

Leitfaden

PDF/A-3 als Trägerformat für ZUGFeRD

Dieser Leitfaden entstand in Zusammenarbeit zwischen der Arbeitsgemeinschaft für wirtschaftliche Verwaltung e.V. (AWV) und der PDF Association für FeRD mit dem Ziel, zu erklären, welche Eigenschaften und Konfigurationen ein PDF/A-3 Dokument aufweisen muss, um konform zu dem ZUGFeRD Standardentwurf des Forum elektronische Rechnung Deutschland, FeRD zu sein.

Der ZUGFeRD-Standard beschreibt ein Dokumenten- und Datenformat für den Austausch von elektronischen Rechnungen. Die Grundlage bildet ein hybrider Ansatz, bei dem eine visuelle Darstellung der Rechnungsdaten über ein PDF/A-3 konformes Dokument erfolgt, während eine Rechnungsdatendatei im XML-Format in das PDF/A-3 Dokument eingebettet wird.

Das resultierende PDF/A-3 Dokument muss intern über folgende Konstrukte verfügen, um die Konformität zu gewährleisten:

- eine PDF/A-3 konforme Struktur, d.h. das Ausgangsdokument ist ohne die eingebetteten Daten bereits PDF/A-3 konform! Dabei spielt die Konformitätsstufe (d.h. 3a, 3b oder 3u) keine Rolle.
- die Einbettung der XML-Datei über eine Beziehung vom Typ "Alternative" mit Bezug auf das gesamte Dokument.
- die Präsenz eines spezifischen XMP Erweiterungsschemas zur Beschreibung des Dokuments als ZUGFeRD-konforme Rechnung sowie der entsprechenden XMP Metadaten

Im folgenden werden diese Prämissen detaillierter beschrieben.

1. PDF/A-3 konforme Struktur

Ein PDF/A-3 konformes Dokument muss die Anforderungen der Norm ISO 19005-3 [IS1901] erfüllen. Darin werden die grundlegenden Unterschiede und Beschränkungen einer A-3 Datei auf Basis des zugrundeliegenden Standards ISO 32000-1 [IS3201], auch bekannt als PDF 1.7, beschrieben. Im wesentlichen sind dies Vorgaben, die bereits in den Vorgängerstandards PDF/A-1 [IS1901] und PDF/A-2 [IS1902] dargestellt sind. Die wichtigsten Besonderheiten einer PDF/A Datei im Vergleich zu einem beliebigen PDF Dokument sind dabei

- Es muss eine Kennung in Form eines XMP Erweiterungsschemas existieren, das die PDF/A-Eigenschaft und den Konformitätsstufe explizit enthält
- Alle Metadaten sind in XMP Form einzubetten. Das verwendete XMP Schema kann entweder aus der Menge vordefinierter Schemata genommen werden oder es muss ein eigenes Schema erstellt und zwingend immer mit den Metadaten zusammen eingebettet werden.
- Alle verwendeten Zeichensätze sind in das PDF/A einzubetten. Zur Optimierung können an Stelle vollständiger Zeichensätze auch nur Subsets der effektiv verwendeten Glyphen eingebettet werden.
- Es dürfen keine Fremddateien wie Filme, Tondateien oder sonstige Binärdateien eingebettet werden, außer über den später beschriebenen A-3 konformen Mechanismus.
- Es dürfen keine aktiven Elemente mehr im PDF/A vorhanden sein. Darunter versteht man z.B. JavaScript für Aktionen oder Flash für Animationen.
- Es sind nur genau definierte Bildformate zur Einbettung zulässig. Dazu zählen CCITT Group 3 und Group 4, JBIG2, JPEG und JPEG2000.
- Es darf keine Verschlüsselung oder sonstige Berechtigungssteuerung im Dokument enthalten sein.

2. Einbettung der XML-Datei

Die Einbettung der Rechnungsdaten im XML-Format erfolgt über ein sog. *File Specification Dictionary* ([IS32001], Kap. 7.11.3). Voraussetzung ist ein gültiger MIME-Type für das einzubettende Dokument. Im Fall von ZUGFeRD ist der MIME-Type immer `text/xml`.

Das *Stream Dictionary* der eingebetteten Datei sollte einen Schlüssel namens `Params` aufweisen. `Params` verweist auf ein Dictionary, das zumindest einen Eintrag `ModDate` besitzt, der das letzte Änderungsdatum der eingebetteten Datei enthält.

Das eingebettete Dokument ist auch in den Objektbaum `Names` aufzunehmen, um konformen PDF-Werkzeugen die Darstellung der Datei zusammen mit zusätzlichen Informationen zu ermöglichen.

Grundsätzlich ist die Einbettung von mehreren Dateien in das PDF/A-3 Dokument möglich. Damit können z.B. neben dem Rechnungsdatendokument auch Informationsdokumente zur Rechnungsprüfung im PDF/A-3 gebündelt werden. Um auf PDF-Ebene kenntlich zu machen, bei welcher der eingebetteten Dateien es sich um das Rechnungsdatendokument handelt, ist der Name des Rechnungsdatendokuments in das entsprechende Metadatenattribut aufzunehmen (s.h. Kap. 3).

2.1. Bezug der Einbettung

Eine eingebettete Datei kann sich prinzipiell auf das gesamte (PDF-)Dokument (Document Level), eine bestimmte Seite (Page Level) oder nur auf ein bestimmtes Objekt im PDF (DPart Level) beziehen. Abhängig von der Bezugsart befindet sich das *File Specification Dictionary* entweder im Document Dictionary, dem Page Dictionary oder dem DPart Dictionary. Die Verknüpfung erfolgt über ein Array namens `AF` (für Associated Files), das in die jeweiligen Dictionaries eingetragen wird und einen Verweis auf das *File Specification Dictionary* enthält.

In der Version 1.0 des ZUGFeRD Standards ist pro PDF/A-3 Dokument nur die Einbindung eines einzigen Rechnungsdatendokuments zulässig. Dementsprechend ist die Bezugsart "Document Level" zu wählen. Die

Einbettung weiterer Dokumente und Dateien, die keine Rechnungsdaten enthalten, ist davon unbetroffen¹.

2.2. Datenbeziehung

Neben der Bezugsart verlangt ISO 19005-3 die Angabe einer Datenbeziehung, d.h. in welchem Verhältnis das eingebettete Dokument zum PDF-Teil, d.h. der Visualisierung, steht. Diese Datenbeziehung wird durch das Tag `AFRelationship` ausgedrückt und kann folgende Werte annehmen:

- `Data` – Die eingebettete Datei enthält Daten, die für die visuelle Darstellung im PDF-Teil verwendet werden, z.B. für eine Tabelle oder einen Graphen.
- `Source` – Die eingebettete Datei enthält die Quelldaten für die daraus abgeleitete visuelle Darstellung im PDF-Teil, z.B. bei einer PDF-Datei, die durch eine XSL-Transformation aus einer (eingebetteten) XML-Quelldatei entsteht oder die MS Word-Datei, aus der das PDF erzeugt wurde.
- `Alternative` – Wenn die eingebetteten Daten eine alternative Darstellung des PDF-Inhalts sind, sollte diese Datenbeziehung verwendet werden.
- `Supplement` – Diese Datenbeziehung wird angewendet, wenn die eingebettete Datei weder als Quelle noch als alternative Darstellung dient, sondern die Datei zusätzliche Informationen z.B. zur einfacheren maschinellen Verarbeitung enthält.
- `Unspecified` – Sofern keine der vorstehenden Datenbeziehungen zutrifft oder eine unbekannte Datenbeziehung besteht, wird diese Datenbeziehung verwendet.

Im Fall von ZUGFeRD muss der Wert `Alternative` betragen. Dadurch soll verdeutlicht werden, dass der steuerrechtlich relevante Inhalt beider Darstellungen übereinstimmt und es sich bei der XML-Datei lediglich um eine andere bzw. alternative Darstellungsform handelt, die sich besser für die maschinelle Verarbeitung eignet.

Abb. 1 verdeutlicht diese Struktur am Beispiel einer ZUGFeRD XML-Rechnung. Die eingebettete Rechnungsdatei hat den Namen `ZUGFeRD_Example_1.xml`. Das Array `/AF` ist Bestandteil des Document Dictionaries (direkt unter Root), weshalb sich die Rechnungsdatei auf das ganze Dokument bezieht. Die

¹ In künftigen Versionen des ZUGFeRD Standards kann diese Beschränkung aufgehoben werden, d.h. es können dann auch mehrere Rechnungen pro PDF/A-3 Dokument zulässig sein.

Datenbeziehung ist Alternative, d.h. die XML-Rechnungsdaten sind eine alternative Form der Darstellung der PDF-Visualisierung.

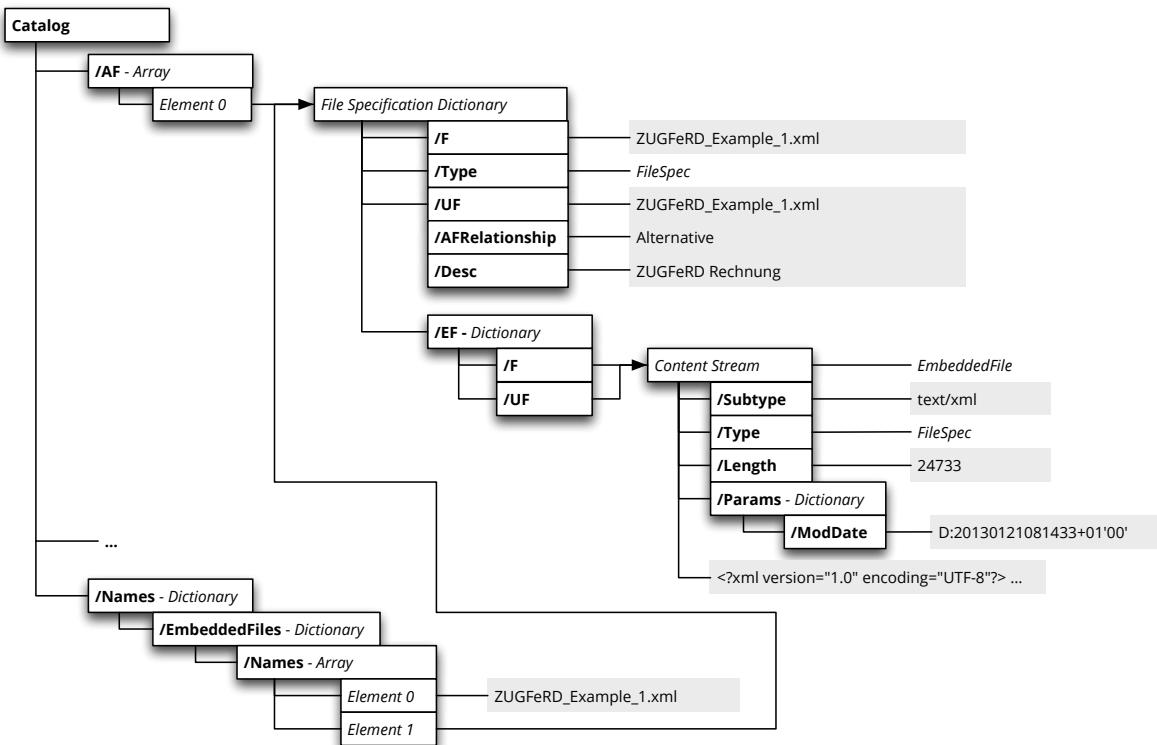


Abb.1 Die Objektstruktur im PDF zur Dateieinbettung (Associate File).

1. PDF/A Erweiterungsschema ZUGFeRD

Die PDF/A-konforme Aufnahme von Metadaten erfordert im Fall, dass die Metadatenattribute benutzerspezifisch sind (d.h. sie sind nicht in den im PDF/A-Standard deklarierten XMP-Schemata enthalten), die Definition eines eigenen Metadatenschemas. Diese Schemadefinition gehorcht den Konventionen für PDF/A Erweiterungsschemas [TN0008], [TN0009]. Sowohl die konkrete Metadatenausprägung als auch das Erweiterungsschema sind in jedem PDF/A Dokument einzubetten. Die Angabe der Referenz auf eine externe Ablage genügt nicht.

Die PDF Association hat für den Einsatz von ZUGFeRD-konformen Rechnungsdokumenten ein entsprechendes Erweiterungsschema definiert. Die Eigenschaften des Erweiterungsschemas sind im folgenden aufgeführt:

Eigenschaft	Wert	Beschreibung
Name des Erweiterungsschemas	ZUGFeRD Schema	
URI	http://www.awv-net.de/ns/zugferd/1/	Dieser Ablageort wird sich im Zuge der weiteren Standardisierung noch ändern!
Schema Präfix	zf	Präfix des Namensraums

Die Felder des Erweiterungsschemas zeigt die nachstehende Tabelle. Die XPath-Ausdrück beziehen sich auf das ZUGFeRD Rechnungsdatenschema [ZF0001].

Feld	Beschreibung	Beispiel
zf:DocumentID	Die Rechnungsnummer aus der XML-Datei (XPath: /HeaderExchangedDocument/ID)	471102

Feld	Beschreibung	Beispiel
zf:DocumentDate	Das Rechnungsdatum aus der XML-Datei (XPath: /HeaderExchangedDocument/ IssueDateTime)	2013-03-05
zf:DocumentType	Der Dokumententyp, enthält bei ZUGFeRD-Rechnungen immer INVOICE	INVOICE
zf:DocumentFileName	Der Dateiname des eingebetteten Rechnungsdatendokuments; muss identisch sein mit dem Wert des /F Eintrags im File Specification Dictionary.	ZUGFeRD_Example_1.xml
zf:Version	Die Version des XML-Schemas der Rechnungsdaten	0.11
zf:ConformanceLevel	Die Ausprägung der XML-Rechnungsdaten entsprechend den Vorgaben von ZUGFeRD (erlaubte Werte: BASIC, COMPLETE, EXTENDED)	BASIC

Das vollständige Erweiterungsschema befindet sich im Anhang.

1.1. Beispiel

Eine exemplarische Belegung (hier mit den Testdaten des CeBIT 2013 Piloten) illustriert die Verwendung innerhalb eines PDF/A Dokuments.

```

<rdf:Description rdf:about="" xmlns:zf="http://www.awv-net.de/ns/zugferd/1/">
  <zf:DocumentID>471102</zf:DocumentID>
  <zf:DocumentDate>2013-03-05</zf:DocumentDate>
  <zf:DocumentType>INVOICE</zf:DocumentType>
  <zf:DocumentFileName>ZUGFeRD_Example_1.xml</zf:DocumentFileName>
  <zf:Version>0.11</zf:Version>
  <zf:ConformanceLevel>BASIC</zf:ConformanceLevel>
</rdf:Description>

```

2. Referenzen

- [TN0008] TechNote 0008: Predefined XMP Properties in PDF/A-1, PDF/A Competence Center, www.pdfa.org/doku.php?id=pdfa:en:techdoc
- [TN0009] TechNote 0009: XMP Extension Schemas in PDF/A-1, PDF/A Competence Center, www.pdfa.org/doku.php?id=pdfa:en:techdoc
- [IS3201] ISO 32000-1, Document management — Portable document format — Part 1: PDF 1.7, www.iso.ch
- [IS1901] ISO 19005-1: Document management — Electronic document file format for long-term preservation — Part 1: Use of PDF 1.4 (PDF/A-1), www.iso.ch
- [IS1902] ISO 19005-2: Electronic document file format for long-term preservation — Part 2: Use of ISO 32000-1 (PDF/A-2), www.iso.ch
- [IS1903] ISO 19005-3: Document management — Electronic document file format for long-term preservation - Part 3: Use of ISO 32000-1 with support for embedded files (PDF/A-3), www.iso.ch
- [AD2004] XMP Specification, Januar 2004, Adobe Systems Incorporated. www.aiim.org/documents/standards/xmpspecification.pdf
- [ZF0001] ZUGFeRD_Invoice_CEBIT.xml, ZUGFeRD CeBIT Pilotpackage, FeRD, Februar 2013

Copyright und Verwendung

Copyright © 2013 AWV, www.awv-net.de

Referenzen und Web-Links zum Originalspeicherort dieses Dokuments sind erlaubt. Die Weiterverteilung oder Vervielfältigung des Dokuments bedarf der ausdrücklichen schriftlichen Erlaubnis des Rechteinhabers.

Kontaktieren Sie bei Fragen zum Inhalt oder zur Weiterverbreitung die Adresse zugferd@awv-net.de.

Dokumentenstatus

2013-01-28 Version 1.0

Autor

Dr. Bernd Wild, FeRD Kernteam und intarsys consulting GmbH

Anhang

PDF/A Erweiterungsschema ZUGFeRD Schema

```
<!-- Copyright 2013 PDF Association
www.pdfa.org
```

This XMP schema specifies XMP entries for inclusion in ZUGFeRD compliant electronic invoice documents based on PDF/A-3.

The following properties of the custom schema are used:

Schema name: ZUGFeRD Schema

Schema namespace URI: <http://www.awv-net.de/ns/zugferd/1>

Preferred schema namespace prefix: zf

Properties in this schema and their types:

DocumentID: Text Correlates to the invoice number

DocumentDate: Date Complete Date according to

<http://www.w3.org/TR/NOTE-datetime>,
e.g. 2008-01-17

DocumentType: Text INVOICE

DocumentFileName: Text Name of the embedded file;
must be equal with the
value of the /F tag in the /EF
structure

Version: Text The actual version of the
ZUGFeRD standard

ConformanceLevel: Text The actual conformance level of the
ZUGFeRD standard,
e.g. BASIC, COMPLETE

Since this example uses a schema beyond the set of predefined XMP 2004 schemas it includes a description of the custom schema according to the PDF/A requirements.

-->

```
<!-- Several XMP entries in the custom schema -->
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
<!-- XMP extension schema container schema for the fm schema -->
<rdf:Description rdf:about=""
```

```

xmlns:pdfaExtension="http://www.aiim.org/pdfa/ns/extension/"
xmlns:pdfaSchema="http://www.aiim.org/pdfa/ns/schema#"
xmlns:pdfaProperty="http://www.aiim.org/pdfa/ns/property#" >

    <!-- Container for all embedded extension schema descriptions -->
    <pdfaExtension:schemas>
        <rdf:Bag>
            <rdf:li rdf:parseType="Resource">
                <!-- Optional description of schema -->
                <pdfaSchema:schema>ZUGFeRD Schema</pdfaSchema:schema>
                <!-- Schema namespace URI -->
                <pdfaSchema:namespaceURI>http://www.awv-net.de/ns/zugferd/1/</pdfaSchema:namespaceURI>

                <!-- Preferred schema namespace prefix -->
                <pdfaSchema:prefix>zf</pdfaSchema:prefix>

                <!-- Description of schema properties -->
                <pdfaSchema:property>
                    <rdf:Seq>
                        <rdf:li rdf:parseType="Resource">
                            <pdfaProperty:name>DocumentID
                            </pdfaProperty:name>
                            <pdfaProperty:valueType>Text
                            </pdfaProperty:valueType>
                            <pdfaProperty:category>external
                            </pdfaProperty:category>
                            <pdfaProperty:description>invoice number;
HeaderExchangedDocument:ID
                            </pdfaProperty:description>
                        </rdf:li>

                        <rdf:li rdf:parseType="Resource">
                            <pdfaProperty:name>DocumentDate</pdfaProperty:name>
                            <pdfaProperty:valueType>Date
                            </pdfaProperty:valueType>
                            <pdfaProperty:category>external
                            </pdfaProperty:category>
                            <pdfaProperty:description>date when invoice was
issued; HeaderExchangedDocument:IssueDateTime
                            </pdfaProperty:description>
                        </rdf:li>

                        <rdf:li rdf:parseType="Resource">

```

```

<pdafProperty:name>DocumentFileName
</pdafProperty:name>
<pdafProperty:valueType>Text
</pdafProperty:valueType>
<pdafProperty:category>external
</pdafProperty:category>
<pdafProperty:description>name of the embedded xml
invoice file<
    /pdafProperty:description>
</rdf:li>

<rdf:li rdf:parseType="Resource">
<pdafProperty:name>DocumentType</pdafProperty:name>
<pdafProperty:valueType>Text
</pdafProperty:valueType>
<pdafProperty:category>external
</pdafProperty:category>
<pdafProperty:description>INVOICE
</pdafProperty:description>
</rdf:li>

<rdf:li rdf:parseType="Resource">
<pdafProperty:name>Version
</pdafProperty:name>
<pdafProperty:valueType>Text
</pdafProperty:valueType>
<pdafProperty:category>external
</pdafProperty:category>
<pdafProperty:description>The actual version of the
ZUGFeRD data
    /pdafProperty:description>
</rdf:li>

<rdf:li rdf:parseType="Resource">
<pdafProperty:name>ConformanceLevel
</pdafProperty:name>
<pdafProperty:valueType>Text
</pdafProperty:valueType>
<pdafProperty:category>external
</pdafProperty:category>
<pdafProperty:description>The conformance level of
the ZUGFeRD data
    /pdafProperty:description>
</rdf:li>

```

```
    </rdf:Seq>
    </pdfaSchema:property>
    </rdf:li>
  </rdf:Bag>
  </pdfaExtension:schemas>
</rdf:Description>
```