



# Formatkonsolidierung im Langzeitarchiv

## PDF/A im In- und Output-Management

LuraTech Europe GmbH  
Carsten Heiermann





## Das Unternehmen LuraTech

LuraTech ist der führende Anbieter von Software-Technologie für JPEG2000 und hoch komprimiertes PDF(/A). Mit Software und Dienstleistungen im Bereich Document-Imaging, Capturing und Kompression liefert LuraTech schnell und kompetent Lösungen für jedes Imaging-Problem. LuraTech setzt auf ISO-Standard-konforme Produkte und bietet erstklassigen Support für Unternehmen und die öffentliche Hand.

- Gegründet 1995 in Berlin
- Eigentümergeführtes Unternehmen
- IST-Preis „groundbreaking products“ der EU für LuraDocument



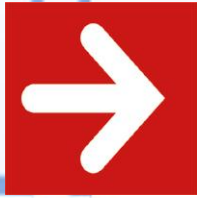
LuraTech

Member of:



[www.luratech.com](http://www.luratech.com)

[info@luratech.com](mailto:info@luratech.com)



## Inhalt

- **PDF/A**
  - Grundlagen, Ziele, Vorteile
  
- PDF/A Arbeitsschritte und Werkzeuge
  - Input – gescannte Dokumente
  - Konvertierung digitaler Dokumente
  - Output – Konvertierung
  - Validierung





## PDF/A – warum ein neuer Standard?

- PDF ist ein universell einsetzbares offenes Format für elektronische Dokumente
  - Mit veröffentlichter Spezifikation, aber (noch) proprietär
- Viele Unternehmen, Behörden und Institutionen schreiben die uneingeschränkte und unveränderte Verfügbarkeit von PDF-Dokumenten über viele Technologie-Generationen vor
- Native, teils private Formate wie MS Word genügen nicht den Anforderungen bei der Langzeitarchivierung
- PDF ist so mächtig und reich an Features, das Archivierungsbemühungen sich stark verkomplizieren
  - Anwenderspezifische Sub-Dialekte wären Konsequenz





## PDF/A – Zielsetzung der Standardisierung

- Langzeitarchivierung – lange oder gar „für immer“
- PDF/A soll sein:
  - Device-/ Software-/ Versions-unabhängig: Inhalte werden immer gleich dargestellt
  - Self Contained: eine PDF/A kompatible Datei beinhaltet alle Komponenten, die zur Darstellung nötig sind
  - Self Documented: eine PDF/A kompatible Datei beschreibt und dokumentiert sich selbst (Metadaten)
  - Transparent: Eine PDF/A kompatible Datei ist mit einfachen Mitteln analysierbar
- Dateiformat zur statischen Erhaltung des Erscheinungsbildes  
-> **fachspezifische Anwendungsrichtlinie!**





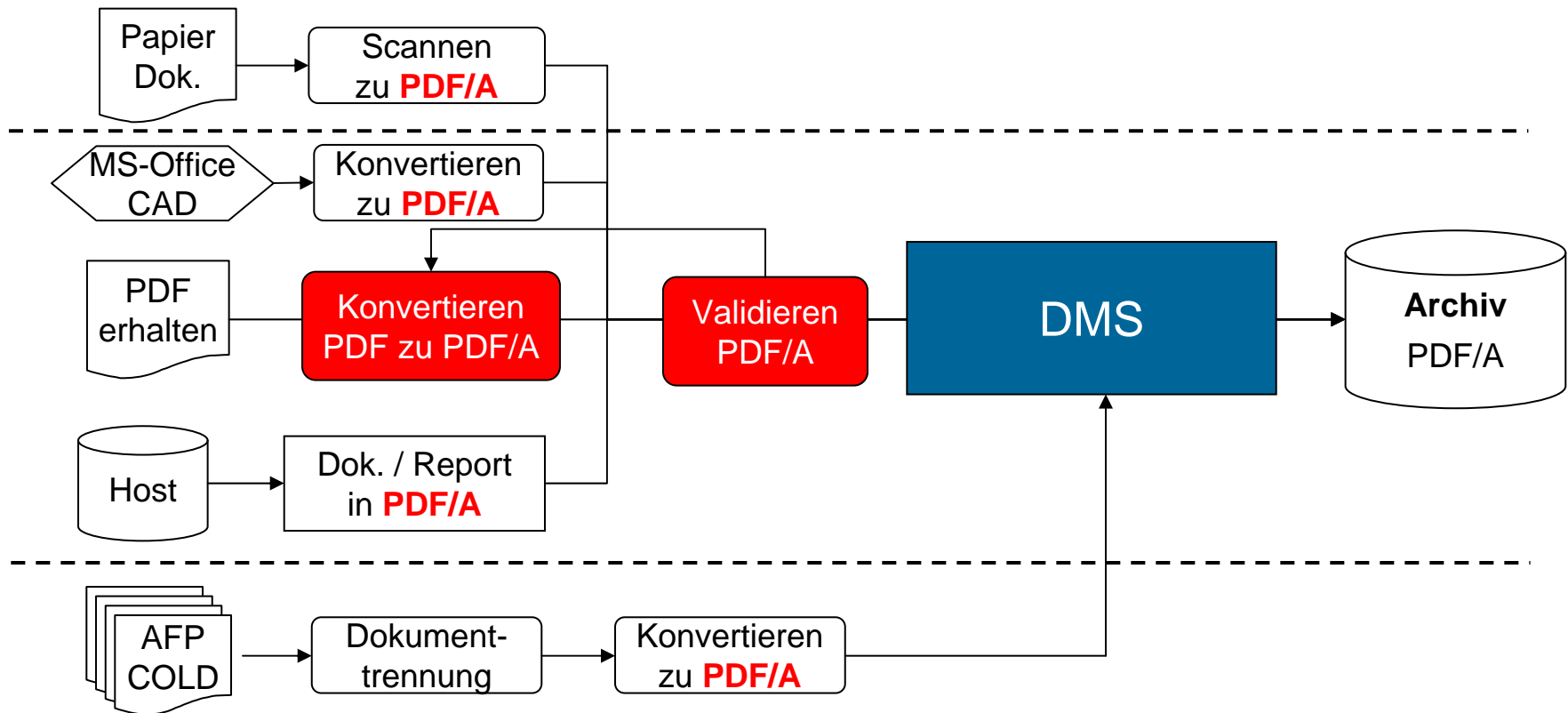
## PDF/A – Vorteile des unabhängigen PDF

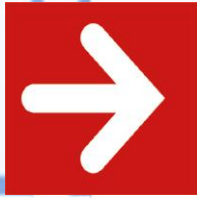
- PDF/A kann Volltext, automatisches Indexing, neue Kompressionsschemata, Metadaten etc.
- Vereinheitlichung des Archivformates
  - Scan in Farbe und schwarz-weiss - Rastergrafik
  - Digitale Dokumente – Vektorgrafik
  - Reine Datenströme XML, AFP, 1403, PCL
- ! Damit Reduzierung auf einen Viewer möglich
- ! Damit Migrationsaufwände geringer
- ! Damit langfristige Lesbarkeit garantiert, auch für Zeiträume, die keine Technologieprognosen erlauben





## Formatvereinheitlichung - das PDF/A Archiv

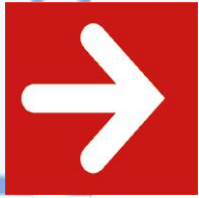




## Inhalt

- PDF/A
  - Grundlagen, Ziele, Vorteile
  
- **PDF/A Arbeitsschritte und Werkzeuge**
  - Input – gescannte Dokumente
  - Konvertierung digitaler Dokumente
  - Output – Konvertierung
  - Validierung





## Erzeugung PDF/A aus gescannten Dokumenten verbesserte Kompression

Bei schwarz/weissen Belegen

- JBIG2 - ISO/IEC 14492
  - Positioniert sich „statt TIFF G4“
  - Eingebettet in PDF/A, verfügbar im Reader

FAX G4



60 kB

JBIG2/lossless



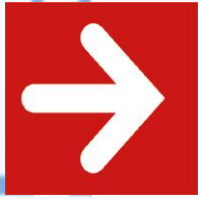
46 kB

JBIG2/lossy



29 kB





## Erzeugung PDF/A aus gescannten Dokumenten verbesserte Kompression

- Sehr hohe Kompression, voller Erhalt der Lesbarkeit
- Erhalt der Farbigkeit und der visuellen Qualität
- Volltextfähig, optimal für Farbe und auch schwarz/weiss

TIFF



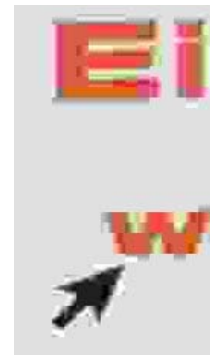
23,8 MB

FAX G4



60 kB

JPEG/PDF



180 kB

LuraDocument  
PDF/A



65 kB

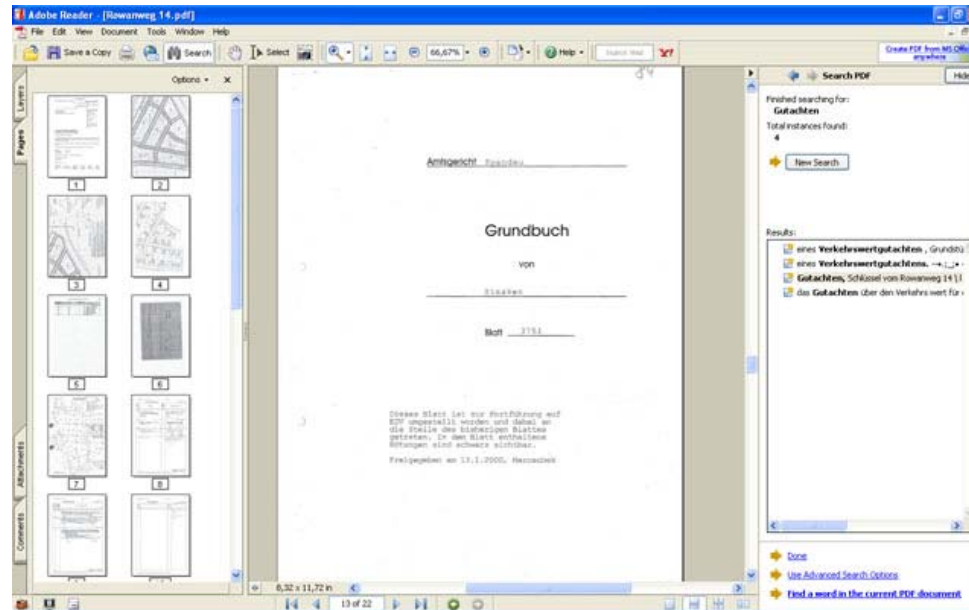


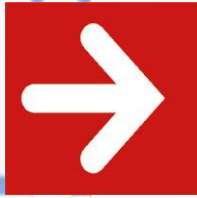


# Erzeugung PDF/A aus gescannten Dokumenten

## Durchsuchbarkeit der Akten auf Dateiebene

- z.B., elektronische Bibliothek, „nach dem Download“
- z.B. Handbuch, Konstruktionsunterlage oder Bauakte im Archiv im Rahmen der Produkthaftung
- z.B. Ein zum Kunden, Steuerberater, Rechtsanwalt gesendetes Dokument





## Erzeugung PDF/A aus gescannten Dokumenten- ein Wort zur Signatur

- Verfahren:
  - Signaturcontainer
  - externe Signaturdatei
  - eingebettete PDF Signatur
- Auch in PDF/A-1 ist die eingebettete Signatur standardkonform möglich (trotz PDF 1.4 Basis)!
- Aber, wie alle PDF-Operationen, muss die Signatur (bzw. die Formularfelder,... dafür) PDF/A konform eingebettet werden.





## Erzeugung PDF/A aus digitalen Dokumenten

- Applikationslogik wird benötigt
  - Rendering privater Formate wie Office oder CAD
  - Unter Umständen Betriebssystem Abhängigkeiten
- Alle Ressourcen werden benötigt
  - Schriften
  - Links auflösen, falls möglich
  - Transparenzen auflösen,....

-> Lösung schon bekannt: virtueller Druckertreiber



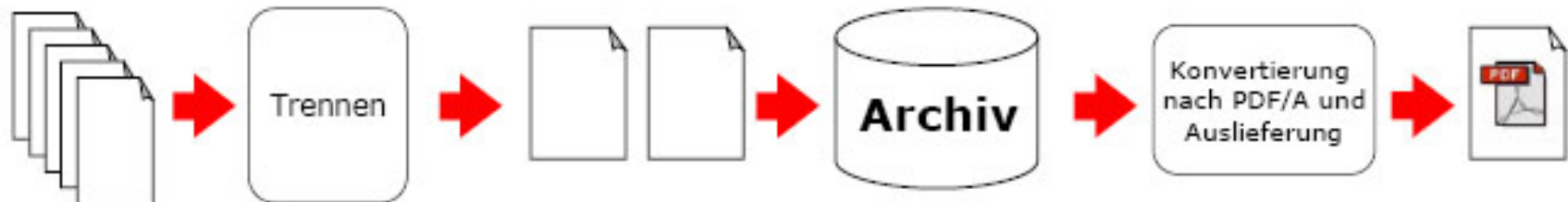


## Konvertierung im Output-Management

- „Reine Lehre“ –PDF/A mit eingebetteten Ressourcen



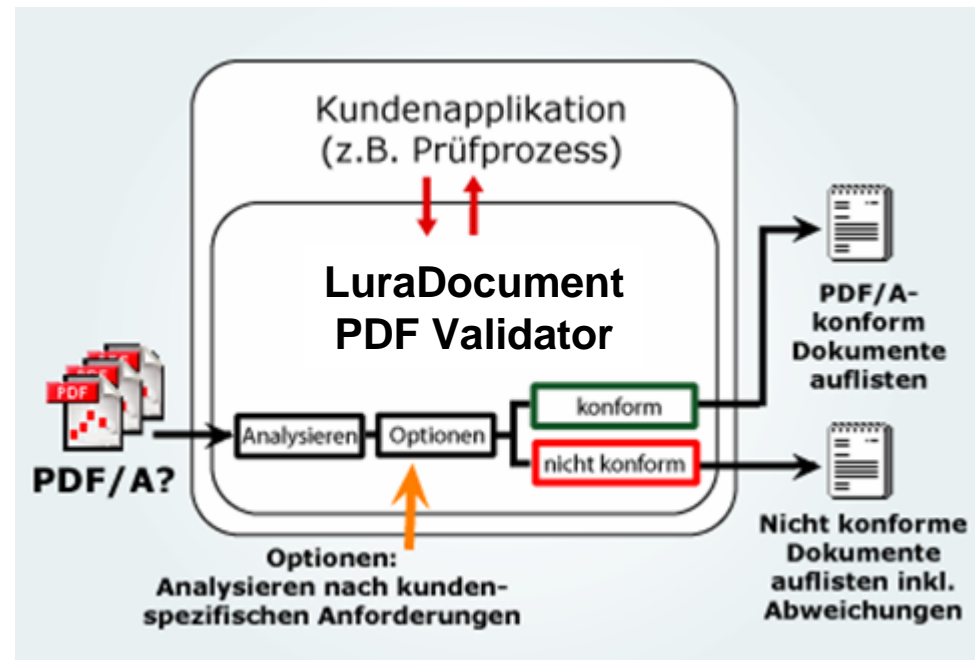
- Späte Konvertierung oder Ressourcenmanagement

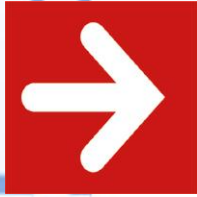




## Validierung – full PDF/A compliant?

- Test „auf PDF/A“, aber was?
  - Standard mit allen Referenzen: 10 Leitz-Ordner!
- Unterschied error / warning
- embedded fonts
- PDF Reference 1.4, PDF/A-1a, PDF/A-1b
- content of embedded XML, XMP metadata, ICC color profiles, PostScript?





## Noch Fragen?

Carsten Heiermann

Tel: 030/3940500, 0171/3377023, **02191/5922440**

c.heiermann@luratech.com

### Zentrale Europa:

LuraTech Europe GmbH  
Kantstr. 21

10587 Berlin  
Deutschland

### Büro Deutschland/West:

LuraTech Imaging GmbH  
Ulmenstr. 22

42855 Remscheid  
Deutschland

### USA:

LuraTech, Inc.  
1276 Lincoln Ave.,  
Suite 104  
San Jose, CA 95125  
USA



**LuraTech**

[www.luratech.com](http://www.luratech.com)

[info@luratech.com](mailto:info@luratech.com)