will allow the user to perform automatic removal of selected metadata and other information using this feature.
   iii. **Electronic signatures.** Filers are encouraged to use electronic signatures as specified in the *Standards for Electronic Access to the Courts*. This includes the use of /s signatures, images of signatures, or signatures that are automatically added to a document using a signature pad. A digital signature created by a PDF editor should not be used (item 2.e.xii above) in documents to be filed with the E-Portal.
   iv. **PDF Metadata**
      1. **Metadata in Document:** Metadata includes information about the PDF document such as author, subject, keywords, date, copyright information, etc. Metadata can be searched using PDF search utilities.
      2. **Introduced Metadata:** Additional metadata may be added to a document as it flows through the judicial process. For example, Time Stamps, Case Number stamps, seals, etc. added after filing are metadata added by the Clerk.
      3. **PDF/A** requires some metadata to function.
      4. The submitter is responsible for document metadata. Training is needed on how the submitter can remove unwanted metadata.