

Creating Usable PDFs: Effective and Ineffective Acrobat Practices

Shlomo Perets,  MicroType

1

Creating Usable PDFs

- Introduction
- Print PDFs
- Online PDFs
- Effective use of interactive features
- Testing your PDFs
- Evaluation checklist
- PDFs instead of printed books?

2

Introduction



3

PDFs Tested

- Over 300 from web sites
 - Direct access: including Adobe, Apple, Boeing, Hewlett-Packard, IBM, Intel, LSI Logic, Mentor Graphics, Motorola, National Semiconductor, Sun, Xerox
- Over 400 through searchpdf.adobe.com & general-purpose search engines
- Over 200 from CDs (document collections, software, books)

4

Most PDFs Authored in

- FrameMaker
- Word
- PageMaker
- Quark

5

Many PDFs are Unacceptable

- Quality of type & graphics
- No interactivity
- Partial interactivity and/or bad links
- Not or only partially searchable
- No bookmarks or inefficiently used
- Excessively large file sizes / unoptimized

6

State of the PDF

- PDF technology supports many unused features
- Those features that are used are often employed inefficiently
- There is poor or no quality control

7

Print PDFs



8

General

- Patterns & hatches
- Lines
- Error messages
 - Font-related problems
 - File corruption
 - Index attached for CD distribution, but is unavailable for download
- Page labels

9

Type

- Type 3 fonts
- TrueType issues
- Multiple Master fonts
- Embedding
- Subsetting
- Font substitution quality

10

Graphics

- Clarity
- Banding
- Excessive downsampling
- Resolution of graphics & image quality at various magnifications
- Consistency of quality (eg reasonable at 200% zoom)
- Rotated pages

11

Printing

- Garbage characters
- Missing fonts
- No output at all
- Errors when printing specific pages to non-Adobe PS printers
- Color distortions

12

Excessive PDF File Sizes

- Can be the result of:
 - Redundant items stored in the PDF
 - Unoptimized PDFs
 - ASCII PDFs (10-20% larger)
 - Numerous named destinations
 - Graphics not handled efficiently by Distiller, PDFWriter
 - Bitmap based PDFs
 - Inefficient PDF creators

13

Online PDFs



14

Design

- Page orientation
- Type
- Fonts
- Tables
- Graphics
- Blank pages (duplex printing?)
- Cropping

15

Display

- Hidden objects / items
- Complex vector objects
- Render time for various pages
 - complex vector objects
 - high-resolution bitmaps
 - fonts
- Greeking of hidden text

16

Opening Mode

- With bookmarks or thumbnails
- Open on relevant page
- Page layout: continuous / single page

17

Document Title

- Displayed in Search Results window
- If document has a “resize to initial page window” view setting:
 - Displayed in Window title
 - Displayed in Window menu
- Spoken by the text-to-speech engine

18

Single PDF vs. Multiple PDFs ①

- Entire book as a single PDF
 - Easier to print/download
 - Somewhat smaller PDF size (especially if many fonts are used & embedded)
 - Might be affected by Acrobat Catalog limitations

19

Single PDF vs. Multiple PDFs ②

- Multiple chapters
 - Search (with meaningful document titles)
 - Add links/bookmarks to contents / index
 - Use multiple windows

20

Find / Search

- TrueType
- Type 3
- Driver-related issues
 - No spaces between letters
 - Missing letters
 - Meaningless text
- Interletter spacing might affect Find and Search functions

21

Acrobat 3 & TrueType

- Text searches do not work on text which uses TrueType fonts
- Copy/paste produces gibberish text
- Potential problems with other Acrobat operations (eg printing with Unix Readers)

22

Acrobat 4 & TrueType

- Windows 95/98/NT:

Works OK if you use Distiller 4.0x
– and –
AdobePS 4.3.1/5.1.2 printer drivers,
– and –
only if PDF is displayed in Reader 4.0x

23

Type 3 Fonts ①

- AdobePS drivers +ATM may result in Type 3
- Can also be the result of custom PostScript fonts with file names not conforming to the “8.3” format
- Are larger than similar Type 1 or TrueType fonts

24

Type 3 Fonts ②

- Might not be searchable
- Can cause missing characters
- Take longer to display or output
- Look bad in low resolution & small sizes

25

Synthetic Bold

- With non-AdobePS drivers – attribute not visible in the PDF file
- With AdobePS drivers – the same text is positioned two or four times – not recommended for online documents

26

Acrobat's Text Recognition

- Garbled/missing characters
- Extra spaces
- "Word Finder"
- Hyphenation
- Output ordering factors

27

Effective Use of Interactive Features



28

Navigational Devices ①

- TOC
 - Short titles; simple & uncluttered layout
 - Appears in one place only
- Index
 - Two-level hierarchy (three-level for a complex product)
 - Select main entry for a topic
 - Subdivide entries when lengthy

29

Navigational Devices ②

- List of Figures, Tables, Equations
- Hierarchical maps
- Bookmarks: document structure is continuously displayed

30

Links ①

- None
- Sizing
- Bad
- Colors

31

Links ②

- Duplicate
- Distribution
- Wording
- Non-rectangular links

32

Link Problems ①

- Links which do nothing
- Links which point to non-existing files (file name changed before placed on web, merged)
- Links which point to non-existing destinations in external files
- Links which are duplicates of other links

33

Link Problems ②

- Links which have inconsistent / too small / offset active area
- Missing links
- Unintentional links
- Limitations of automatic web-links
- HTML to PDF Links

34

Bookmark Problems ①

- Missing bookmarks
- Too many bookmarks
- Wrong hierarchy
- Expanded hierarchy
- Order not useful
- Non-unique beginning

35

Bookmark Problems ②

- No direct bookmarks to Contents, Index, List of Figures, List of Tables
- No bookmarks to letter headings in lengthy indexes
- Paragraph tag appears
- Paragraph numbering
- All caps

36

Bookmarks – Additional Uses

- Can have creative uses
 - Link to web site, e-mail addresses
 - Launch external text/source files
 - Multimedia
 - As menus
- Top bookmark can identify publication
- Can have spacers (empty bookmarks)

37

Thumbnails

- **Effective in PDFs with illustrations and tables**

38

Articles

- **Required?**
 - **Hiding navigational buttons or important items**
 - **2-up pages**
 - **Multiple columns**
- **Articles in multiple columns**
 - **Must relate to columns if you have more than one column**

39

Notes

- **“Readme” notes with informal explanation, printing instructions**
- **Layered notes covering sections of a page, disclosing information when minimized**

40

Form Fields

- **Popups**
- **Non-printing buttons**
- **Watermarks printed but not displayed**

41

Video / Sound

- Better learning
- Reach non-verbal learners
- Complex subjects
- Illustrate processes

42

How-To Section

- Show your audience how to use the documentation
- Specific issues relating to the document collection
- Sample search queries
- Pointers to additional info

43

Testing your PDFs



44

Things Can Go Wrong... ①

- Don't use development station
- Display test
 - View your PDF using Reader and Acrobat 4 (Acrobat 3?)
 - Standard installations, minimum fonts, no plug-ins
 - Relevant operating systems

45

Things Can Go Wrong... ②

■ Print test

- Test on several types of printers:
Adobe PostScript, Non-Adobe PostScript compatible, PCL (non-PostScript), video drivers
- Color – if applicable
- TrueType / Multiple Master fonts
- Placed PDFs?

46

Local Viewer Preferences ①

- Smooth text and monochrome images (anti-aliasing)
- Default zoom
- Max “fit visible” magnification
- Substitution fonts (effective after restarting computer)

47

Local Viewer Preferences ②

- Display images larger than 128K (a gray box appears if this option is turned off)
- Use calibrated color for display
- Greek text below __ pixels
- Open cross-file links in same window

48

Different PDF Destinations

- Depending on the purpose of the PDF, you will require different configurations:
- CD-ROM (compact, yet print decently)
- Web-based viewing (compact, fast)
- High-end 4-color press

49

Testing Interactive Features ①

- **Checking interactivity manually is not practical**
 - In files with numerous links & bookmarks
 - Because not all bad links result in error messages
 - Frequent releases mean repeated testing

50

Testing Interactive Features ②

- **Link checking utilities (interactive/batch)**
 - Ari's Link Checker (www.dionis.com)
 - Auditor, part of Compose (www.ambia.com)
- **Link checkers do not check all aspects (wrong actions, link area, missing links, duplicates)**

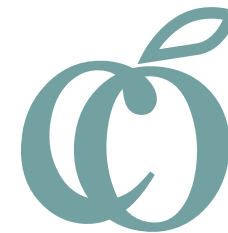
51

Testing Interactive Features ③

- **Need to check visual attributes of links**
 - font
 - consistent use of color
 - border
 - links in hypergraphics
- **When testing, always use/simulate end-user environment**
- **Verify web links periodically**

52

Evaluation Checklist



53

Document ①

- Layout, line- and page-breaks
- Overlooked/missing properties
- Page count; empty trailing pages?
- Orientation (rotation) & dimensions of each page in the PDF file
- Order of pages
- Blank pages (duplex printing?)
- Cropping: is it possible to eliminate excess white space (if applicable)?

54

Document ②

- Page labels (Acrobat 4)
- File size: is it reasonable?
- Document information Fields (Title, Subject, Keywords)
- Security (specific actions, general access)
- Optimization
- Find/Search: Can we select/locate text (test different fonts, in different pages)?

55

Display ①

- Initial settings
 - Initial page to be displayed
 - Bookmarks, thumbnails, full screen
 - Zoom
- Render time for various pages
- Approximate per-page download time
- Default view settings

56

Display ②

- Are the default page view and links to specific page views appropriate for different monitor sizes and resolutions?
- Appropriate when Acrobat is used as a browser plug-in?

57

Fonts

- **Is the text legible? (check for Type 3 fonts, especially in low resolutions)**
- **Font embedding and subsetting**
- **Substituted fonts and substitution quality**
- **Subsetting (affects editing; unique font names)**
- **Synthetic styles (bold, italic, outline)**

58

Graphics

- **Image quality at various magnifications**
- **Excessive downsampling?**
- **Banding?**
- **Clarity?**
- **Consistency of graphics quality throughout**

59

Colors

- **Similarity to color in original files**
- **Text**
- **Graphics**
- **Color spaces**
- **Are colors acceptable when viewed on monitors of different color depths or across platforms?**

60

Printing

- **Quality (text/graphics): clear/legible?**
- **Output devices**
 - **Test on major types of printers**
 - **Different resolutions**
 - **Color output (if applicable)**
- **Buttons/icons: are buttons printed?**
- **Unix: Multiple Master, TrueType fonts**

61

Interaction and Navigation ①

- Position and location of links: sizing & offset
- Functionality of links and buttons (within the current PDF and external files)
- Table(s) of contents and other lists: Hyperactive? Easily accessible?
- Redundant links?
- Index

62

Interaction and Navigation ②

- **Bookmarks**
 - Content/paragraphs included
 - Hierarchy
 - Order of items
 - Initial state (collapsed/expanded)
 - Paragraph numbers?
- **Thumbnails**

63

Interaction and Navigation ③

- **Articles**
 - Present?
 - Necessary?
 - Order of connection

64

Interaction and Navigation ④

- Visual attributes of links: color, border
- Base URL
- Pop-ups
- Is directory structure maintained when delivering the document collection to the Web server or a CD-ROM?
- PDX (cross-document, full-text search): Updated after all PDFs are finalized

65

PDFs Instead of Printed Books?



66

Publisher's View:

- Cheaper
- Faster to produce
- Frequent releases

67

User's View:

- If printed in full – less compact and useful
- Lower quality (type, graphics)

68

PDFs Instead of Printed Books

- For PDFs to effectively replace printed manuals, you must provide added value!
 - Design / Layout
 - Hyperlinked Contents, Index, List of Figures, List of Tables
 - Find / cross-document search
 - Useful bookmarks
 - Where applicable: thumbnails, notes and pop-ups, video and sound

69

Interactive Documentation

- High-quality online documentation does not just “happen” without a vision
- Includes aspects that are not part of printed documentation (eg navigation)
- Must develop quality-control procedures specifically for print and online documents

70

Whose Comfort?

Author / Editor / Producer

or

Reader / End-User

71

Ideally...

- The conversion process should be transparent
- Applications should natively support all relevant PDF features

72

PDF = Final Format!

- PDF is essentially a non-revisable format
 - Paragraphs converted to lines, single line might be converted to multiple segments
 - No document structure and flow
 - Subsetting disables even minute changes
 - Some Acrobat/Exchange editing operations might have unexpected results

73

Using Acrobat

- Features added in Acrobat “disappear” when you redistill (after changing source files)
- It is not practical to re-insert features every time the source file is redistilled
- Quality control considerations

74

PDF Production Line

- Rely on automatically-generated data as much as possible
- Add other features to your source files using Acrobat commands:
Create once, distill again and again and again...