PDF/A1-b compliance

(mPDF >= 4.3)

PDF/A1-b is a file format for the long-term archiving of electronic documents. It is based on the PDF Reference Version 1.4 from Adobe Systems Inc. and is defined by ISO 19005-1:2005

A key element to this reproducibility is the requirement for PDF/A documents to be 100% self-contained. All of the information necessary for displaying the document in the same manner every time is embedded in the file. This includes, but is not limited to, all content (text, raster images and vector graphics), fonts, and color information. A PDF/A document is not permitted to be reliant on information from external sources (e.g. font programs and hyperlinks).

Other key elements to PDF/A compatibility include:

- All fonts must be embedded and also must be legally embeddable for unlimited, universal rendering. This also applies to the Adobe standard fonts such as Times, Courier or Helvetica.
- Colorspaces specified in a device-independent manner.
- Encryption is not allowed.
- Use of standards-based metadata is mandated.
- PDF Version must be specified as 1.4

The standard specifies two levels of compliance for PDF files:

- PDF/A-1a - Level A compliance in Part 1
- PDF/A-1b - Level B compliance in Part 1

PDF/A-1b has the objective of ensuring reliable reproduction of the visual appearance of the document. PDF/A-1a includes all the requirements of PDF/A-1b and additionally requires that document structure be included (also known as being "tagged"), with the objective of ensuring that document content can be searched and repurposed.

A PDF/A document can be identified as such through PDF/A-specific metadata located in the "http://www.aiim.org/pdfa/ns/id/" namespace. However, claiming to be PDF/A and being so are not necessarily the same.

> **Important**: mPDF is **_not guaranteed_** to produce fully PDF/A1-b compliant files in all circumstances. It is the users responsibility to check compliance if this is essential.

mPDF and PDF/A1-b compliance

You can make mPDF produce mPDF/A1-b compliant files by setting:

```
$this->PDFA = true; // Edit in config.php or set at runtime
```

Some changes were made in mPDF 4.3 which affect all files (PDF/A or not) to improve compliance with PDF-1.4 specification:

- file version changed to PDF-1.4*
- newline removed at end of file after %%EOF
- an /ID object is added to the PDF trailer (this is optional in PDF-1.4. spec.) when the file is not encrypted
- a binary file marker (consisting of a comment line with 4 characters > 127 ASCII) is added just after the first line (recommended n PDF-1.4 spec.)

The PDF file version was changed to 1.5 in previous versions when 'active forms were introduced as an experiment'. These are no longer viable, and the rest of mPDF generated files meet 1.4 specification.

Colorspaces and ICC Profiles
PDF files handle colours internally in a number of different colorspaces. PDFA files must have one defined (and embedded) colour profile for one colorspace.

The 3 colour RGB colorspace is used most commonly in PDF/A documents, although you could select a 4-colour CMYK colorspace - but you cannot use both together, and SPOT colours are not permitted. RGB colorpsace is the default in mPDF for a PDF/A1-b document.

It is likely that most content in a PDF file will be RGB unless you have specifically defined otherwise e.g. a CMYK JPG image file, or colours defined as CMYK() etc.

Colorspaces can be altered - see restrictColorSpace.

An ICC Color profile must be embedded in the file.

The ICC Color profile will likely change the appearance of colours in your document. The ICC profile that is included with mPDF is an open licence color profile from the International Color Consortium (http://www.color.org) which seems to change colours a little as possible for an sRGB profile. Other sources of free ICC profiles are:

- http://www.eci.org
- http://www.color.org/profiles2.xalter
- http://www.adobe.com/support/downloads/iccprofiles/iccprofiles_win.html

You can change the color profile used by mPDF by adding a `*.icc` file to the `[mpdf]/iccprofiles/` folder. Then edit `config.php` e.g.

```
$this->ICCProfile = 'eciRGB_v2';
```

Fonts
All fonts must be embedded in a PDF/A file. mPDF normally uses the core Adobe fonts (Helvetica, Times, Courier, Symbol and Zapfdingbats) in a number of situations, and these font files are not available to embed. This affects:

- *$useCoreFontsOnly* is not permitted
- lists `<ul>` cannot use Zapfdingbats font for bullets
- form element `<select>` cannot use Zapfdingbats character as the drop-down symbol
- ISBN/ISSN barcode cannot use Helvetica for the ISBN number (above the bar)

Also, the fonts GB, BIG5, UHC and SJIS which reference the free CJK font pack from Adobe cannot be used

Transparency and Annotations
Because PDF/A-1b has the objective of ensuring reliable reproduction of the visual appearance of the document, no objects may be transparent or semi-transparent. This affects:

- watermarks text or images
- images with the CSS property *opacity* set
- annotation markers

Making it happen!

When `$this->PDFA` is set to true, the following changes will automatically and silently happen:

- list bullets will use the characters from the current font (if available) e.g. • ? ?
- the default font (which in some circumstances is set to a core Adobe font) is set to an alternative embedded font
- annotations (including hyperlinks) have their 'flag' changed to /F 28 which should force printing (this seems to make no difference but is required for PDFA compliance)
- a /Charset entry is included for each font when using embedded subsets
- a Metadata object is added to the file
- an ICC sRGB Color Profile is added as /OuputIntent

Only the first change may make a discernible change to the displayed document. All other changes are to the file structure.

Warnings
Some problems can be fixed by mPDF, but will cause a change to the appearance of your document. By default they will generate warning messages. Once you have assessed the warnings, you can direct mPDF to make these changes automatically by setting in your script:

```
$mpdf->PDFAauto = true; // Overrides warnings making changes when possible to force PDFA1-b
compliance
```

The following issues will cause a warning message when you try to generate a PDFA file:

Problem detected	Action taken by mPDF
List bullets cannot be substituted from current font	Bullets substituted by a hypen "-"

Problem detected	Action taken by mPDF
Character substitution (*$useSubstitutions* or *$useSubstitutionsMB*) is enabled. This would attempt to substitute missing characters using core Adobe fonts which are disallowed.	Character substitution is disabled
Annotation markers are present; these cannot be semi-transparent and may hide underlying text.	Annotation markers are placed in the right margin
Images which have CSS property opacity set less than 1 - PDFA does not allow transparency. NB GIF or PNG images with a simple transparency set (not alpha-channel mask) are valid.	Opacity is changed to 1 (no transparency)
JPG images in CMYK colour are not permitted.	Converted to RGB (only if GD library installed). NB This may significantly change the colour and appearance of the image
Core fonts (Times, Helvetica, Courier) are specified directly e.g. in CSS stylesheet	Font will be substituted with an available sans, serif, or mono font.
Barcode of ISBN, ISSN or EAN-13 type which specifies the code to appear above the barcode (normally uses Helvetica font).	Substitutes available sans-serif font
Form element `<select>` which normally uses a Zapfdingbats character for the drop-down symbol.	An equal sign "=" will be substituted if ▼ is not available in the default sans-serif font.

Fatal Errors

Some issues cannot be fixed automatically by mPDF and will generate fatal errors:

- *$useCoreFontsOnly* is set as **TRUE** (cannot embed core fonts)
- BIG5, SJIS, UHC or GB fonts cannot be used (cannot be embedded)
- Watermarks - text or image - are not permitted (transparency is disallowed so will make text unreadable)
- Using CMYK colour in functions SetTextColor() SetDrawColor() or SetFillColor()
- PNG images with alpha channel transparency ('masks' not allowed)
- encryption is enabled

See Also

- PDF/X-1a compliance
- PDFA - Create PDF/A1-b compliant document
- PDFX - Create PDF/X-1a compliant document
- PDFAauto - Specify whether to automatically fix issues to create PDF/A1-b compliant document
- PDFXauto - Specify whether to automatically fix issues to create PDF/X-1a compliant document
- ICCProfile - Specify the ICC profile for the chosen colorspace used in the document
- restrictColorSpace - Specify whether to automatically limit the colorspaces used

Useful resources

On-line PDFA Validator: http://www.validatepdfa.com/online.htm

Useful info on PDFA: http://www.pdfa.org/doku.php?id=pdfa:en:techdoc

Printed on Wed 05 Aug 2015 12:10:53 GMT +0100 (DST)