Riksarkivet

PREFORMA

Author	Department	Phone	Date	Version	Page
Benjamin Yousefi	OFI	+46 (0) 10 476 72 98	2017/06/04	0.x	1 (6)

TESTFILE	Note
correct	Version: 5.2.7.2 Build ID: 5.2.7-1 CPU Threads: 8; OS Version: Linux 4.10; UI Render: default; VCL: gtk3; Locale: sv-SE (en_US.UTF-8); Calc: single -- Adobe Acrobat X 10.1.16 (Win

1. Description

This file attempts to represent a correct implementation of a PDF/A document *within the context of the TCs (Technical Class) defined in PREFORMA*.

The following method is applied.
1. All correct PDF/A must not contain any .
2. All correct PDF/A can contain an inverted TC or cannot contain a specific TC.
3. All correct PDF/A must contain specific inverted TC.

2. Fundamental error

A fundamental error is an error that is not allowed in any version of PDF/A and therefore shall not exist in any "correct" PDF/A. The qualifier for these errors are of two types: passive and active.

Passive qualifiers refer to those errors which are not usually included in the PDF/A unless specifically applied, such as, annotations, encodings, actions. Active qualifiers refer to those errors that could be included in a PDF/A during creation or conversion, for example, failure to create metadata, define color space, or embed font. Passive qualifiers can be avoided by simply not applying them, while active qualifiers must be checked to ensure that they have not occurred. The distinction is practical as, for example, a poor PDF/A creator could apply inappropriate functionality or use inappropriate settings.

Passive qualifiers		Active qualifiers	
TC003	Annotation Sound	TC038	File header - Has %PDF M.n in the beginning of the document file format.
TC004	Annotation Movie	TC039	File header - Has %PDF 1.0-1.7 in the beginning of the document file format.
TC005	Annotation Screen	TC040	Keyword spacings - Has spacings around keywords 'obj', 'endobj', 'stream', 'endstream', 'xref'.
TC006	Annotation 3D	TC041	Does not have: Color space
TC007	Encoding LZW	TC042	Device dependent - Does not have a Device dependent color space.

Author	Department	Phone	Date	Version	Page
Benjamin Yousefi	OFI	+46 (0) 10 476 72 98	2017/06/04	0.x	2 (6)
TESTFILE *correct*	Note Version: 5.2.7.2 Build ID: 5.2.7-1 CPU Threads: 8; OS Version: Linux 4.10; UI Render: default; VCL: gtk3; Locale: sv-SE (en_US.UTF-8); Calc: single -- Adobe Acrobat X 10.1.16 (Win				

Passive qualifiers		Active qualifiers	
TC008	Encoding Crypt	TC047	Metadata (DocumentInfo) – not in sync with the XMP metadata.
TC009	Encoding Crypt	TC048	Metadata (XMP) - key is missing.
TC010	Image Encoding Interpolation	TC049	Undefined (in the schema) custom XMP properties
TC011	Image Alternative/Proxy	TC050	XMP PDF/A version identifier missing
TC013	Transitions	TC051	<pdfaid:part><1-3> incorrect format
TC015	Import/Link to External Resource	TC052	<pdfaid:conformance><BUA> incorrect format
TC018	Executable PostScript	TC055	Embed fonts - Does not have fonts embedded
TC019	Form Action	TC056	Embed fonts Text rendering mode 3 - Does not have fonts not rendered embedded.
TC020	XFA Forms	TC058	Character identification of font subsets - Does not have Keys CharSet and CIDSet.
TC021	Action Launch	TC060	Font metric missing
TC022	Action Sound	TC062	Data after EOF
TC023	Action Movie		
TC024	Action Hide		
TC025	Action ResetForm		
TC026	Action ImportData		
TC027	Action Javascript		
TC028	Action Set-state		
TC029	Action No-op		
TC030	Action SetOCGState		
TC031	Action Rendition (play media)		
TC032	Action Trans(action)		

Author	Department	Phone	Date	Version	Page
Benjamin Yousefi	OFI	+46 (0) 10 476 72 98	2017/06/04	0.x	3 (6)

TESTFILE	Note
correct	Version: 5.2.7.2 Build ID: 5.2.7-1 CPU Threads: 8; OS Version: Linux 4.10; UI Render: default; VCL: gtk3; Locale: sv-SE (en_US.UTF-8); Calc: single --- Adobe Acrobat X 10.1.16 (Win

Passive qualifiers		Active qualifiers
TC033	Action GoTo3DView	
TC034	Named Action - Has Named Action besides NextPage, PrevPage, FirstPage, LastPage.	
TC035	Encryption	
TC036	Permission	

2.1. Conclusion

Create 2 files(1 training, 1 evaluation): a PDF used as base for subsequent test files.

• Can be created in LibreOffice and MSO2003+AAXP.

• Confirm that active qualifiers are not applied.

3. Conformance version

Functionality that is allowed in a PDF/A *conformance version* is therefore correct for that and only that version of PDF/A.

TC	Definition	PDF/A		
		1	*2*	*3*
TC001	Annotation FileAttachment	✖	✖	✔
TC002	Annotation FileAttachment non-PDF/A	✖	✔	✔
TC012	Document Optional content	✖	✔	✔
TC014	Transparency	✖	✔	✔
TC016	Document Attachment	✖	✖	✔
TC017	Document non-PDF/A Attachment	✖	✔	✔

TC037	.notdef	✔	✖	✖
TC053	Embed composite fonts CMaps - defined mappings from Unicode encodings to character collections	✖	✔	✔
TC054	Embed non-predefined CMaps - defined mappings from Unicode encodings to character collections	✔	✖	✖
~~TC061~~	~~Character encoding symbolic TrueType fonts - Has an Encoding entry in the font dictionary for any Symbolic TrueType fonts, or the #cmap# table in the embedded font program neither contains exactly one encoding nor contains at least the Microsoft Symbol (3,0 # Platform ID=3, Encoding ID=0) encoding.~~	✔	✖	✖

3.1. Summary

1. A correct PDF/A-1 may contain TC 37, 54, ~~61~~, while these TCs are errors in PDF/A-2 and -3.

2. A correct PDF/A-2 -3 may contain TC 2, 12, 14, 17, 53, while these TCs are errors in PDF/A-1.

3. A correct PDF/A-3 may contain TC 1, 2, 12, 14, 16, 17, 53, while these TCs are errors in PDF/A-1 - 2.

3.2. Conclusion

Create at least 4 files(2 training, 2 evaluation):

- Create 1 with classes 37, 54 and ~~61~~ (PDF/A-1).

  ○ Can be created in ~~LibreOffice~~ and MSO2003+AAXP.

  ○ (37) ⬚   Cambria (.notdef *and* can be mapped to Unicode)

  ○ (54) 海砂利水魚の、長久命の長助。五劫の擦り切れ

- Create 1 with classes 2, 12, 14, 17, 53 (PDF/A-2).

  ○ Can be created in AAXP.

- If possible, create 1 with classes 1, 2, 12, 14, 16, 17, 53 (PDF/A-3).

  ○ Can be created in AAXP.

4. Conformance level

Functionality that is required in a PDF/A *conformance level* is therefore correct for that PDF/A conformance level.

4.1. u

TC059	Unicode Character map - [Does not have] characters mapped to Unicode table

4.2. a

TC043	Logical structure - [Does not have] structured content
TC044	RoleMap - Has custom tags [but without mapping them to the] (mapped to) standard tags.
TC045	Hierarchy - [Does not have the] logical structure of the document
TC046	Structure type - [Does not have] (has) structure type for every structure element.

4.3. Summary

1. A correct PDF/A-2u -3u contains 59.

2. A correct PDF/A-1a -2a -3a contains 43, 44, 45, 46, 59.

4.4. Conclusion

Create at least 6 files(3 training, 3 evaluation):

- Create 1 with classes 59 (PDF/A -2u).

  ○ Can be created in AAXP.

Author	Department	Phone	Date	Version	Page
Benjamin Yousefi	OFI	+46 (0) 10 476 72 98	2017/06/04	0.x	6 (6)
TESTFILE *correct*	Note Version: 5.2.7.2 Build ID: 5.2.7-1 CPU Threads: 8; OS Version: Linux 4.10; UI Render: default; VCL: gtk3; Locale: sv-SE (en_US.UTF-8); Calc: single -- Adobe Acrobat X 10.1.16 (Win				

- Create 2 with classes 43, 44, 45, 46, 59 (PDFA -1a -2a).

  - Can be created in LibreOffice+AAXP and MSO2003+AAXP.

If possible, one of each for PDF/A-3 (PDF/A -3u and -3a).