

Ein Sachstandsbericht Digitale Archive TIFF oder PDF/A



Gliederung des Vortrags

- Digitaler Geräte und Medien
- Zum Stand der EDV -Technik
- Zur Datenmenge
- Zur Datenspeicherung
- Digitale Archive
- Dateiformate (Langzeitarchivierung)
- Anwendung in unserem Verein
- Ausblick



Digitale Geräte und Medien

Ist nichts für mich !
Benutze ich nur selten !



Digitale Geräte und Medien 2005



Papstwahl 2005

Digitale Geräte und Medien 2013



Papstwahl 2013

Benutzung digitaler Geräte und Medien

- Navigation im Auto
- Telefonieren mit dem Handy
- Auskunft über das Handy erfragen
- Musik, Film, fotografieren über Handy, Tablet, Laptop (Internetdienste)
- Internet nutzen (Mail, Termine, Adressen, einkaufen, Bankdienste)
- „Skype“ Bildtelefon
- Push Nachrichten „WhatsApp“
- Berufliche Nutzung (schreiben, rechnen lesen, fotografieren)
- Tickets im Internet kaufen usw. usw.
- Die Entwicklung ist nicht abgeschlossen



Zum Stand der EDV Technik 1980 - 2016

- Bei geringeren Kosten ist die Leistungsfähigkeit der EDV-Geräte gestiegen.
- **Kosten eines Desktop PC: von 13.000 DM auf 500 € gefallen!**
- Datenübertragung von großen Datenmengen ist in Milli-Sekunden möglich.
- Datenspeicherung und Datensicherung ist preisgünstig im Cloud möglich



Zur Datenspeicherung 1980 - 2015

- **1980:** 1 Petabyte aus 1.000.000 IBM 3380 Festplatten
- 1980 benötigte man, noch 1.000.000 IBM-3380-Festplatten mit einer Kapazität von jeweils 1 GByte, welche zusammen Kosten in Höhe von 40 Milliarden US-Dollar verursachten. Das Gewicht der reinen Festplatten, also ohne Stromversorgung und Server-Racks, beläuft sich in diesem Beispiel auf 250.000 Tonnen. (1 GByte = 40.000 US-Dollar)
- Heute benötigt man für 1 Petabyte Speicherplatz lediglich 500 Festplatten des Typs Western Digital Eco Green 2TB, welche zusammen nur 365 Kilogramm schwer und 125.000 US-Dollar teuer sind. (1 TB = 125 US-Dollar)
- Bei gleichen Kosten (40 Milliarden US-Dollar) für den Datenspeicher hat sich die für diesen Betrag zu speichernde Datenmenge um das **320.000 fache** vergrößert.

Datenspeicher 1980 - 2015



Datenspeicher heute



Zur Menge der Daten

- Der Datenbestand verdoppelt sich alle zwei Jahre
- Seit etwa elf Jahren werden Daten häufiger digital als analog gespeichert.
- Seitdem hat auch die Vielfalt der Daten-Typen zugenommen. Dazu gehören Videos, Audiodateien, Texte in E-Mails, Nachrichten bei [Twitter](#), PowerPoint-Präsentationen oder Suchanfragen bei Google.
- **Die Zukunft gehört der Cloud (NAS)**



Zum Stand der Softwareentwicklung

- Programme können Daten schneller verwalten und große Mengen lesbare Daten auswerten



Digitale Archive

Warum ?

Doch nicht für mich!



Digitale Archive

- Private Sammlungen
 - Familiendokumente
 - Kochrezepte
 - Fotosammlungen
- Firmenarchive
- Behördenarchive
- Öffentliche Archive
 - Staatsarchive
 - Bibliotheken
- Vereinsarchive
 - Buch- und Schriftensammlung



Dateiformate für die Langzeitarchivierung

- **JPEG** (Format für Bilder)
- **TIFF** (Archivformat für Bilder)

- **PDF** (Format für Texte mit Bilder/ Grafiken)
- **PDF/ A** (Archivformat für lesbare Texte mit Bilder)



Dateiformate für die Langzeitarchivierung

- **JPEG** (Format für Bilder)
Bedingt geeignet; Metadaten verwenden

- **TIFF** (Archivformat für Bilder)
 - Für Langzeitarchivierung geeignet
 - Große Dateien
 - Metadaten unbedingt verwenden



- **PDF** (Format für Texte mit Bilder / Grafiken)
- PDF ist **kein** ISO Standard
- PDF Dokumente unterstützen **keine** Volltextsuche
- PDF Dokumente bleiben **nicht** fristlos gültig



Dateiformate für die Langzeitarchivierung

- **PDF/A** (Archivformat für lesbare Texte mit Bilder)
 - Für Langzeitarchivierung erforderlich
 - Metadaten unbedingt verwenden
 - PDF/A ist ein ISO Standard
 - PDF/A Dokumente unterstützen Volltextsuche



PDF/A (Archivformat für lesbare Texte mit Bilder

- 1.PDF/A ist ein ISO Standard
- 2.PDF/A Dokumente unterstützen Volltextsuche
- 3.Alle Informationen, die für eine Wiedergabe benötigt werden, sind im Dokument enthalten
- 4.PDF/A spart Speicherplatz
- 5.Optimale Sicherheit dank digitalen Signaturen
- 6.PDF/A kann die Richtlinien für behindertengerechte Informationsangebote erfüllen
- 7.PDF/A Dokumente bleiben fristlos gültig
8. PDF/A ist in weiten Kreisen akzeptiert
- 9.Zuverlässige Werkzeuge sind bereits erhältlich



Digitales Archiv im Verein

- Texte digital lesbar und durchsuchbar.
- Texte im Cloud zugänglich.
- Texte über Findbücher am eigenen PC zugänglich
- Öffentlich Archive und Bibliotheken digital vorhanden.
- Dokumente können für die Forschung genutzt werden.



Anwendung im Verein

- Unsere Homepage www.ursella.org
- Denkmale in Oberursel www.geschichtedigital.de
- Buchdruck in Oberursel www.ursellis.de
- Unsere Mitteilungen 1963 – 2014 Internet und DVD
- Faktenbücher



Ausblick

- Digitale Dokumente werden in vielen Bereichen erstellt und genutzt. (Stichwort: Digitale Akte)
- Archivierung erfolgt in digitaler Form. (Zeitungen, Akten usw.)
- Geschichtsforschung mit digitalen Dokumenten.
- Langzeitarchivierung erforderlich.
- Suche im Internet (digital) Standard
- Händische Archive weiterhin notwendig.



Besten Dank



1) PDF/A ist ein ISO Standard

Ein ISO Standard ist eine international erarbeitete Norm, die das optimale Verfahren definiert, um eine Spezifikation zu erfüllen. Das Dateiformat PDF/A ist seit 2005 ein ISO Standard und speziell auf die Anforderungen der Archivierung zugeschnitten. Als offener Standard ist er unabhängig von Plattformen oder Anbietern. PDF/A basiert auf jahrelanger Erfahrung mit PDF (seit 1993).

Dieses Format wird in allen Branchen verwendet und ist allgemein etabliert. Der PDF/A Standard garantiert zusammen mit den unzähligen Betrachtern (z.B. Adobe Acrobat/Reader) auch in Zukunft die Lesbarkeit PDF/A konformer Dokumente. Der PDF/A Standard enthält alle Informationen, um heute, aber auch in Zukunft einen PDF/A konformen Betrachter zu erstellen. PDF/A wird laufend weiterentwickelt. Im Jahre 2011 wurde der 2. Teil und 2012 bereits der 3. Teil des Standards freigegeben.

2) Standardisierte Metadaten sind im Dokument eingebettet

Alle Informationen, die dazu dienen ein Dokument zu beschreiben (sogenannte Metadaten), werden direkt im PDF/A Dokument eingebettet. Diese Metadaten werden dabei in einem standardisierten, austauschbaren Format (XMP, offener XML Standard) festgehalten. Metadaten anderer Formate (z.B. PNG, PostScript oder TIFF) können mühelos in das PDF/A Dokument übertragen werden. Die Verwendung (oft proprietärer) Systeme, um Metadaten separat zu speichern, wird überflüssig.

3) PDF/A Dokumente unterstützen Volltextsuche

Die Möglichkeit, eine Volltextsuche durchzuführen, ist auch Bestandteil des Standards. Der Text eines digital erstellten Inhalts ist im Dokument selbst integriert. Das gilt auch für Scans, die mit optischer Zeichenerkennung OCR (Optical Character Recognition) eingelesen worden sind. Das PDF/A Dokument speichert den erkannten Text (als Unicode) wie auch das Originalbild. So bleibt das visuelle Erscheinungsbild unverändert und die Suchfunktion erhalten.

4) Alle Informationen, die für eine Wiedergabe benötigt werden, sind im Dokument enthalten

PDF/A Dokumente sind selbständig und unabhängig: Elemente (Schriften, Farbprofile, usw.), die für eine einwandfreie Wiedergabe benötigt werden, sind im Dokument enthalten. Ein PDF/A Dokument darf keine Verweise zu externen Quellen aufweisen. Einfache informative Verweise wie zum Beispiel Links zu Webseiten sind erlaubt.

5) PDF/A spart Speicherplatz

Auch wenn PDF/A Dokumente mehr Informationen als Bilddateien (z.B. TIFF) enthalten, brauchen sie dank effizienten Kompressions- und Farbtrennungsverfahren in der Regel weniger Speicherplatz.

6) Optimale Sicherheit dank digitalen Signaturen

Die Kombination von PDF/A und digitalen Signaturen stellt sicher, dass das Dokument nicht manipuliert wurde, sondern den ursprünglichen, vom Autor beabsichtigten Inhalt darstellt. Das gewährleistet Rechtssicherheit für die Langzeitarchivierung.

7) PDF/A kann die Richtlinien für behindertengerechte Informationsangebote erfüllen (USA: Rehabilitation Act, Section 508)

Der PDF/A Standard teilt sich in Unterstandards. Der Unterstandard ‚Part A‘ (PDF/A-1a, PDF/A-2a und PDF/A-3a) stellt die ausführlichste Variante dar, welche Informationen beinhaltet, um Menschen mit besonderen Bedürfnissen besser zu unterstützen (Barrierefreiheit). So lässt sich z. B. der Text als Unicode darstellen, und eignet sich damit für computererzeugte Sprachsynthese. Zudem bleibt die Dokumentenstruktur jederzeit erhalten; die Leserichtung sowie eine Beschreibung der Bilder in Worten sind zu diesem Zweck im Standard inbegriffen. Andere Unterstandards sind weniger anspruchsvolle Varianten, genügt aber meistens. Typische Anwendungsbereiche sind gescannte Dokumente und die Umwandlung digital erzeugter Dokumente nach PDF/A.

8) PDF/A Dokumente bleiben fristlos gültig

Das ISO Standard Komitee erweitert den Standard alle zwei bis vier Jahre. Das heisst nicht, dass bestehende PDF/A Dokumente umgewandelt werden müssen; sie bleiben immer gültig. ISO darf den PDF/A Standard nicht widerrufen.

9) PDF/A ist in weiten Kreisen akzeptiert

In Europa und Asien wird PDF/A bereits empfohlen oder vorgeschrieben. Regierungen, Organisationen und Unternehmen definieren das PDF/A Format als Standard für Langzeitarchivierung. Auch in Nordamerika existieren solche Richtlinien, v. a. bei Gerichten, Bibliotheken und Kernkraftwerken. Die PDF Association trägt massgeblich zur Verbreitung des Format zu.

10) Zuverlässige Werkzeuge sind bereits erhältlich

Dank der langjährigen Erfahrung mit PDF und der schnellen Reaktion der wichtigsten Hersteller von PDF Technologien, gibt es bereits eine Vielfalt an PDF/A konformen Software Werkzeugen auf dem Markt. Ein entsprechender Validator hilft zu überprüfen, ob ein Dokument dem PDF/A-Standard einhält.

Verein für Geschichte und Heimatkunde Oberursel (Taunus) e.V.

Digitaler Stammtisch

Vortrag am 08.11.2016, Stadthalle Oberursel, Raum

Digitale Datenarchive im Internet – Heute und in Zukunft

Erschließung und Zugriff auf digitale Dokumente durch das Dateiformat PDF/A

Referent: Erwin Beilfuss

Anleitungen zur Erstellung von PDF /A 1b

www.pdf-tools.com

<http://www.pdf-tools.com/public/downloads/whitepapers/Whitepaper-PDFA-Standard-ISO-19005-DE.pdf>

https://www.uni-muenster.de/imperia/md/content/publizieren/anleitung_pdf_pdfa.pdf

Anleitung und Beratung für die digitale Archivierung

<https://archivberatung.hessen.de/>

Digitale Archive und Bibliotheken in Deutschland, Stand: 11.2016

<http://www.lagis-hessen.de/>

<http://www.europeana.eu/portal/de>

<http://www.bundesarchiv.de/index.html.de>

<http://www.bundesarchiv.de/gedenkbuch/directory.html.de>

<https://www.hessen.de/fuer-besucher/geschichte-des-landes-hessen>

<https://www.deutsche-digitale-bibliothek.de/?lang=de>

<https://www.archivportal-d.de/>

<https://www.digitale-sammlungen.de/>

<http://www.ifskb.de/node/313>

<http://www.zvdd.de/startseite/>

<portal.kobv.de/>