

veraPDF

nestor-Workshop Formaterkennung
Kiel, 29.06.2017
Martin Hoppenheit



Was ist veraPDF?

Software zur Validierung von PDF/A-Dateien

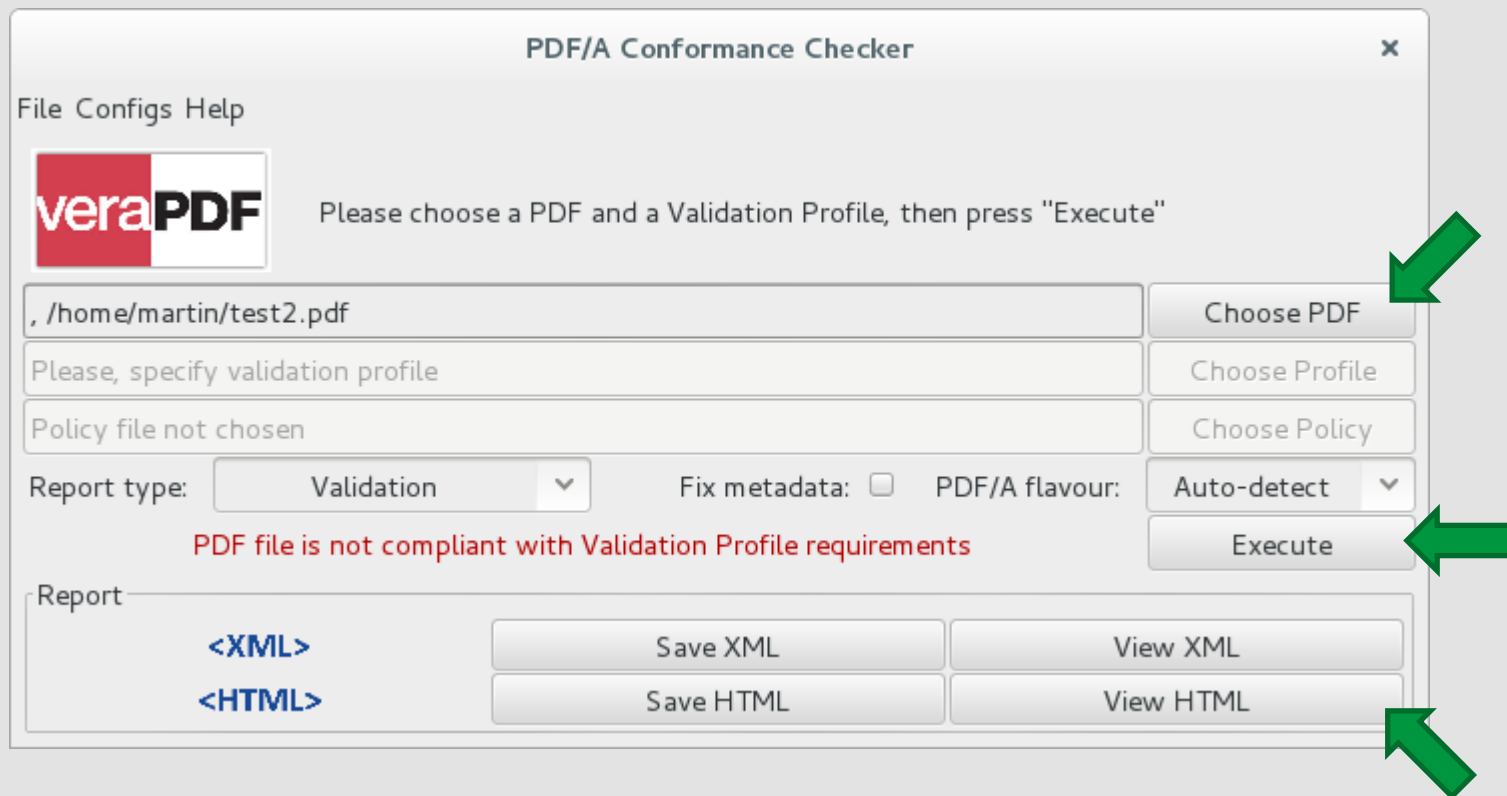
- Validierung bzgl. aller Varianten des PDF/A-Standards
- Keine vollständige Validierung von PDF allgemein (ohne /A)
- Reparatur von PDF-Metadaten (Metadata Fixer)
- Überprüfung bzgl. eigener Regeln (Policy Checker)
- Open Preservation Foundation + PDF Association
- Projektstart 2014, Version 1.0 Anfang 2017, aktuell Version 1.6
- Open Source (GPL und Mozilla Public License)
- GUI, CLI, Java-Library

Installation

- **Voraussetzung: Java**
- **Installer auf dem USB-Stick (Software-Repository/veraPDF) oder von <http://verapdf.org/software/>**
- **Entpacken der zip-Datei**
- **Ausführen der Datei verapdf-install.bat (Windows) bzw. verapdf-install (Linux)**

Graphische Oberfläche (GUI)

Aufruf: verapdf-gui



Graphische Oberfläche (GUI)

HTML-Report

Validation information

Rule	Status
Specification: ISO 19005-1:2005, Clause: 6.3.8, Test number: 1	
The font dictionary shall include a ToUnicode entry whose value is a CMap stream object that maps character codes to Unicode values, as described in PDF Reference 5.9, unless the font meets any of the following three conditions: (*) fonts that use the predefined encodings MacRomanEncoding, MacExpertEncoding or WinAnsiEncoding, or that use the predefined Identity-H or Identity-V CMaps; (*) Type 1 fonts whose character names are taken from the Adobe standard Latin character set or the set of named characters in the Symbol font, as defined in PDF Reference Appendix D; (*) Type 0 fonts whose descendant CIDFont uses the Adobe-GB1, Adobe-CNS1, Adobe-Japan1 or Adobe-Korea1 character collections.	Failed
1 occurrences	
Glyph	
toUnicode != null	
root/document[0]/pages[1](1 0 obj PDPPage)/contentStream[0](2 0 obj PDContentStream)/operators[2 90]/usedGlyphs[0](ABCDEE+Arial 177 0)	

Kommandozeile (CLI)

Dokumentation anzeigen:

```
verapdf --help
```

Dokument validieren:

```
verapdf dokument.pdf
```

Kommandozeile (CLI)

PDF/A-Variante vorgeben:

```
verapdf --flavour 2b dokument.pdf
```

Text statt XML als Ausgabeformat, kurz und knapp:

```
verapdf --format text -f 2b dokument.pdf
```

Alle *.pdf-Dateien in einem Verzeichnisbaum validieren:

```
verapdf --recurse verzeichnis
```

Weiterführende Links

- Projektseite: <http://verapdf.org>
- GitHub: <https://github.com/verapdf>
- Dokumentation: <http://docs.verapdf.org/>

Vielen Dank für Ihr Interesse!

martin.hoppenheit@lav.nrw.de

