

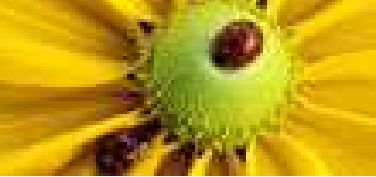
PDF in Theorie und Praxis

Erstellen • Analysieren • Bearbeiten • Drucken

Ralf König, ralk@hrz.tu-chemnitz.de

Technische Universität Chemnitz, Fakultät für Informatik

<http://www.tu-chemnitz.de/informatik>



Inhalt

● Inhalt

● Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

Ende.

Theorie

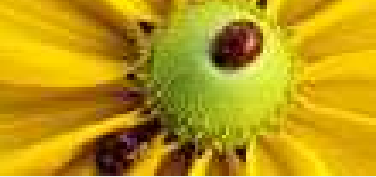
Entwicklung und Geschichte
Vergleich mit PostScript
Struktur von PDF-Dateien

Praxis

Erstellen von PDF-Dateien
Analysieren von PDF-Dateien
Bearbeiten von PDF-Dateien
Drucken von PDF-Dateien

Philosophie Spezialthemen

Quellen



● Inhalt

● Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

Ende.

■ PDF Reference, Fourth Edition, Version 1.5

<http://partners.adobe.com/asn/tech/pdf/specifications.jsp>

■ Thomas Merz, c't 07/2002, Ein Rezept für alle – Was im Portable Document Format steckt

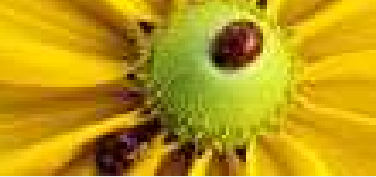
■ Thomas Merz, c't 01/2003, Sorglos-Drucksache – PDF/X soll eine reibungslose Produktion garantieren

■ Thomas Merz, PostScript & Acrobat/PDF, Springer 1997, ISBN 3-540-60854-0.

■ Thomas Merz, Die PostScript- & PDF-Bibel, d.punkt Verlag 2002, ISBN 3-935320-01-9.

■ `news://comp.text.pdf`

■ Troubleshooting von PDF-Dateien im URZ



- Inhalt
- Quellen

Was ist PDF?

- PDF - Ein paar Beispiele
- Entwicklung von PDF
- Gemeinsamkeiten von PostScript und PDF
- Unterschiede zwischen PostScript und PDF
- Konvertierung zwischen PostScript und PDF

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

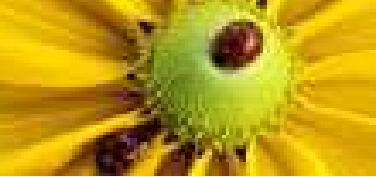
PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

Ende.

Was ist PDF?



PDF - Ein paar Beispiele

- Inhalt
- Quellen

Was ist PDF?

● PDF - Ein paar Beispiele

- Entwicklung von PDF
- Gemeinsamkeiten von PostScript und PDF
- Unterschiede zwischen PostScript und PDF
- Konvertierung zwischen PostScript und PDF

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

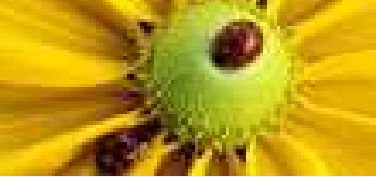
Rechtliches

Spezialthemen

Ende.

PDF ist ein Container-Format für verschiedenste digitale Inhalte.

- elektronische Dokumente mit druckbarem Inhalt (ähnlich PostScript)
- gelocht und mit runden Ecken? ;-)
- mit Animation
- mit Klang
- mit eingebetteten beliebigen Dateien (ähnlich einem Zip-Archiv)
- als Formular zum Ausfüllen, mit selbstberechnenden Feldern
- als kleine Anwendung (Taschenrechner, Spiel)
- als E-Book mit erweiterten Maßnahmen für Digital Rights Management (DRM)



Entwicklung von PDF

- Inhalt
- Quellen

Was ist PDF?

- PDF - Ein paar Beispiele
- Entwicklung von PDF
- Gemeinsamkeiten von PostScript und PDF
- Unterschiede zwischen PostScript und PDF
- Konvertierung zwischen PostScript und PDF

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und Troubleshooting

Rechtliches

Spezialthemen

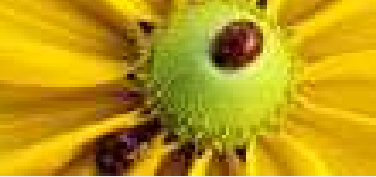
Ende.

- geschaffen von Adobe-Mitgründer John Warnock, 1990 (Projektname: Camelot, später Carousel)
- begann als vereinfachte Version von PostScript, später immer mehr Web-Features

PostScript 1		1985	
	.ai format	1988	
PDF 1.0	Acrobat 1	1993	214 S.
PDF 1.1	Acrobat 2	1994	298 S.
PDF 1.2	Acrobat 3	1996	394 S.
PDF 1.3	Acrobat 4	1999	696 S.
PDF 1.4	Acrobat 5	2001	978 S.
PDF 1.5	Acrobat 6	2003	1172 S.
PDF 1.6?	Acrobat 7	2005?	(Reader Beta seit 15.9.)

Genauer:

<http://www.prepressure.com/pdf/history/history01.htm>



Gemeinsamkeiten von PostScript und PDF

- Inhalt
- Quellen

Was ist PDF?

- PDF - Ein paar Beispiele
- Entwicklung von PDF
- Gemeinsamkeiten von PostScript und PDF
- Unterschiede zwischen PostScript und PDF
- Konvertierung zwischen PostScript und PDF

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

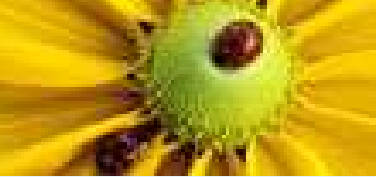
PDF drucken und Troubleshooting

Rechtliches

Spezialthemen

Ende.

- beide eignen sich zur Beschreibung physischer Seiten mit grafischen Elementen als Inhalt als Basis für den Druck
- ähnliche Ausdrucksmöglichkeiten für Zeichen-Anweisungen, gleiche Logik (Adobe Grafik-Modell)
- Namen der Operatoren wurden in PDF abgekürzt (PS: „moveto“, PDF: „m“, .ai-Format)
- Grund: in PostScript-Dateien hatte jeder sein eigenes Dictionary aufgebaut



Unterschiede zwischen PostScript und PDF

- Inhalt
- Quellen

Was ist PDF?

- PDF - Ein paar Beispiele
- Entwicklung von PDF
- Gemeinsamkeiten von PostScript und PDF
- Unterschiede zwischen PostScript und PDF
- Konvertierung zwischen PostScript und PDF

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und Troubleshooting

Rechtliches

Spezialthemen

Ende.

PostScript

stack-basierte Programmiersprache mit allen Eigenschaften

nur druckbare Objekte

Seiten hängen voneinander ab (lineare Struktur)

nur Schriften und Bilder sind sinnvoll einzubetten

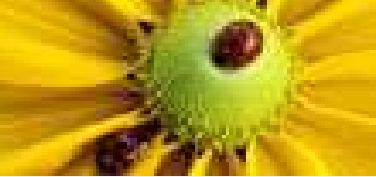
PDF

Seitenbeschreibungssprache (+ JavaScript für begrenzte Programmierbarkeit)

zusätzliche Web-Features: Hyperlinks, Formulare, Verschlüsselung, mehr Kompressionsalgorithmen

Seiten sind voneinander unabhängig (Baum-Struktur)

digitale Objekte fast jeden Typs können eingebettet werden (Plug-In mechanismus in AR)



Konvertierung zwischen PostScript und PDF

- Inhalt
- Quellen

Was ist PDF?

- PDF - Ein paar Beispiele
- Entwicklung von PDF
- Gemeinsamkeiten von PostScript und PDF
- Unterschiede zwischen PostScript und PDF
- Konvertierung zwischen PostScript und PDF

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

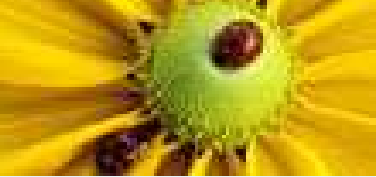
PDF drucken und Troubleshooting

Rechtliches

Spezialthemen

Ende.

- PS \Rightarrow PDF (z.B. zur Bereitstellung im Web):
 - ◆ interpretieren des PS, ausrechnen aller Zeichenkommandos
 - ◆ ersetzen aller Makros durch ihr Ergebnis bei Ausführung
 - ◆ ausrechnen eines *Graphics State* für jede Seite (macht die Seiten unabhängig voneinander)
 - ◆ komprimieren der eingebetteten Schriften und Bilder
 - ◆ pdfmark-Mechanismus für Anweisungen in der PostScript-Quelle an den Distiller
- PDF \Rightarrow PS (eg. beim Drucken auf einen PostScript-Drucker):
 - ◆ dekomprimieren der druckbaren Seiteninhalte, wie Text und Vektorgrafik
 - ◆ hinzufügen eines Makro-Dictionary zum Ummappen der Operatoren
 - ◆ ersetzen mancher Operatoren
 - ◆ dekomprimieren der eingebetteten Schriften und Bilder
 - ◆ Rest (Javascript Code, etc.) wird meist verworfen.



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

- Struktur
- Beispiel

PDF erstellen

PDF analysieren

PDF bearbeiten

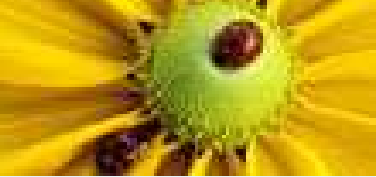
PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

Ende.

PDF-Struktur



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

● Struktur

● Beispiel

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

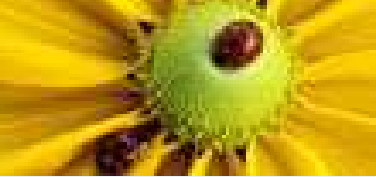
Ende.

Folien von Jim King (Adobe)

http://www.planetpdf.com/planetpdf/pdfs/seiboldseminars/ny02/king_jim_PDFAsPackageAnnot.pdf

Zusammenfassung:

- Graph von nummerierten Objekten mit Referenzen
- jedes Objekt hat eine Objekt-ID und ein Revision-Level, z. B.
`75 0 obj`
- Streams können komprimiert oder umkodiert werden
- `xref` Tabelle, um die Objekte schnell zu finden \Rightarrow erschwert Modifikationen
- aber: die meisten PDF-Tools versuchen eine Reparatur der `xref` Tabelle



Beispiel

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

● Struktur

● **Beispiel**

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

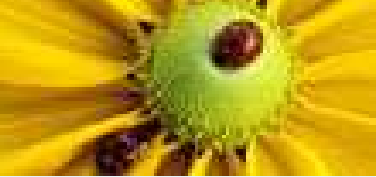
Ende.

\LaTeX -Einzeiler:

```
\documentclass{article}
\begin{document}
Hello World
\end{document}
```

PS-Datei: test1.ps (12 kB)

PDF-Datei: test1.pdf (7 kb)



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

- Mehrere Wege
- 1. Erstellung über Postscript
- 2. Nativer PDF-Export
- 3. Tools für Bilder
- PDF-Bibliotheken

PDF analysieren

PDF bearbeiten

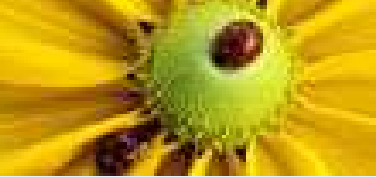
PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

Ende.

PDF erstellen



Mehrere Wege

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

● Mehrere Wege

- 1. Erstellung über Postscript
- 2. Nativer PDF-Export
- 3. Tools für Bilder
- PDF-Bibliotheken

PDF analysieren

PDF bearbeiten

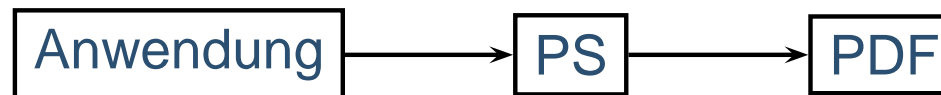
PDF drucken und
Troubleshooting

Rechtliches

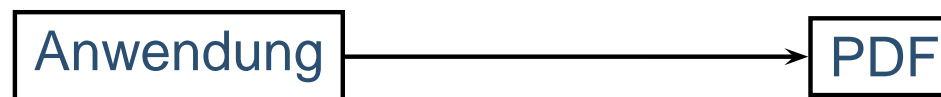
Spezialthemen

Ende.

1. über PostScript und Konvertierung

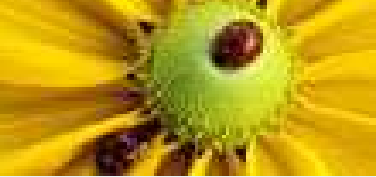


2. direkt aus einer Anwendung exportieren

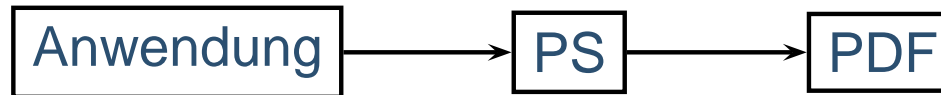


3. für Einzelbilder gibt es auch weitere Tools (z.B. tiff2pdf, jpeg2pdf, svg2pdf, png2pdf)





1. Erstellung über Postscript



Adobe Distiller: konvertiert von Postscript nach PDF (viele Optionen)

Linux

- Ghostscript: ps2pdf ist ähnlich zum Distiller. Möglichst neue Ghostscript-Version (AFPL 8.14) nutzen.
- Bequemer PDF-Druckertreiber (Alternative zu Adobe Distiller):
 - ◆ Linux: meist schon eingerichtet nach Installation, wandelt PS nach PDF mit Hilfe von GhostScript
 - ◆ Windows: es gibt mehrere Wrapper für Ghostscript, z.B. FreePDF
 - ◆ für MS Office gibt es GhostWord, ähnlich zu Adobe PDFMaker (per pdfmark: Links erhalten, Bookmarks generieren)

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

- Mehrere Wege
- 1. Erstellung über Postscript
- 2. Nativer PDF-Export
- 3. Tools für Bilder
- PDF-Bibliotheken

PDF analysieren

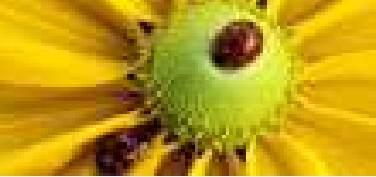
PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

Ende.



2. Nativer PDF-Export

Anwendung

PDF

Adobe PDFWriter: arbeitet auch als Druckertreiber, konvertiert direkt nach PDF ohne Umweg über Postscript (nur Basisfunktionalität, aber schneller als Distiller)

Programme unter Linux mit PDF-Export

- OpenOffice (eigene Bibliothek) (suboptimal, z.B. Kreise zusammengesetzt aus Geradenstücken)
- Scribus (eigene libpdf) (sieht ganz gut aus, aber beschränkter Umfang)
- pdfTeX (eigene Bibliothek) (sehr mächtig)
- vTeX/free (sogar eigener PostScript-Interpreter GeX)
- Sketch (über reportlab)
- `lout` (Dokument-Formattierer)

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

- Mehrere Wege
- 1. Erstellung über Postscript
- 2. Nativer PDF-Export
- 3. Tools für Bilder
- PDF-Bibliotheken

PDF analysieren

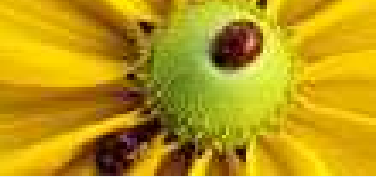
PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

Ende.



3. Tools für Bilder



Konvertieren (vor allem aus Bildformaten):

- `tiff2pdf`, `tiff2ps`
- `jpeg2ps`, `jpeg2pdf`
- `png2pdf`
- `convert` aus Imagemagick
- SVG Transcoder aus Batik (SVG to PDF)

Spezialkonverter sind den Allzweck-Tools meist überlegen, z. B. `tiff2pdf` erhält G4 Kompression, `convert` (nutzt Ghostscript) nicht

`tiff2pdf` eignet sich ganz gut zum Erstellen von gescannten Büchern in PDF

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

- Mehrere Wege
- 1. Erstellung über Postscript
- 2. Nativer PDF-Export
- 3. Tools für Bilder
- PDF-Bibliotheken

PDF analysieren

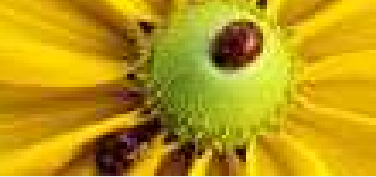
PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

Ende.



PDF-Bibliotheken

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

- Mehrere Wege
- 1. Erstellung über Postscript
- 2. Nativer PDF-Export
- 3. Tools für Bilder
- PDF-Bibliotheken

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

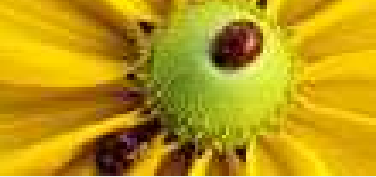
Rechtliches

Spezialthemen

Ende.

- vor allem für Entwickler interessant
- Scribus hat begonnen, eine libpdf zu schreiben, es gibt aber auch schon viele andere Ansätze
- bis jetzt keine Konvergenz erkennbar
- xpdf (C++), pdflib lite (C), reportlab (Python), iText, pjax, pdfbox (alle Java), fpdf (PHP), ...
- Link-Sammlung:

<http://rnvs.informatik.tu-chemnitz.de/twiki/bin/view/Main/FreePdfUtilitiesAndLibraries>



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

- Streams
- Streams entpacken
- Werkzeuge für die Analyse

PDF bearbeiten

PDF drucken und
Troubleshooting

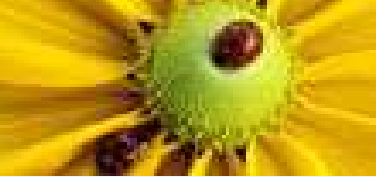
Rechtliches

Spezialthemen

Ende.

PDF analysieren

Streams



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

- Streams
- Streams entpacken
- Werkzeuge für die Analyse

PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

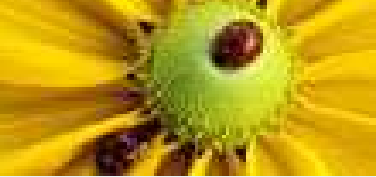
Spezialthemen

Ende.

- Ziel: „Gutes“ von „schlechtem“ PDF unterscheiden.
- Gutes PDF: für die entsprechenden Objekte werden auch die passenden Datentypen genutzt, es wird nicht unnötig konvertiert (Bsp. Kreis zu Linienstücken)
- wenig redundanter Code, dazu analysiert man die Content Streams
- Bilder in gut komprimierenden Formaten, dazu analysiert man die XObjects
- content stream enthält **Text**, **Vektorgraphik** and **Referenzen zu XObjects**

```
BT
    /F1 24 Tf
    1 0 0 1 260 330 Tm
    (Hello World)Tj
ET
100 100 200 200 re
```

Streams entpacken



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

- Streams
- Streams entpacken
- Werkzeuge für die Analyse

PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

Ende.

vorher: die interessanten Objekte im PDF sind gepackt

nachher: alle sinnvoll zu entpackenden Objekte sind entpackt
und können editiert werden

- einzelnes Objekt entpacken:

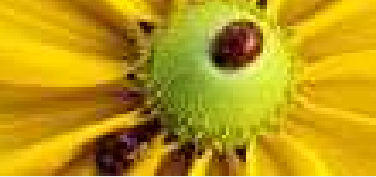
```
pdftosrc <file> <object-nr>
```

- gesamtes Dokument: aus Multivalent:

```
tool.pdf.Uncompress, tool.pdf.Compress
```

- alternativ mit `pdftk uncompress, compress`

Werkzeuge für die Analyse



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

- Streams
- Streams entpacken
- **Werkzeuge für die Analyse**

PDF bearbeiten

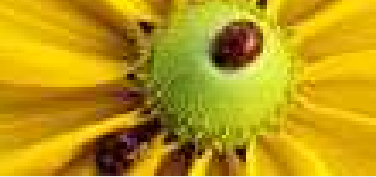
PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

Ende.

- **Multivalent PDF Tools:** `tool.pdf.Info`,
`tool.pdf.Decrypt`, `tool.pdf.Validate`,
`tool.font.view`
- **pdftk (PDF Tool Kit):** `pdftk dump_data`, `update_info`,
`repair`, `unpack_files`
- `pdfinfo`, `pdffonts (xpdf)`
- **Editor:** `xemacs`, `vim`
- **Viewer:** `acroread`, `xpdf`, `gpdf`, `gv`, `ggv`,
`kghostview`
- **Rasterizer:** `gs`, `pdftoppm (xpdf)`
- **Extrahieren von Ressourcen:**
 - ◆ **Rasterbilder:** `pdfimages (xpdf)`
 - ◆ **Text:** `pdftotext (xpdf)`
 - ◆ **Vektorgraphik (am einfachsten: ganze Seite extrahieren,
dann MediaBox oder CropBox ändern)**



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

- Auswahl und Anordnung der Seiten
- Information hinzufügen
- Nutzung von PDF-Bibliotheken
- Grafisches Editieren

PDF drucken und
Troubleshooting

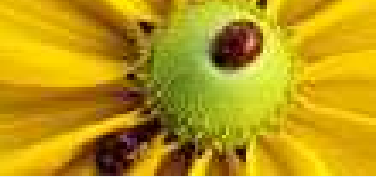
Rechtliches

Spezialthemen

Ende.

PDF bearbeiten

Auswahl und Anordnung der Seiten



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

● Auswahl und Anordnung der Seiten

● Information hinzufügen

- Nutzung von PDF-Bibliotheken
- Grafisches Editieren

PDF drucken und Troubleshooting

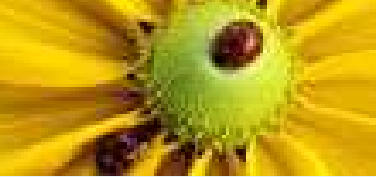
Rechtliches

Spezialthemen

Ende.

Ausschießen ist selbst mit Acrobat (pur) nur schwer zu machen. Acrobat kann man aber skripten, es gibt dafür mehrere Plug-Ins.

- `pdfpages` (\LaTeX package)
- `pdfnup.c` (Wrapper um `pdf \LaTeX` und `pdfpages`)
- `pdfjam` = (`pdf90`, `pdfnup`, `pdfjoin`) (Wrapper um `pdf \LaTeX` und `pdfpages`)
- `tool.pdf.Split`, `Merge`, `Impose` (Multivalent): Seiten umarrangieren, n-up Drucken, Seiten auf einem Bogen anordnen
- `pdftk cat burst`: Seiten umarrangieren, Einzel-PDF-Seiten erzeugen



Information hinzufügen

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

- Auswahl und Anordnung der Seiten
- **Information hinzufügen**
- Nutzung von PDF-Bibliotheken
- Grafisches Editieren

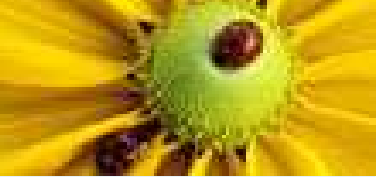
PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

Ende.

- Lesezeichen hinzufügen: `mbtpdfasm o`
- Stempel (z.B. fürs Druckdatum) hinzufügen: `mbtpdfasm t, T`
- Watermarks: `pdftk background`
- pdfL^AT_EX mit `pdfpages` und `eso-pic`



Nutzung von PDF-Bibliotheken

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

- Auswahl und Anordnung der Seiten
- Information hinzufügen
- Nutzung von PDF-Bibliotheken
- Grafisches Editieren

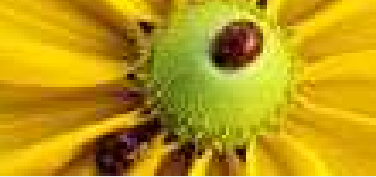
PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

Ende.

- für eigene Tools, die es bis jetzt noch nicht gibt
- manche (nicht alle!) der oben genannten Bibliotheken eignen sich auch für den PDF-Import
- dann kann man Konverter bauen, die das PDF nach den eigenen Wünschen umbauen



Grafisches Editieren

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

- Auswahl und Anordnung der Seiten
- Information hinzufügen
- Nutzung von PDF-Bibliotheken
- Grafisches Editieren

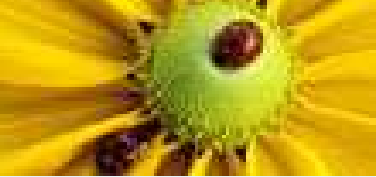
PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

Ende.

- Scribus (PDF-Import seit V1.2, vorher PS-Import)
- `pstoedit` (Umwandeln in andere Vektorformate) und dann Vektorzeichenprogramme verwenden



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

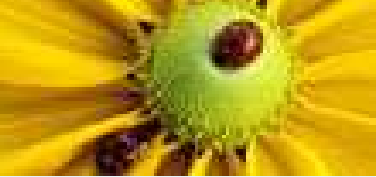
- Drucken
- Druck-Werkzeuge
- Verbreitete Probleme beim Drucken
- Lösungsansätze

Rechtliches

Spezialthemen

Ende.

PDF drucken und Troubleshooting

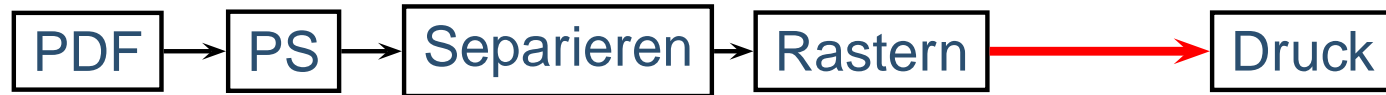


Drucken



Was passiert wo?

1. Host: Druckstrom vorrastern, in einer druckerspez. Sprache an den Drucker schicken (von Ghostscript gerastert, GDI)



2. Host: Wandlung nach PostScript plus PPD, dann PS (oder PCL) an den Drucker schicken



3. Direkt: PDF zum Drucker auf neueren PS 3 Druckern (kostenpflichtiger Zusatz); Druckeinstellungen über externen Kanal, oder über LPR-Queues (wird drucker-intern zuerst in PS gewandelt und dann gerastert, i.a. kein direktes Rendering)



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

● Drucken

● Druck-Werkzeuge

● Verbreitete Probleme beim
Drucken

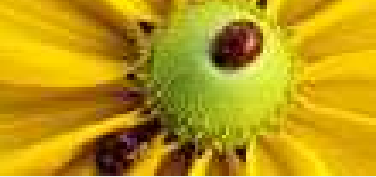
● Lösungsansätze

Rechtliches

Spezialthemen

Ende.

Druck-Werkzeuge



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

- Drucken
- Druck-Werkzeuge

- Verbreitete Probleme beim Drucken
- Lösungsansätze

Rechtliches

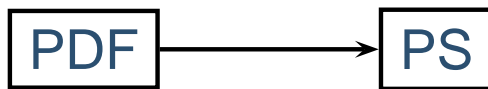
Spezialthemen

Ende.

Preprocessing: Ausschieß-Werkzeuge für Seitenauswahl, Broschürendruck wurden bereits genannt

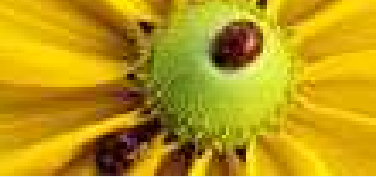
Wichtig: PDF enthält (idealerweise) keine drucker-abhängigen Einstellungen, Ausnahmen (Profi-Bereich):

- Job Ticket (hat nicht so richtig Fuß gefasst)
- OPI (Open Prepress Interface) (veraltet)
- Output-Intent, Rendering-Intent



- `pdftops (xpdf)`
- `acroread -toPostScript`
- `pdf2ps (gs)`

Dann mit PostScript-Viewer (`gv`, `ggv`, `kghostview`, `gsview`) prüfen.



Verbreitete Probleme beim Drucken

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

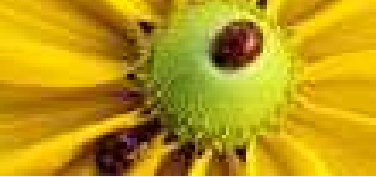
- Drucken
- Druck-Werkzeuge
- Verbreitete Probleme beim Drucken
- Lösungsansätze

Rechtliches

Spezialthemen

Ende.

- problematische Eingabe-Dateien durch mangelhafte Konverter (Komplexität wächst, z.B. Zerlegen von GIF-Bildern mit Transparenz in tausende Bitmap-Bildchen durch Distiller, 9 Seiten 20.000 Bilder)
- Verschlüsselung und Zugriffseinstellungen (Dokument lässt sich nicht richtig drucken, wegen Zugriffsbeschränkungen lässt es sich nicht ohne weiteres korrigieren)
- fehlende Schriften (nicht eingebettet) und Abbildungen (lo-res statt hi-res)
- Probleme mit Standardschriften (unterschiedliche Namen und Kodierungen)
- eingebettete Schriften fehlerhaft
- Profi-Druck: falscher Farbraum (RGB statt CMYK), Probleme bei Overprint/Trapping



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

- Drucken
- Druck-Werkzeuge
- Verbreitete Probleme beim Drucken
- **Lösungsansätze**

Rechtliches

Spezialthemen

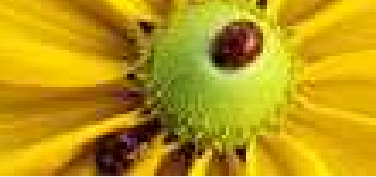
Ende.

„Preflighting“

- Analyse der Dateien vor dem Druck
- Reaktion: Ablehnen des Jobs, automatische Korrektur oder menschliche Eingriffe in den Druckprozess
- Selbst teure kommerzielle Tools rechnen sich im grafischen Gewerbe.
- Verbessertes grafisches Preview.
- Was machen wir an der Uni? (gs)

PDF/X:

standardisierter Workflow im grafischen Gewerbe (siehe auch Folie „PDF/X“)



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

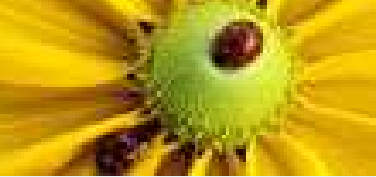
Rechtliches

- Zum Vergleich: Adobe Produkte
- Copyright von PDF
- Viele eigene Patente auf den Viewer
- Welche externe Software hat Adobe genutzt?
- Freie Software

Spezialthemen

Ende.

Rechtliches



Zum Vergleich: Adobe Produkte

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

- Zum Vergleich: Adobe Produkte
- Copyright von PDF
- Viele eigene Patente auf den Viewer
- Welche externe Software hat Adobe genutzt?
- Freie Software

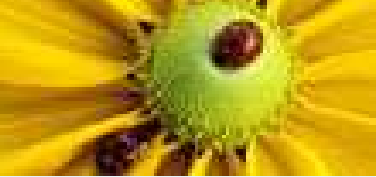
Spezialthemen

Ende.

Reader	kostenlos mit restriktiver Lizenz
Acrobat	verschiedene Versionen, Server-Einsatz nicht erlaubt
Acrobat Document Server	Server-Einsatz
AR Extension Server	Zusatzfunktionen für AR freischalten
Adobe Acrobat SDK	Bibliothek für Software-Entwicklung
Adobe Capture	OCR von gescannten Seiten

Gerade die Server Produkte sind sehr teuer, vor allem das Erzeugen von PDF mit Adobe-Software. Bei den Viewern hängt Adobe bei UNIX deutlich hinterher (MacOS und Win haben neueste Version).

Copyright von PDF



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

- Zum Vergleich: Adobe
Produkte

● Copyright von PDF

- Viele eigene Patente auf den
Viewer
- Welche externe Software hat
Adobe genutzt?
- Freie Software

Spezialthemen

Ende.

- Datenstrukturen und Operatoren sind (C) Adobe
- PDF spec, Abschnitt 1.4: Intellectual Property
- Jeder darf die Datenstrukturen nutzen, solange er sich an folgende Vorgaben hält:
 - ◆ Dokumentenzugriffsrechte müssen beachtet werden
 - ◆ Copyright-Notice von Adobe muss enthalten sein
 - ◆ Zugriff auf Dokumente im Widerspruch zu den Dokumentenzugriffsrechte gilt als Verstoß des Copyright

In der Vergangenheit war Adobe immer recht großzügig mit den Spezifikationen. Allerdings wurden schon öfter Teile zurückgehalten. Nun sind Teile nicht mehr (oder bewusst noch nicht) veröffentlicht (Reader Extensions).

Viele eigene Patente auf den Viewer

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

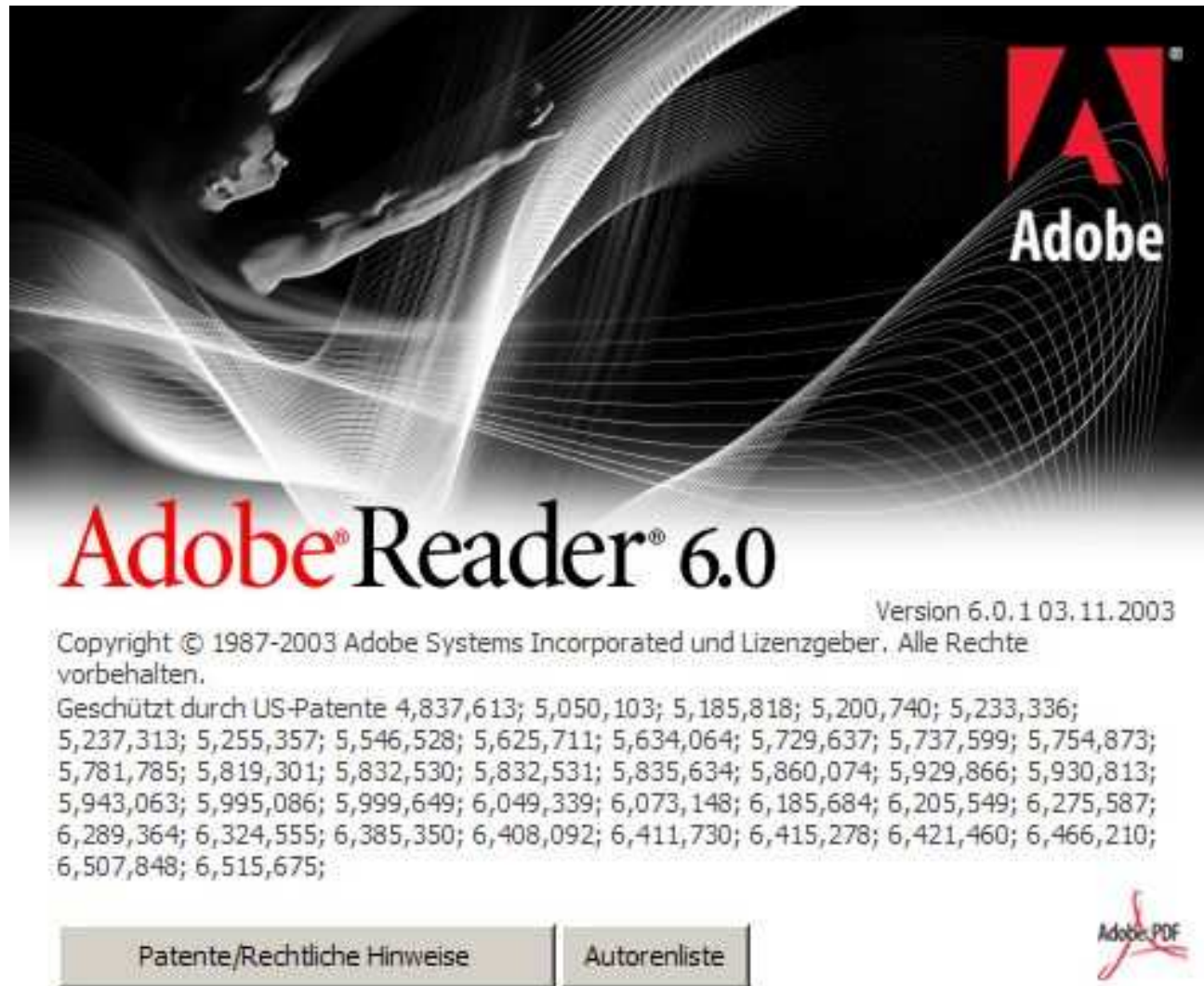
PDF drucken und
Troubleshooting

Rechtliches

- Zum Vergleich: Adobe Produkte
- Copyright von PDF
- Viele eigene Patente auf den Viewer
- Welche externe Software hat Adobe genutzt?
- Freie Software

Spezialthemen

Ende.



The image shows the splash screen for Adobe Reader 6.0. It features a black background with a white wireframe graphic of a person's head and neck. The Adobe logo is in the top right corner. The text 'Adobe Reader 6.0' is prominently displayed in the center. Below it, the version number 'Version 6.0.1 03.11.2003' is shown. A copyright notice follows: 'Copyright © 1987-2003 Adobe Systems Incorporated und Lizenzgeber. Alle Rechte vorbehalten.' This is followed by a list of US patents: 'Geschützt durch US-Patente 4,837,613; 5,050,103; 5,185,818; 5,200,740; 5,233,336; 5,237,313; 5,255,357; 5,546,528; 5,625,711; 5,634,064; 5,729,637; 5,737,599; 5,754,873; 5,781,785; 5,819,301; 5,832,530; 5,832,531; 5,835,634; 5,860,074; 5,929,866; 5,930,813; 5,943,063; 5,995,086; 5,999,649; 6,049,339; 6,073,148; 6,185,684; 6,205,549; 6,275,587; 6,289,364; 6,324,555; 6,385,350; 6,408,092; 6,411,730; 6,415,278; 6,421,460; 6,466,210; 6,507,848; 6,515,675;'. At the bottom, there are two buttons: 'Patente/Rechtliche Hinweise' and 'Autorenliste'. The Adobe PDF logo is in the bottom right corner.

Adobe Reader® 6.0

Version 6.0.1 03.11.2003

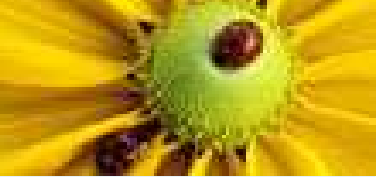
Copyright © 1987-2003 Adobe Systems Incorporated und Lizenzgeber. Alle Rechte vorbehalten.

Geschützt durch US-Patente 4,837,613; 5,050,103; 5,185,818; 5,200,740; 5,233,336; 5,237,313; 5,255,357; 5,546,528; 5,625,711; 5,634,064; 5,729,637; 5,737,599; 5,754,873; 5,781,785; 5,819,301; 5,832,530; 5,832,531; 5,835,634; 5,860,074; 5,929,866; 5,930,813; 5,943,063; 5,995,086; 5,999,649; 6,049,339; 6,073,148; 6,185,684; 6,205,549; 6,275,587; 6,289,364; 6,324,555; 6,385,350; 6,408,092; 6,411,730; 6,415,278; 6,421,460; 6,466,210; 6,507,848; 6,515,675;

Patente/Rechtliche Hinweise

Autorenliste

Adobe PDF



Welche externe Software hat Adobe genutzt?

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

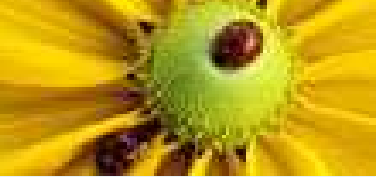
- Zum Vergleich: Adobe Produkte
- Copyright von PDF
- Viele eigene Patente auf den Viewer
- Welche externe Software hat Adobe genutzt?
- Freie Software

Spezialthemen

Ende.

Auswahl:

- Code aus einigen Open Source Projekten (libtiff, Apache, X Consortium, XFree)
- viele amerikanische Unis (es steht nicht genau da, welche Software)
- Berkeley (Sockets)
- Mach Kernel (wahrscheinlich für MacOS X port)
- Eric Young (Kryptographie, wahrscheinlich SSL, https, WebCapture)
- David Koblas (GIF Dekomprimierung, wahrscheinlich WebCapture)
- Netscape, Hewlett-Packard, RSA, Palm, Microsoft



Freie Software

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

- Zum Vergleich: Adobe Produkte
- Copyright von PDF
- Viele eigene Patente auf den Viewer
- Welche externe Software hat Adobe genutzt?
- Freie Software

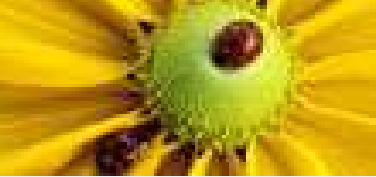
Spezialthemen

Ende.

- Eine sehr lange Liste steht auf:

<http://rnvs.informatik.tu-chemnitz.de/twiki/bin/view/Main/FreePdfUtilitiesAndLibraries>

- Was vor allem noch fehlt sind grafische Editierwerkzeuge und Werkzeuge zum Ausfüllen von PDF-Formularen.
- PDF-Konverter könnten besser werden.
- Offenbar ist vielen Entwicklern das Format nicht frei genug.



Spezialthemen

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

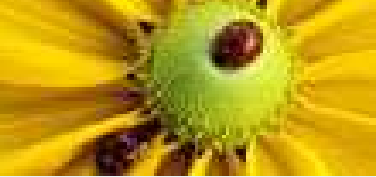
Rechtliches

Spezialthemen

- Sicherheit und DRM
- Verschlüsseln von
PDF-Dateien
- Von Adobe PDF abgeleitete
Formate
- PDF-Formulare

Ende.

Sicherheit und DRM



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

● Sicherheit und DRM

- Verschlüsseln von
PDF-Dateien
- Von Adobe PDF abgeleitete
Formate
- PDF-Formulare

Ende.

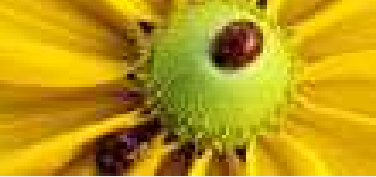
- es gibt mehrere Security Handler
- Standard Security
 - ◆ praktische Sicherheit gering
 - ◆ RC4, 40 Bits (nicht mehr zeitgemäß), jetzt bis 128 Bit
 - ◆ sinnvoll nur, wenn user password *und* owner password gesetzt sind
- Public-Key mit Zertifikaten (kaum eingesetzt)
- externe Firmen können weitere Verfahren bei Adobe anmelden
- E-Books (Adobe DRM): eigenes Verfahren, schon gebrochen (Elcomsoft), nicht veröffentlicht

Externe Methoden meist mit Wrappern um AR (Snake-Oil?):

<http://www.locklizard.com>,

<http://www.linkdata.com/index.htm#pdfcops>

Verschlüsseln von PDF-Dateien



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

- Sicherheit und DRM
- **Verschlüsseln von PDF-Dateien**
- Von Adobe PDF abgeleitete Formate
- PDF-Formulare

Ende.

■ Standard Security Handlers: 40 Bit (seit PDF 1.1) oder 128 Bit (seit PDF 1.4)

◆ Ghostscript:

```
http://casper.ghostscript.com/~ghostgum/pdftips.htm#encrypt
```

◆ ps2pdf14 -sOwnerPassword#owner

```
-sUserPassword#user -dEncryptionR#3
```

```
-dKeyLength#128 -dPermissions#-3904 in.ps
```

```
out.pdf
```

Permissions: Print, Extract text and graphics, ...

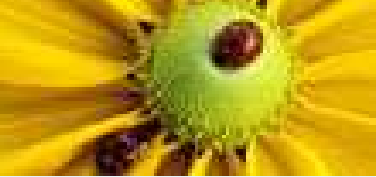
◆ Multivalent `tool.pdf.Encrypt`

◆ `pdftk encrypt_40bit encrypt_128bit allow`

■ Password Recovery:

◆ <http://www.elcomsoft.com/prs.html>

◆ RC4 Challenge



Von Adobe PDF abgeleitete Formate

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

- Sicherheit und DRM
- Verschlüsseln von
PDF-Dateien
- Von Adobe PDF abgeleitete
Formate
- PDF-Formulare

Ende.

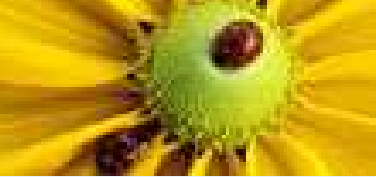
Spezielle „stripped-down“ PDFs:

■ PDF/X - PDF Exchange

- ◆ Untermenge von Adobe PDF, speziell für den zuverlässigen Datenaustausch in der Druckvorstufe
- ◆ ISO 15930 „Graphic technology – Prepress digital data exchange“
- ◆ Verleger fordern zunehmend PDF/X, v. a. von Anzeigenkunden
- ◆ Details: http://www.pdf-x.com/pdfx_123_1.php

■ PDF/A - PDF for Archival Storage

- ◆ Ziel: ISO-Standard
- ◆ Details: http://www.aiim.org/pdf_a/



PDF-Formulare

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

- Sicherheit und DRM
- Verschlüsseln von
PDF-Dateien
- Von Adobe PDF abgeleitete
Formate
- PDF-Formulare

Ende.

Erstellen

- OpenOffice (etwas verstecktes Menü mit den Formular-Elementen)
- pdfLaTeX
- grafisch: Scribus (Entwicklung)

Ausfüllen

Problem:

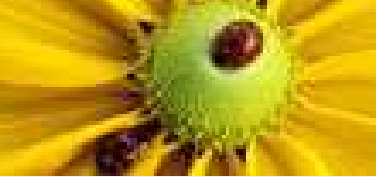
`acroread` arbeitet nicht bei `LANG=de_DE.UTF-8`.

Workaround:

Deshalb wird in `/usr/bin/acroread` auf `LANG=C` umgeschaltet. Folge: Umlaute lassen sich nicht eingeben.

`pdftk fill_form` – Formulare automatisiert ausfüllen

`pdftk flatten` – Formulardaten mit Hintergrund verbinden



- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

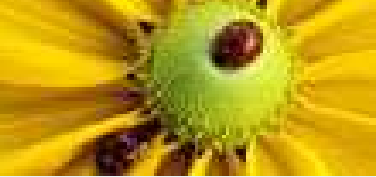
Rechtliches

Spezialthemen

Ende.

- Fragen?
- Diese Präsentation

Ende.



Fragen?

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

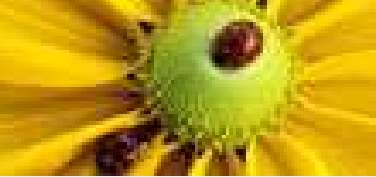
Spezialthemen

Ende.

● Fragen?

● Diese Präsentation





Diese Präsentation

- Inhalt
- Quellen

Was ist PDF?

PDF-Struktur

PDF erstellen

PDF analysieren

PDF bearbeiten

PDF drucken und
Troubleshooting

Rechtliches

Spezialthemen

Ende.

- Fragen?
- Diese Präsentation

- \LaTeX , dvips, ps2pdf
 - ◆ `latex slides.tex`
 - ◆ `dvips -o slides.ps slides.dvi`
 - ◆ `ps2pdf slides.ps`
- `documentclass: prosper.cls`
- `zusätzlich: HA-prosper.sty`
- `Stil: HAPHA.sty` (kommt mit HA-prosper)
- `Grafiken: pstricks, pst-node`