

**KOST** Koordinationsstelle für die dauerhafte Archivierung  
elektronischer Unterlagen

---

Ein Gemeinschaftsunternehmen von Schweizer Archiven

## Formatvalidierung: KOST-Val

nestor Workshop: "Formaterkennung,  
Formatvalidierung und Tools"  
ZBW Kiel,  
29.06.2017

Claire Röthlisberger-Jourdan

# Inhaltsverzeichnis

- Was ist KOST-Val
- Funktionsprinzip / Drittvalidatoren
- Eingabe
- Ausgabe
- Anwendung

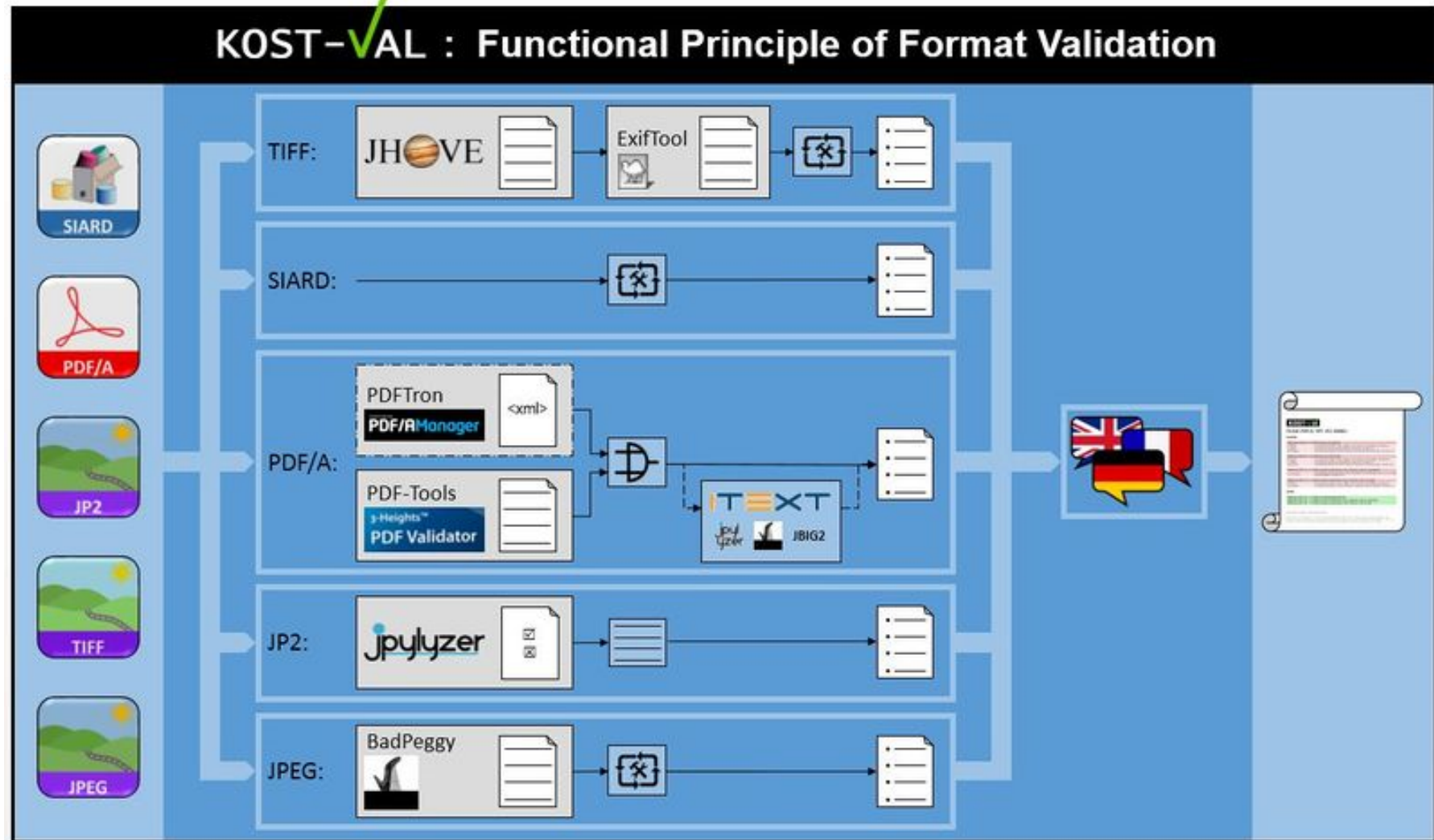
# Was ist KOST-Val



KOST-Val

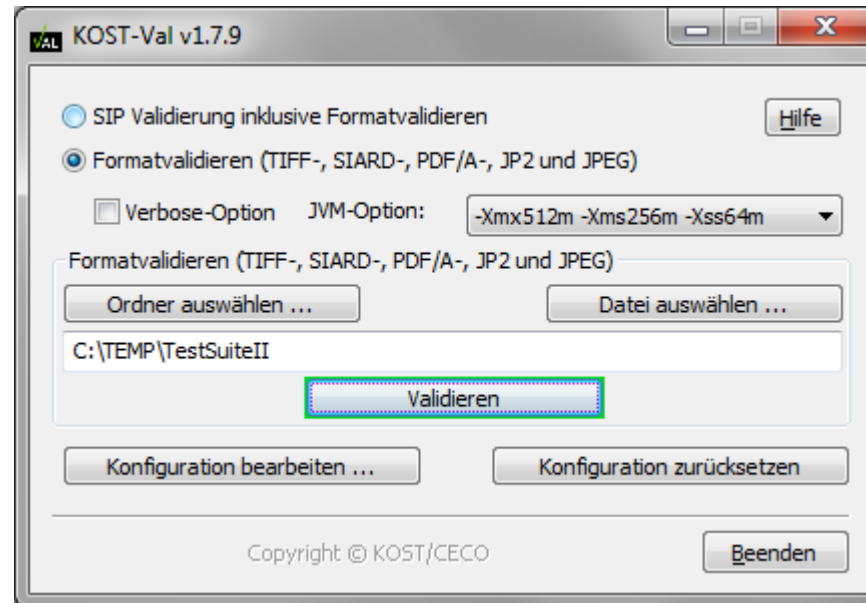
- Erstmals veröffentlicht: 2012
- Support & Entwicklung: KOST
- Lizenz: GPL 3+
- Unterstützte Formate: PDF/A, SIARD, TIFF, JP2, JPEG und SIP (eCH-0160)
- Nutzungsmöglichkeiten: GUI & CMD
- Ausgabe: XML inkl. XSL

# Funktionsprinzip / Drittvalidatoren



# Eingabe

- GUI



- CMD

vereinfachter Aufruf:

```
kostval --format <Ordner / Datei>
```

Aufruf:

```
resources\jre6\bin\java.exe -jar kostval_de.jar --format C:\TEMP\TestSuiteII
```

# Ausgabe

## • XML

```
<?xml version="1.0" encoding="ISO-8859-1"?><?xml-stylesheet type="text/xsl" href="kost-val.xsl"?><KOSTValLog>
<Infos><Start>19.06.2017 17:36:03</Start>
<End>19.06.2017 17:36:08</End>
<FormatValOn>PDF/A, TIFF, JP2, SIARD, JPEG</FormatValOn>
<Info>KOST-Val v1.7.9, Copyright (C) 2012-2017 Claire Roethlisberger (KOST-CECO), Christian Eugster, Olivier
Debenath, Peter Schneider (Staatsarchiv Aargau), Markus Hahn (coderslagoon), Daniel Ludin (BEDAG AG). This
program comes with ABSOLUTELY NO WARRANTY. This is free software, and you are welcome to redistribute it
under certain conditions; see GPL-3.0_COPYING.txt for details</Info></Info>
<Format><Info><Summary>Von den 11 Dateien sind 1 (9.091%) valid, 7 (63.636%) invalid und 3 (27.273%) wurden
nicht validiert.</Summary><Info><Message>Folgende Dateien wurden nicht validiert (.gif):
</Message></Info><Info><Message> - C:\TEMP\TestSuiteII\GIF\Table1.gif</Message></Info><Info><Message> -
C:\TEMP\TestSuiteII\GIF\header.gif</Message></Info><Info><Message> - C:\TEMP\TestSuiteII\GIF\map.gif
</Message></Info></Info>
<Validation>
<ValType>Validierung: PDFA</ValType>
<ValFile>C:\TEMP\TestSuiteII\PDFA\Auftragsuebersicht.pdf</ValFile>
<PdfaVL>-2B</PdfaVL>
<Error><Modul>C) Grafiken</Modul><Message>Das Dokument enthaelt geratespezifische Farbraeume. [iCategory_3]
</Message></Error>
<Error><Modul>D) Schriften</Modul><Message>Das Dokument enthaelt Schriften ohne eingebettete Font-Programme
oder Codierungsinformation (CMAPs). [iCategory_8]</Message></Error>
<Error><Modul>H) Metadaten</Modul><Message>Die Dokumentmetadaten sind nicht vorhanden, inkonsistent oder
beschadigt. [iCategory_16]</Message></Error>
<Invalid>invalid</Invalid>
</Validation>
<Validation>
<ValType>Validierung: PDFA</ValType>
<ValFile>C:\TEMP\TestSuiteII\PDFA\gpl2.pdf</ValFile>
<PdfaVL>-1B</PdfaVL>
<Error><Modul>E) Transparenz</Modul><Message>Das Dokument enthaelt Transparenz. [iCategory_10]
</Message></Error>
<Invalid>invalid</Invalid>
</Validation>
<Validation>
<ValType>Validierung: PDFA</ValType>
<ValFile>C:\TEMP\TestSuiteII\PDFA\isartor-FG.pdf</ValFile>
<PdfaVL>-1B</PdfaVL>
<Error><Modul>F) Annotationen</Modul><Message>Das Dokument enthaelt unbekannte Annotation-Typen.
[iCategory_11]</Message></Error>
<Error><Modul>F) Annotationen</Modul><Message>Das Dokument enthaelt Multimedia-Annotationen (Ton, Filme).
[iCategory_12]</Message></Error>
<Error><Modul>G) Aktionen & Interaktionen</Modul><Message>Das Dokument enthaelt andere Action Types als
solche fuer die Navigation (Launch, JavaScript, Resetform, etc.). [iCategory_15]</Message></Error>
<Invalid>invalid</Invalid>
</Validation>
<Validation>
<ValType>Validierung: PDFA</ValType>
<ValFile>C:\TEMP\TestSuiteII\PDFA\isartor-J.pdf</ValFile>
<PdfaVL>-1B</PdfaVL>
<Error><Modul>F) Annotationen</Modul><Message>Das Dokument enthaelt Annotations oder Formularfelder mit
```

## XML inkl. XSL

### KOST-Val

#### Format ( PDF/A, TIFF, JP2, SIARD, JPEG ):

Von den 11 Dateien sind 1 (9.091%) valid, 7 (63.636%) invalid und 3 (27.273%) wurden nicht validiert.

Folgende Dateien wurden nicht validiert (.gif):

- C:\TEMP\TestSuiteII\GIF\Table1.gif

- C:\TEMP\TestSuiteII\GIF\header.gif

- C:\TEMP\TestSuiteII\GIF\map.gif

#### Invalid:

##### Validierung: PDFA-2B -> C:\TEMP\TestSuiteII\PDFA\Auftragsuebersicht.pdf

C) Grafiken	Das Dokument enthaelt geratespezifische Farbraeume. [iCategory_3]
D) Schriften	Das Dokument enthaelt Schriften ohne eingebettete Font-Programme oder Codierungsinformation (CMAPs). [iCategory_8]
H) Metadaten	Die Dokumentmetadaten sind nicht vorhanden, inkonsistent oder beschadigt. [iCategory_16]

##### Validierung: PDFA-1B -> C:\TEMP\TestSuiteII\PDFA\gpl2.pdf

E) Transparenz Das Dokument enthaelt Transparenz. [iCategory\_10]

##### Validierung: PDFA-1B -> C:\TEMP\TestSuiteII\PDFA\isartor-FG.pdf

F) Annotationen	Das Dokument enthaelt unbekannte Annotation-Typen. [iCategory_11]
F) Annotationen	Das Dokument enthaelt Multimedia-Annotationen (Ton, Filme). [iCategory_12]
G) Aktionen & Interaktionen	Das Dokument enthaelt andere Action Types als solche fuer die Navigation (Launch, JavaScript, Resetform, etc.). [iCategory_15]

##### Validierung: PDFA-1B -> C:\TEMP\TestSuiteII\PDFA\isartor-J.pdf

F) Annotationen	Das Dokument enthaelt Annotations oder Formularfelder mit mehrdeutigen oder ohne notwendige Erscheinungsbilder. [iCategory_14]
G) Aktionen & Interaktionen	Das Dokument enthaelt andere Action Types als solche fuer die Navigation (Launch, JavaScript, Resetform, etc.). [iCategory_15]

##### Validierung: TIFF -> C:\TEMP\TestSuiteII\TIFF\551ad8c6c639cc59940903b0428f47.tif

B) Jhove	Die Jhove-Validierung wurde nicht bestanden ( Status: Well-Formed, but not valid).
B) Jhove	Jhove ErrorMessage: ProjectedCSType value out of range: 2180
C) Komprimierung	Compression: Adobe Deflate ist nicht zulaessig.
G) Kacheln	TileWidth: 256 ist nicht zulaessig.
G) Kacheln	TileLength: 256 ist nicht zulaessig.

##### Validierung: TIFF -> C:\TEMP\TestSuiteII\TIFF\6453732434a8a2358a3c895d962bdc2.tif

G) Kacheln	TileWidth: 128 ist nicht zulaessig.
G) Kacheln	TileLength: 128 ist nicht zulaessig.

##### Validierung: TIFF -> C:\TEMP\TestSuiteII\TIFF\9b286add20871bbbf1601997b429344.tif

B) Jhove	Die Jhove-Validierung wurde nicht bestanden ( Status: Well-Formed, but not valid).
B) Jhove	Jhove ErrorMessage: JPEGLossy not defined for JPEG compression
B) Jhove	Jhove ErrorMessage: For PhotometricInterpretation, SamplesPerPixel must be >= 3, equals1
C) Komprimierung	Compression: JPEGLossy (old-style) ist nicht zulaessig.
C) Komprimierung	Compression: JPEG ist nicht zulaessig.
D) Farbraum	PhotometricInterpretation: YCbCr ist nicht zulaessig.

#### Valid:

##### Validierung: TIFF -> C:\TEMP\TestSuiteII\TIFF\bd8582b004aa2560f5bccc4f3d6.tif

19.06.2017 17:36:03 - 19.06.2017 17:36:08

KOST-Val v1.7.9, Copyright (C) 2012-2017 Claire Roethlisberger (KOST-CECO), Christian Eugster, Olivier Debenath, Peter Schneider (Staatsarchiv Aargau), Markus Hahn (coderslagoon), Daniel Ludin (BEDAG AG). This program comes with ABSOLUTELY NO WARRANTY. This is free software, and you are welcome to redistribute it under certain conditions; see GPL-3.0\_COPYING.txt for details

# Anwendung

- Anwendungsmöglichkeiten
  - Formatvalidierung diverser Formate gleichzeitig
  - SIP-Validierung inkl. Formatvalidierung
  - Einbindung in einen Prozess möglich
- Erfahrungswerte
  - Konfigurierbar je nach Archiv und Anforderung
  - Erweiterung nur mit guten Validatoren und bei Bedarf
- Stärken und Schwächen (keine Eigenbewertung)

**KOST** Koordinationsstelle für die dauerhafte Archivierung  
elektronischer Unterlagen

---

Ein Gemeinschaftsunternehmen von Schweizer Archiven

## Claire Röthlisberger-Jourdan

Koordinationsstelle für die Archivierung elektronischer Unterlagen

KOST

c/o Schweizerisches Bundesarchiv

Archivstrasse 24

CH-3003 Bern

T +41 79 464 07 41

E [claire.roethlisberger@kost.admin.ch](mailto:claire.roethlisberger@kost.admin.ch)

W [www.kost-ceco.ch](http://www.kost-ceco.ch)