



<rdf:Description rdf:about="" xmlns:pdfaorg="http://pdfa.org/ns/presentations/">
<pdfaorg:Title>

XMP Metadata in PDF/A: Requirements and benefits

</pdfaorg:Title>

<pdfaorg:Presenter>

Dietrich von Seggern

</pdfaorg:Presenter>

<pdfaorg:PresenterCompany>

callas software GmbH

</pdfaorg:PresenterCompany>

<pdfaorg:PresenterCompanyExtension>

Founding member of PDF/A Competence Center

</pdfaorg:PresenterCompanyExtension>

</rdf:Description>



About callas software



- ❖ **Technology for PDF since 1997**
- ❖ **active contribution to ISO standardization of PDF since 1999**
- ❖ **Products**
 - ❖ **„Preflight“ in Adobe Acrobat Professional**
 - ❖ **pdfaPilot**
 - ❖ **pdfToolbox 4**
- ❖ **Available as**
 - ❖ **Plug-In for Acrobat**
 - ❖ **Server**
 - ❖ **Command line module**
 - ❖ **Library (C/C++, .Net, Java)**
 - ❖ **on Windows, Linux, Sun, AIX, Mac OS X**



Content



- ❖ **What is XMP Metadata?**
- ❖ **What are the requirements of PDF/A with respect to XMP Metadata?**
- ❖ **Demonstration**
 - ❖ **Analysis and validation of XMP Metadata**
 - ❖ **Creation and editing of XMP Extension Schemas**



About Metadata



- ❖ **“Information about data”**
- ❖ **Organizing metadata in document processing:**
 - ❖ **Where to place (inside or in an extra file)**
 - ❖ **Structure (Format)**
 - ❖ **Content**
- ❖ **Some Standards**
 - ❖ **Dublin Core (Office)**
 - ❖ **Title**
 - ❖ **Author**
 - ❖ **Keywords**
 - ❖ **Creation and modification date**
 - ❖ **IPTC for news**
 - ❖ **EXIF for images**



XMP “Extensible Metadata Platform” by Adobe



- ❖ **May be used in various file formats**
 - ❖ **TIFF, PostScript, JPEG, CreativeSuite etc.**
 - ❖ **Survives format conversions (most of the time)**
- ❖ **In a PDF it may be associated with the document but also with images, fonts, icc profiles, pages...**
- ❖ **Metadata properties are organized in schemas**



XMP Schemas



- ❖ **A schema contains one or more properties**
- ❖ **Is identified by means of a namespace URI**
 - ❖ **E.g.**
<http://pdfa.org/namespace/presentations/>
- ❖ **For every property**
 - ❖ **Name**
 - ❖ **Value type**
 - ❖ **String, Integer etc.**
 - ❖ **Also structured types or arrays**
 - ❖ **Description**



Predefined XMP Schemas



- ❖ **General schemas that are defined in the XMP Specification**
 - ❖ **Dublin Core Schema**
 - ❖ <http://purl.org/dc/elements/1.1/>
 - ❖ **XMP Basic Schema**
 - ❖ **XMP Media Management Schema**
 - ❖ **EXIF Schema**
 - ❖ **...**



Extending XMP



- ❖ **For Metadata that is not covered in predefined schemas**
 - ❖ **Metadata standards for industries or associations**
 - ❖ **Metadata structures from companies**
- ❖ **Recommendation: Do not invent XMP properties that are defined in the XMP specification (or anywhere else).**



Requirements of the PDF/A Standard



- ❖ **All XMP properties must be**
 - ❖ **either predefined in the relevant XMP Specification**
 - ❖ **defined in an embedded Extension Schema (name, value type, description ...)**
- ❖ **XMP Extension Schemas can be “translated” into XMP syntax for embedding**
 - ❖ **Description is *always* linked to the data**
 - ❖ **Forces complete and consistent definition**



PDF/A Validation and Conversion



❖ Validation

- ❖ **Predefined properties against XMP specification**
- ❖ **All other properties against the definitions of the embedded Extension Schema**

❖ Conversion

- ❖ **Not predefined XMP Metadata**
 - ❖ either **remove** or
 - ❖ **create and embedd XMP Extension Schema**



Demonstration



- ❖ **Analysis and validation of XMP Metadata**
 - ❖ **Acrobat 9**
 - ❖ **Preflight**
- ❖ **Creation and editing of XMP Extension Schemas**
 - ❖ **Remove Metadata using callas pdfaPilot 2**
 - ❖ **Create a XMP Extension Schema and embedd it using pdfaPilot 2**



Summary



Summary



- ❖ **XMP Metadata is flexibel and resistant (format conversion).**



Summary



- ❖ **XMP Metadata is flexibel and resistant (format conversion).**
- ❖ **XMP comes with a set of predefined schemas but may be extended.**



Summary



- ❖ **XMP Metadata is flexibel and resistant (format conversion).**
- ❖ **XMP comes with a set of predefined schemas but may be extended.**
- ❖ **PDF/A requires that a XMP Extension Schema is embedded for all properties that are not predefined.**



Summary



- ❖ **XMP Metadata is flexibel and resistant (format conversion).**
- ❖ **XMP comes with a set of predefined schemas but may be extended.**
- ❖ **PDF/A requires that a XMP Extension Schema is embedded for all properties that are not predefined.**
- ❖ **An XMP Extension Schema Editor manages XMP schemas, properties and types and allows you to embedd the schemas when converting to PDF/A.**



Further information



- ❖ **TechNotes at pdfa.org**
 - ❖ **0008 Predefined XMP Schemas**
 - ❖ **0009 XMP Extension Schemas**
- ❖ **Proceedings summary from the 3. International PDF/A Conference**
 - ❖ **E-mail to: rainer.lemke@pdfa.org**
 - ❖ **Non-Members: 129 EURO (+ VAT)**
 - ❖ **Members: 15 EURO (+ VAT)**



PDF/A Competence Center Webinars

Thank you for your attention!



Dietrich von Seggern
d.seggern@callassoftware.com

callas software GmbH
Founding member of PDF/A Competence Center



Dietrich von Seggern
Business Development
Manager

