PDFlib GmbH München, Germany www.pdflib.com



Copyright © 2005-2008 PDFlib GmbH. All rights reserved.

PDFlib GmbH Franziska-Bilek-Weg 9, 80339 München, Germany www.pdflib.com

This publication and the information herein is furnished as is, is subject to change without notice, and should not be construed as a commitment by PDFlib GmbH. PDFlib GmbH assumes no responsibility or liability for any errors or inaccuracies, makes no warranty of any kind (express, implied or statutory) with respect to this publication, and expressly disclaims any and all warranties of merchantability, fitness for particular purposes and noninfringement of third party rights.

Adobe, Acrobat, and PostScript are trademarks of Adobe Systems Inc. AIX, IBM, OS/390, WebSphere, iSeries, and zSeries are trademarks of International Business Machines Corporation. ActiveX, Microsoft, Windows, and Windows NT are trademarks of Microsoft Corporation. Apple, Macintosh and TrueType are trademarks of Apple Computer, Inc. Unicode and the Unicode logo are trademarks of Unicode, Inc. Unix is a trademark of The Open Group. Java and Solaris are a trademark of Sun Microsystems, Inc. Other company product and service names may be trademarks or service marks of others.

Thank you for using PDFlib FontReporter, a free Acrobat plugin provided by PDFlib GmbH. PDFlib GmbH offers software for creating and processing PDF documents. Please visit our Web site to learn more about our products.

You can use PDFlib FontReporter free of charge; however, it is not in the public domain. This software cannot be sold or redistributed (whether for a fee or at no charge), either stand-alone or in combination with any other product, without the express written permission of PDFlib GmbH.

Although PDFlib FontReporter is not a commercial product, we strive to provide high quality. If you run into problems you are encouraged to contact us at support@pdflib.com.



Contents

- 1 Installing PDFlib FontReporter 5
- 2 Working with FontReporter 7
- 2.1 What can you do with FontReporter? 7
- 2.2 Overview of PDF Font Formats 9
- 2.3 Contents of a Font Report 11
- 2.4 Investigate PDF Problems with FontReporter 14
- 2.5 Error Messages 15

A Revision History 17

1 Installing PDFlib FontReporter

Requirements. The PDFlib FontReporter plugin works with Acrobat 6/7/8 Standard and Professional on Windows and Mac, and Acrobat 9 Standard, Pro and Pro Extended on Windows. The plugin doesn't work with Acrobat Elements or any version of Acrobat Reader/Adobe Reader.

Installing FontReporter on Windows. All plugin-related files must be copied to the subdirectory »PDFlib FontReporter« in the Acrobat plugin folder. This is done automatically by the plugin installer, but can also be done manually. A typical location of the plugin folder is as follows:

C:\Program Files\Adobe\Acrobat 9.0\Acrobat\plug_ins\PDFlib FontReporter

Installing FontReporter for Acrobat 6/7/8 on the Mac. With Acrobat 6/7/8 the plugin folder is not visible in the finder. Make sure that Acrobat is not running and follow these steps:

- ► Extract the plugin files by double-clicking the disk image (.dmg).
- ► Locate the Acrobat application icon in the finder. It is usually located in a folder which has a name similar to the following:

/Applications/Adobe Acrobat 8.0 Professional

- ► Single-click on the Acrobat application icon and select File, Get Info.
- ▶ In the window that pops up click the triangle next to Plug-ins.
- ► Click Add... and select the FontReporter folder from the folder which has been created in the first step. Note that this folder will not immediately show up in the list of plugins, but only when you open the info window next time.

Multi-lingual Interface. FontReporter supports multiple languages in the user interface and generated font reports. Depending on the application language of Acrobat, FontReporter will choose its interface language automatically. Currently English and German interfaces are available. If Acrobat runs in any other language mode, Font-Reporter will use the English interface.

Trouble-shooting. If the FontReporter plugin doesn't seem to work, make sure that in *Edit, Preferences, [General...], Startup* the »Use only certified plug-ins« box is unchecked.

2 Working with FontReporter

2.1 What can you do with FontReporter?

FontReporter is a useful tool if you are interested in fonts within PDF documents. It provides font- and encoding-related information which will helps in a variety of situations:

- ► analyze printing problems (e.g. a particular font causes printing errors)
- ► investigate text extraction problems (e.g. copying text from a PDF results in garbage)
- ▶ visualize Unicode mappings for a font
- ► find flaws in the PDF creation workflow (e.g. printer driver converted a PostScript Type 1 font to Type 3)
- ► test whether ToUnicode mapping tables (required for PDF/A compliance) are present
- ► identify logos and symbols which are represented as text in a PDF
- ► learn which fonts are contained in a PDF, and which glyphs they contain (e.g. the file size is too large because some fonts ended up in the PDF unintentionally)
- ► check font subsets to see which glyphs are contained in the subset
- ► learn more about PDF font technology

Using FontReporter is as easy as bringing up the menu *Plug-Ins, PDFlib FontReporter..., Create Font Report* in Acrobat. This will create a font report for all pages of the current PDF document as a separate PDF. Two pages from typical font reports are shown in Figure 2.1.

Fig. 2.1 Sample font reports

PDFIIb FontReporter	PIAPOC+ThesisAntiqua-Normal, Type1, Embedded, Subset Encoding: custom, symbol bit, ToUnicode table, CharSet table											*DFIID FontReporter	Shir End Cha	oding aracte	: Ider r colle	m, CII ntity-H ection glyph	l : Adol			2													
W,		x1	x2	хЗ	x4	x5	х6	x7	x8	х9	xΑ	хВ	хС	хD	хE	xF	W.		x1	x2	хЗ	x4	х5	х6	x7	x8	х9	хA	хВ	хС	хD	хE	xF
Ox																	Ox																
1x																	1x																
2x	E9-804 U=0020								() parantipe			eenna Unoted	hyphen unicoso	paried Unoxide	/	2x																
Зх		1	2	3		5	6	7	8		colon U+000A	personal on U-0229					3x																
4x		A	В	C	D	E	F	G	H	Ī	Ţ	K	L	M	N	0	4x														い		
5x	P		R	S	T	U	V	W									5x						が										
6x		a	b	C	d	e	f	g	h	İ		k	Ī	m	n 3-one	O 0+000F	6x			す						た					つ		7
7x	p	q	r	S	t	u	V	W	X	y	Z	0-000			0+0000		7x	で				に			の								
8x											U+OSFA						8x									H		හ්			ゃ		
9x	0-2016																9x					る							を			ア	1
Ax	0=2019											((Ax			ウ	I		オ							グ	ケ		\neg
Вх												>> >>					Вх		サ	ザ	シ	ジ	ス		セ		ソ		タ				ツ
Сх												0-coe#					Сх			テ	デ	٢			Ξ								
Dx																	Dx		フ	ブ							ポ	マ	111	4	×	Ŧ	
Ex																	Ex		ュ		3		ラ	リ	ル	レ							ン
Fx																	Fx																

Supported PDF and font formats. FontReporter supports all PDF versions up to PDF 1.8, the file format created by Acrobat 9. All font and encoding formats available in PDF are supported, as well as all types of embedded font data.

Advantages over Acrobat's font properties panel. All versions of Acrobat including Adobe Reader provide font information via *File, Document Properties..., Fonts.* However, Acrobat's font overview is limited in use; FontReporter provides the following advantages compared to Acrobat's font list:

- ► FontReporter provides much more information about each font
- ► FontReporter deals with CJK font names even on Western systems
- ► FontReporter provides glyph tables containing the glyphs of a font along with their widths, names, and Unicode values
- ► FontReporter presents the output as a PDF document so that you can save or print it
- ► FontReporter is guaranteed to process the full document, regardless of which pages have already been displayed in Acrobat

PDF text extraction with PDFlib TET. FontReporter is an auxiliary tool to our PDFlib Text Extraction Toolkit (TET). TET is software for extracting the text contents of PDF documents. It is available both as a standalone program and a programming library/component which can be integrated into existing software. TET extracts text from all kinds of PDF documents and normalizes the text to Unicode. FontReporter can be used to create Unicode mapping tables for PDF documents which do not contain enough information for extracting text, or which contain wrong Unicode mapping tables. Fully functional evaluation versions of TET are available for download from www.pdflib.com.

TET PDF IFilter. TET PDF IFilter extracts text and metadata from PDF documents and makes it available to search and retrieval software on Windows. This allows PDF documents to be searched on the local desktop, a corporate server, or the Web. TET PDF IFilter is based on the patented PDFlib Text Extraction Toolkit (TET). TET PDF IFilter is a robust implementation of Microsoft's IFilter indexing interface. It works with all search and retrieval products which support the IFilter interface, e.g. SharePoint and SQL Server. Fully functional evaluation versions of TET PDF IFilter are available for download from www.pdflib.com.

Free TET Plugin. The TET Plugin is a free companion to the FontReporter Plugin. It can be installed in Adobe Acrobat and allows interactive use of the Text Extraction Toolkit (TET) with any PDF document that is currently open in Acrobat. Using the TET plugin you can access TET's functionality and experiment with TET options. The TET plugin can freely be downloaded from <code>www.pdflib.com</code>.

A Revision History

Revision history of this manual

Date	Changes
July 3, 2008	► Minor changes for FontReporter 1.3 (new: support for Acrobat 9 on Windows)
March 26, 2007	► Minor changes for FontReporter 1.2 (new: support for Acrobat 8 on Mac OS X)
January 30, 2006	► Minor additions for FontReporter 1.1
February 14, 2005	► Initial version for FontReporter 1.0.0

Known problems in this version. We are currently aware of the following minor problems:

- ► In rare cases the glyphs of Type 3 fonts may appear too small or too large, or not on the baseline.
- ► Although PDF Producer entries with non-Latin characters will be displayed properly in the bookmarks, they will appear garbled in the overview page.
- ► Sometimes not all glyph names are shown for simple fonts with the predefined encodings *WinAnsi* or *MacRoman*.
- ► Glyph names for built-in encodings are not shown, nor those for custom encodings which are based on a font's built-in encoding.
- ► Unicode values will not be shown for CID fonts with standard CJK character collections and simple fonts with the predefined encodings *WinAnsi* or *MacRoman*. However, since these are fixed mappings no document-specific information is lost.