datasheet

PDFlib, PDFlib+PDI, PDFlib Personalization Server (PPS) 9.2

What is PDFlib?

PDFlib is the leading developer toolbox for generating and manipulating files in the Portable Document Format (PDF). PDFlib's main targets are dynamic PDF creation on a Web server or any other server system, and to implement »Save as PDF« in existing applications. You can use PDFlib to dynamically create PDF documents from database contents, similar to dynamic Web pages. PDFlib has proven itself in a wide range of other use cases as well. Application programmers need only decent graphics or print output experience to be able to use PDFlib quickly. Since PDFlib frees you from the technicalities of the PDF file format, you can focus on acquiring the data and arranging text, graphics, and images on the page.

The PDFlib product family is available in three different flavors: PDFlib, PDFlib+PDI (PDF Import), and the PDFlib Personalization Server (PPS) with the PDFlib Block Plugin for Adobe Acrobat.

PDFlib

PDFlib offers all functions required to generate PDF documents with text, graphics, images, and interactive elements such as annotations or bookmarks. Use PDFlib for the following and many other tasks:

► add »Save as PDF« capability to your application
► create PDF documents on a Web server in real time
► create database reports in PDF
► take advantage of advanced typography and full Unicode and encoding support for text output
► advanced color management functionality
► convert TIFF, JPEG, or other image formats as well as SVG graphics to PDF
► automatically format tables with all kinds of cell contents
► create PDF/X-1/3/4/5 documents for commercial printing
► create PDF/A-1/2/3 for archiving
► Create PDF/VT for transactional printing
► create Tagged PDF and PDF/UA-1 for accessibility

PDFlib+PDI (PDF Import)

PDFlib+PDI includes all PDFlib functions, plus the PDF Import Library (PDI). With PDI you can open existing PDF documents and incorporate some pages into the PDFlib output. Use PDFlib+PDI for all PDFlib tasks plus the following:
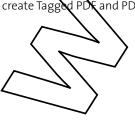
► impose multiple PDF pages on a single sheet for printing
► add text such as headers, footers, stamps, or page numbers to existing PDF pages
► place images, e.g. company logo, on existing pages
► add barcodes to existing PDF pages
► assemble existing PDF pages
► add content to PDF/A, PDF/X or PDF/UA documents

PDFlib Personalization Server (PPS) and PDFlib Block Plugin

The PDFlib Personalization Server (PPS) includes PDFlib+PDI plus additional functions for variable data processing using PDFlib Blocks. PPS makes applications independent from layout changes. The designer creates the page layout and converts it to PDF. She takes into account areas as placeholders for variable text and images. In Acrobat she drags a rectangular Block for each area using the PDFlib Block Plugin. Each Block contains a variety of Block properties, such as font size, color, image scaling. The PDFlib Block Plugin offers a Preview feature which shows the results of filling Blocks according to the specified properties.

The developer writes code to fill PDFlib Blocks with text, images, vector graphics or PDF pages. He doesn't need to know the formatting or position of a Block. Use PPS for all PDFlib+PDI tasks plus the following:

► customize direct mailings with text and images
► fill templates for transactional and statement processing
► personalize promotional material with address data
► generate individual parts catalogs from a database
► produce customized documentation for multiple products

What's new in PDFlib 9?

PDF/A-2 and PDF/A-3

PDFlib supports two additional parts of the PDF/A standard for archiving. PDF/A-2 is based on PDF 1.7 and supports transparency, JPEG 2000 compression, layers, and many other features. While PDF/A-2 allows embedding of PDF/A-1 and PDF/A-2 documents, PDF/A-3 allows embedding of arbitrary file types.

Tagged PDF and PDF/UA

Creating Tagged PDF is much easier through various convenience features, such as abbreviated tagging and automatic tagging of Artifacts. PDFlib's table formatter automatically tags tables. Tagged PDF documents including structure elements can be imported with PDI.

Accessible documents can be created according to the PDF/UA-1 standard (Universal Accessibility). PDF/UA is based on PDF 1.7 and improves Tagged PDF for accessibility similar to WCAG 2.0 (Web Content Accessibility Guidelines) in the Web world.

PDF/X

PDFlib supports PDF/X-1a, PDF/X-3 and PDF/X-4/4p as well as PDF/X-5n for exchange of n-colorant production files, e.g. in the packaging industry.

PDF/VT

PDF/VT is a standard for optimized PDF for variable and transactional printing. PDFlib creates output which conforms to PDF/VT-1, PDF/VT-2 or PDF/VT-2s according to ISO 16612-2 for Variable Document Publishing (VDP). Document Part Metadata (DPM) can be attached according to the PDF/VT standard.

Scalable Vector Graphics (SVG)

PDFlib imports vector graphics in the SVG format. SVG is the standard format for vector graphics on the Web, and is supported by all mainstream browsers. PDFlib 9.1 introduces support for ICC profiles, CMYK and spot color in SVG.

Font handling and text output

PDFlib's font engine and text processing have been enhanced in several ways:

▶ ideographic variation sequences (IVS) for CJK variant glyphs
▶ WOFF fonts (Web Open Font Format), a new container format for TrueType and OpenType fonts specified by the W3C
▶ support for all Unicode normalization forms (NFC, NFKC etc.)
▶ automatically create UPR font configuration files with all fonts found in an arbitrary number of directories

Import PDF documents with PDFlib+PDI

The following features are new in the PDF Import library PDI:

▶ Tagged PDF documents including structure elements can be imported.
▶ Layer definitions can be imported.

PDFlib Personalization Server (PPS) and Block Plugin

The following features are new in PPS:

▶ The Block type »Graphics« can be used to fill PDFlib Blocks with SVG graphics.
▶ PDFlib Blocks can be imported with PPS in addition to the Block Plugin.

Create PDFlib Blocks programmatically

In addition to creating PDFlib Blocks interactively with the PDFlib Block Plugin, PDFlib Blocks can be created programmatically with PPS. Existing PDFlib Blocks in imported documents can be copied to the generated PDF output. These features enable advanced document composition workflows where templates for PPS can themselves be built programmatically.

PDF Object Creation API (POCA)

POCA provides a set of methods for creating low-level PDF objects which are included in the generated PDF output. POCA can be used for the following purposes:

▶ create Document Part Metadata (DPM) for PDF/VT
▶ programmatically create PDFlib Blocks for use with PPS

Multimedia contents

PDFlib can create rich media annotations with Sound, Movie or 3D content. The multimedia content can be controlled with JavaScript and PDF actions. The following new multimedia features are available:

▶ rich media annotations
▶ rich media execute actions

Enhanced encryption algorithm

PDFlib supports PDF encryption according to Acrobat X/XI/DC. This encryption scheme is based on AES-256 and is specified in PDF 1.7 extension level 8 and PDF 2.0 according to ISO 32000-2.

What's new in PDFlib 9.1 and 9.2?

PDFlib/PDFlib+PDI/PPS 9.1 introduces new features related to color handling:

▶ Support for n-colorant color spaces (DeviceN and NChannel)
▶ PDF/X-5n for exchange of n-colorant production files, e.g. in the packaging industry
▶ SVG color extension for ICC profiles, spot and DeviceN color as well as Gray/RGB/CMYK device color
▶ PANTONE Extended Gamut Coated (XGC) spot colors and PANTONE Plus 2016 update
▶ color gradients with an arbitrary number of stop colors for enhanced color blends
▶ color gradients between different spot colors, e.g. a blend of PANTONE colors

PDFlib 9.2 introduces additional features:

▶ .NET Core language binding
▶ convenience features for PDF/UA and PDF/X creation
▶ improved import of Tagged PDF pages
▶ optimized TrueType subsetting with reduced file size
▶ reduced memory requirements for PDFlib Mini Edition (ME)

Common Features in PDFlib, PDFlib+PDI, and the PDFlib Personalization Server

PDF flavors	PDF 1.4 – PDF 1.7 extension level 8 and PDF 2.0 (Acrobat 5–DC)
	Linearized (web-optimized) PDF for byteserving over the Web
	High-volume output and arbitrary PDF file size (beyond 10 GB)
ISO standards for PDF	ISO 32 000-1: standardized version of PDF 1.7
	ISO 32 000-2: PDF 2.0 (including dated revision ISO 32000-2:2019)
	ISO 15 930: PDF/X-1/3/4/5 for the graphic arts industry
	ISO 19 005-1/2/3: PDF/A-1/2/3 for archiving
	ISO 16612-2: PDF/VT-1/2 for variable and transactional printing
	ISO 14289-1: PDF/UA-1 for universal accessibility
Fonts	TrueType (TTF and TTC) and PostScript Type 1 fonts
	OpenType fonts with PostScript or TrueType outlines (TTF, OTF, OTC)
	WOFF fonts (Web Open Font Format), a W3C-specified container format for fonts on the Web
	Support for dozens of OpenType layout features for Western and CJK text output, e.g. ligatures, small caps, old-style numerals, swash characters, simplified/traditional forms, vertical alternates
	Access fonts which are installed on Windows or macOS (»host fonts«)
	Font embedding for all font types; subsetting for TrueType, OpenType, and Type 3 fonts
	User-defined (Type 3) fonts for bitmap fonts or custom logos
	EUDC and SING fonts (glyphlets) for CJK Gaiji characters
	Fallback fonts (use missing glyphs from another font)
Text output	Text output in different fonts; underlined, overlined, and strikeout text
	Glyphs in a font can be addressed by numerical value, Unicode value, or glyph name
	Kerning for improved character spacing
	Artificial bold, italic, and shadow text
	Create text on a path
	Configurable replacement of missing glyphs
Accessibility	Create Tagged PDF for accessibility, page reflow, and improved content repurposing
	Tagging of interactive elements, e.g. annotations and form fields
	Automatic table and artifact tagging
	PDF/UA-1 for universal accessibility
	Additional structure element types and attributes
Internationalization	Full Unicode support for page content, interactive elements, and file names; support for all Unicode normalization forms
	Support for a variety of 8-bit and legacy multi-byte CJK encodings (e.g. Shift-JIS; Big5)
	CJK fonts and CMaps for Chinese, Japanese, and Korean text
	Ideographic variation sequences (IVS) for CJK variant glyphs
	Vertical writing mode for Chinese, Japanese, and Korean text
	Character shaping for complex scripts, e.g. Arabic, Thai, Devanagari
	Bidirectional text formatting for right-to-left scripts, e.g. Arabic and Hebrew
SVG vector graphics	Import vector graphics in SVG format; ICC profiles, CMYK and spot colors in SVG, CSS
Images	Load BMP, GIF, PNG, TIFF, JBIG2, JPEG, JPEG 2000, and CCITT raster images
	Query image information (pixel size, resolution, ICC profile, clipping path, etc.)
	Use clipping path in TIFF and JPEG images
	Use alpha channel (transparency) in TIFF and PNG images
	Image masks (transparent images with a color applied), colorize images with a spot or DeviceN color
Color	Grayscale, RGB (numerical, hexadecimal strings, HTML color names), CMYK, CIE L*a*b* color
	Integrated support for PANTONE® (incl. PANTONE+) and HKS® colors

	DeviceN (n-colorant) color space based on process or spot colors
	User-defined spot colors
	Color gradients between an arbitrary number of process colors or spot colors
Color management	ICC-based color with ICC profiles; support for ICC 4 profiles
	Rendering intent for text, graphics, and raster images
	ICC profiles as output intent for PDF/A and PDF/X; multi-colorant (xCLR) profiles for PDF/X-5n
Archiving	PDF/A-1a/1b, PDF/A-2a/b/u and PDF/A-3a/b/u
	XMP extension schemas for PDF/A
Graphic arts	PDF/X-1a, PDF/X-3, PDF/X-4, PDF/X-4p, PDF/X-5p, PDF/X-5pg, PDF/X-5n
	Embedded or externally referenced output intent ICC profile
	External graphical content (referenced pages) for PDF/X-5p and PDF/X-5pg
	Settings for overprint, text knockout, etc.
Variable Document Printing (VDP)	PDF/VT-1, PDF/VT-2, and PDF/VT-2s for variable and transactional printing
Textflow Formatting	Format text into one or more rectangular or arbitrarily shaped areas with hyphenation (user-supplied hyphenation points required), font and color changes, justification methods, tabs, leaders, control commands
	Advanced line-breaking with language-specific processing
	Flexible image placement and formatting
	Wrap text around images or image clipping paths
Table formatting	Table formatter places rows and columns, and automatically calculates their sizes according to a variety of user preferences. Tables can be split across multiple pages.
	Table cells can hold single- or multi-line text, images, vector graphics, PDF pages, path objects, annotations, and form fields
	Table cells can be formatted with ruling and shading options
	Flexible stamping function
	Matchbox concept for referencing the coordinates of placed images or other objects
Vector graphics	Common vector graphics primitives: lines, curves, arcs, ellipses, rectangles, etc.
	Smooth shadings (color blends) between multiple process or spot colors, pattern fills and strokes
	Transparency (opacity) and blend modes
	External graphical content (Reference XObjects) for variable document printing
	Reusable path objects and clipping paths imported from images
Layers	Optional page content which can selectively be displayed
	Annotations and form fields can be placed on layers
Security	Encrypt PDF document or attachments with 128/256-bit AES or RC4 128-bit encryption
	Unicode passwords
	Document permission settings (e.g. printing or copying not allowed)
Interactive elements	Create form fields with all field options and JavaScript
	Create barcode form fields
	Create actions for bookmarks, annotations, page open/close and other events
	Create bookmarks with a variety of options and controls
	Page transition effects, such as shades and mosaic
	Create all PDF annotation types, such as PDF links, launch links (other document types), Web links
	Named destinations for links, bookmarks, and document open action
	Create page labels (symbolic names for pages)
Multimedia	Embed 3D animations in PDF
	Embed Sound, Movie and 3D content in PDF and control it with JavaScript
	Navigators for custom presentation of PDF portfolios
Georeferenced PDF	Create PDF with geospatial reference information
Metadata	Document information: common fields (Title, Subject, Author, Keywords) and user-defined fields

	Create XMP metadata from document info fields or from client-supplied XMP streams
	User-supplied custom XMP metadata
	Process XMP image metadata in TIFF, JPEG, JPEG 2000 images and SVG graphics
Programming	Language bindings for C, C++, COM, Java, .NET and .NET Core, Objective-C, Perl, PHP, Python, RPG, Ruby,
	Virtual file system for supplying data in memory, e.g., images from a database
	Generate PDF documents on disk file or directly in memory (for Web servers)
Embedded Systems	PDFlib Mini Edition (ME) with reduced memory requirements

Additional Features in PDFlib+PDI and the PDFlib Personalization Server

PDF input (PDI)	Import pages from existing PDF documents
	Import all PDF versions up to PDF 1.7 extension level 8 (Acrobat X/XI/DC) and PDF 2.0
	Import documents which are encrypted with any of PDF's standard encryption algorithms
	Query information about imported pages
	Clone page geometry of imported pages (e.g. BleedBox, TrimBox, CropBox)
	Delete redundant objects (e.g. identical fonts) across multiple imported PDF documents
	Repair malformed input PDF documents
	Copy PDF/A or PDF/X output intent from imported PDF documents
	Import pages from Tagged PDF documents including structure elements
	Import layer definitions (optional content)
pCOS interface	pCOS interface for querying details about imported PDF documents

Additional Features in the PDFlib Personalization Server (PPS)

Variable Document Printing (VDP)	PDF personalization with PDFlib Blocks for text, image, PDF data, or SVG vector graphics
	Create PDFlib Blocks programmatically with PPS
	Copy PDFlib Blocks from imported documents1
PDFlib Block Plugin	PDFlib Block Plugin for creating PDFlib Blocks interactively in Acrobat on Windows and macOS
	Preview PPS Block filling in Acrobat
	Copy Blocks to Preview file
	Snap-to-grid for interactively creating or editing Blocks in Acrobat
	Clone PDF/X or PDF/A properties of the Block container
	Convert PDF form fields to PDFlib Blocks for automated filling
	Textflow Blocks can be linked so that one Block holds the overflow text of a previous Block
	PANTONE® and HKS® spot color names integrated in the Block plugin
	Support for Retina displays on macOS

Supported Development Environments

PDFlib is everywhere – it runs on practically all computing platforms. We offer 32- and 64-bit variants for all common flavors of Windows, macOS, Linux and Unix, as well as for IBM iSeries/i5 and zSeries mainframes.

The PDFlib core is written in highly optimized C code for maximum performance and small overhead. Via a simple API (Application Programming Interface) the PDFlib functionality is accessible from a variety of development environments:

► COM
► C and C++
► Java
► .NET and .NET Core
► Objective-C
► Perl
► PHP
► Python
► RPG (IBM i5/iSeries)
► Ruby

Benefits of using PDFlib Software

Rock-solid Products

Tens of thousands of programmers worldwide are working with our software. PDFlib products meet all quality and performance requirements for server deployment. All products are suitable for robust 24x7 server deployment and unattended batch processing.

Speed and Simplicity

PDFlib products are incredibly fast – up to thousands of pages per second. The programming interface is straightforward and easy to learn.

PDFlib Products all over the World

Our products support all international languages as well as Unicode. They are used by customers in all parts of the world.

Professional Support

If there's a problem, we will try to help. We offer commercial support to meet the requirements of your business-critical applications. By adding support you will have access to the latest versions, and have guaranteed response times should any problems arise.

Licensing

We offer various licensing programs for server licenses, integration and site licenses, and source code licenses. Support contracts for extended technical support with short response times and free updates are also available.

About PDFlib GmbH

PDFlib GmbH is completely focused on PDF technology. Customers worldwide use PDFlib products since 1997. The company closely follows development and market trends, such as ISO standards for PDF. PDFlib GmbH products are distributed all over the world with major markets in North America, Europe, and Japan.
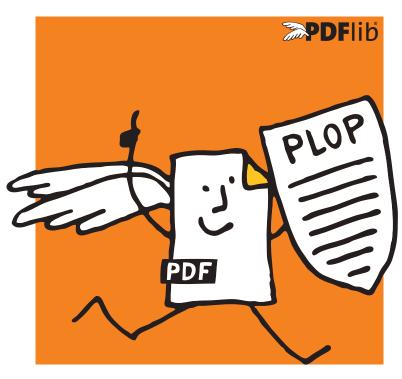
Contact

Fully functional evaluation versions including documentation and samples are available on our Web site. For more information please contact:

PDFlib GmbH
Franziska-Bilek-Weg 9, 80339 München, Germany
phone +49 • 89 • 452 33 84-0, fax +49 • 89 • 452 33 84-99
sales@pdflib.com
www.pdflib.com

PDFlib PLOP 5.3

PDF Linearization, Optimization, Protection

What is PDFlib PLOP?

PDFlib PLOP is a versatile tool for linearizing, optimizing, repairing, analyzing, encrypting and decrypting PDF documents. PLOP linearization and optimization features create efficient and small PDF documents for fast Web delivery. PLOP protection features encrypt or decrypt PDF documents and apply or remove permission settings, such as »printing not allowed« or »content extraction not allowed«. PLOP's repair mode automatically detects damaged PDF documents and fixes the problems if possible. PLOP analysis features can be used to query arbitrary properties of a PDF document. Document info entries and XMP metadata can be retrieved and set in a PDF/A and PDF/X conforming manner.

PDFlib PLOP Features

Linearization

With PDFlib PLOP you can linearize a PDF document for fast delivery over the Web (byteserving). Byteserving increases the perceived download speed since the first page is already visible while the remainder of the document is downloaded in the background.
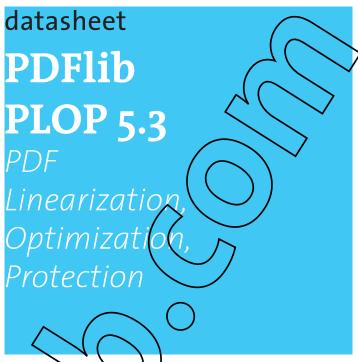
Optimization

PLOP can significantly reduce the file size of a PDF document without affecting quality. It achieves this by removing unnecessary or redundant identical objects, such as repeatedly embedded fonts, images, identical ICC color profiles, etc.

Password Security

PLOP can apply user and master passwords, and set permission restrictions to prevent the document from being printed with Acrobat, disallow text extraction or modification, etc.

PLOP supports strong AES-256 encryption and Unicode passwords. With PLOP you can:

► encrypt a PDF document with user or master password, or both;
► remove PDF encryption (if you know the master password);
► add or remove permission restrictions, e.g. »text extraction not allowed« (if you know the master password).

Certificate Security

PLOP also supports certificate security which works similar to e-mail encryption. PDF documents can be encrypted for a number of recipients, where each recipient is identified with a certificate (public key). Only legitimate recipients can open the document with the corresponding digital ID (private key). Since certificate security is an integral part of the PDF format specification it works in Acrobat and Adobe Reader without the need for any plugin or other custom software.

Certificate security offers several advantages over password security:

► no passwords must be distributed to the document recipients;
► individual permission restrictions can be specified for each recipient or groups of recipients;
► recipients cannot pass on document passwords to unauthorized third parties.

With PLOP you can encrypt PDF documents with certificates, apply permission restrictions, and decrypt documents which are protected with certificate security (provided you have a suitable digital ID).

The advantages of certificate security can be leveraged e.g. for securely distributing transaction documents or commercial content.

Repair Mode

PLOP detects various structural PDF problems and automatically repairs the document. PLOP can also fix certain XMP problems and optionally remove damaged XMP metadata.

PDF Analysis with pCOS

The PLOP library includes the pCOS programming interface for querying details about a PDF document, e.g.

► extract document info entries and XMP metadata
► list fonts and other resources
► query page sizes
► query security status, encryption scheme, permission settings
► many more PDF properties (see separate pCOS datasheet).

PLOP also includes the pCOS command-line tool.

XMP Metadata

Metadata (»data about data«) is an important topic in many areas of application software. XMP (Extensible Metadata Platform) is an XML-based framework with many predefined metadata properties. XMP is integrated in Acrobat/PDF, and much more powerful than simple document info entries. XMP is required for PDF/A and other ISO standards. Many industry groups have published XMP-based recommendations for vertical applications.

With PLOP you can insert XMP metadata in PDF documents and extract XMP from PDF. Inserted XMP is validated to ensure correct and standard-conforming output. If the input document conforms to PDF/A or PDF/X, PLOP ensures the XMP rules set forth in the respective standard.

XMP insertion with PLOP can be used in the following and other situations:

- ▶ Add XMP metadata to PDF/A documents, including support for XMP extension schemas as defined in the PDF/A standard.
- ▶ Add XMP metadata describing the scanning process for digitized legacy documents, also according to PDF/A.
- ▶ Add XMP metadata according to the Ghent Workgroup (GWG) Ad Ticket scheme.
- ▶ Add company-specific XMP metadata.

Document Info Entries

With PLOP you can add new document information entries or replace the values of existing info entries. Both predefined and custom entries can be set. If the input document contains XMP document metadata, all predefined info entries are synchronized to the XMP metadata in order to keep the metadata consistent (this is a requirement of PDF/A).

PDF Versions and Standards

PLOP supports all relevant PDF versions and standards:

- ▶ PLOP processes all PDF versions up to Acrobat DC, i.e PDF 1.7 (ISO 32000-1) up to extension level 8. PLOP can also process documents according to PDF 2.0 (ISO 32000-2).
- ▶ PLOP is aware of the PDF/A-1/2/3 (ISO 19005) archiving standards: if the input document conforms to PDF/A, the output document is guaranteed to conform as well. PLOP fully supports XMP extension schemas as required by PDF/A.
- ▶ Similarly, PLOP is aware of the PDF/X-1a/3/4/5 (ISO 15930) print production standards, PDF/VT-1/2 (ISO 16612-2) for variable and transactional printing and PDF/UA-1 (ISO 14289) for accessible PDF.

PLOP Library or Command-Line Tool?

PLOP is available as a programming library (component) for various development environments, and as a command-line tool for batch operations. Library and command-line tool offer similar features, but are suitable for different deployment tasks.

The PLOP programming library is used...

...for integration into desktop or server applications. Examples for using the library with all supported language bindings are included in the PLOP package. Since the PLOP library accepts PDF input documents from a disk file or directly in memory, it can easily be combined with other products.

The PLOP programming library also contains the pCOS interface for retrieving a variety of PDF document properties.

The PLOP command-line tool is suited...

...for batch processing PDF documents. It doesn't require any programming, but offers powerful command-line options which can be used to integrate it into complex workflows. The PLOP command-line tool can also be called from environments which do not support the use of the PLOP library.

Supported Development Environments

PDFlib PLOP is everywhere – it runs on practically all computing platforms. We offer 32-bit and 64-bit packages for all common flavors of Windows, macOS, Linux and Unix, as well as for IBM zSeries mainframe systems. Versions for iOS and Android are also available.

The PLOP core is written in highly optimized C and C++ for maximum performance and small overhead. Via a simple API (Application Programming Interface) the PLOP functionality is accessible from a variety of development environments:

- ▶ COM for use with VB, ASP, etc.
- ▶ C and C++
- ▶ Java
- ▶ .NET for use with C#, VB.NET, ASP.NET, etc.
- ▶ Objective-C
- ▶ Perl
- ▶ PHP
- ▶ Python
- ▶ Ruby

PLOP DS for digitally signing PDF

The extended version PLOP DS supports all features of PLOP plus the ability to apply digital signatures to PDF documents. Please refer to the separate PLOP DS datasheet for more information.

Benefits of using PDFlib Software

Rock-solid Products

Tens of thousands of programmers worldwide successfully use our software. PDFlib products meet all quality and performance requirements for server deployment. All products are suitable for robust 24x7 server deployment and unattended batch processing.

Speed and Simplicity

PDFlib products are incredibly fast – up to thousands of pages per second. The programming interface is straightforward and easy to learn.

PDFlib Products all over the World

Our products support all international languages as well as Unicode. They are used by customers in all parts of the world.

Professional Support

If there's a problem, we will try to help. We offer commercial support to meet the requirements of your business-critical applications. By adding support you will have access to the latest versions, and have guaranteed response times should any problems arise.

Licensing

We offer various licensing options for server licenses, integration and site licenses, and source code licenses. Support contracts for extended technical support with short response times and free updates are also available.
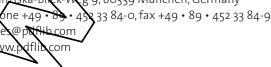
About PDFlib GmbH

PDFlib GmbH is completely focused on PDF technology. Customers worldwide use PDFlib products since 1997. The company closely follows development and market trends, such as ISO standards for PDF. PDFlib GmbH products are distributed all over the world with major markets in North America, Europe, and Japan.
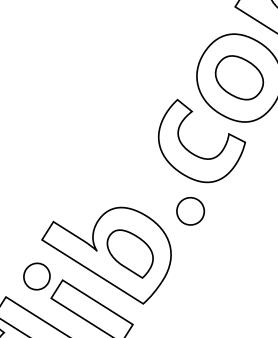
Contact

Fully functional evaluation versions including documentation and samples are available on our Web site. For more information please contact:
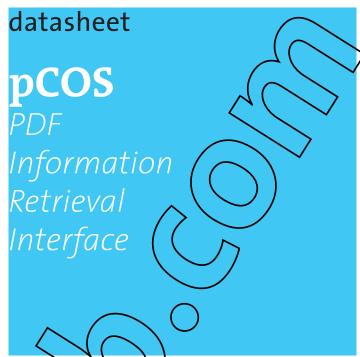
PDFlib GmbH
Franziska-Bilek-Weg 9, 80339 München, Germany
phone +49 • 89 • 452 33 84-0, fax +49 • 89 • 452 33 84-99
sales@pdflib.com
www.pdflib.com

pCOS
PDF Information Retrieval Interface

What is the pCOS Interface?

pCOS provides a simple and elegant facility for retrieving any information from a PDF document which is not part of the page contents (page contents an be extracted with PDFlib TET). The pCOS interface is not a standalone product, but an integrated part of the following products:

► PDFlib+PDI and PDFlib Personalization Server (PPS) (pCOS is not included in the base PDFlib product)
► Text and Image Extraction Toolkit (PDFlib TET)
► In PDFlib TET PDF IFilter you can use pCOS to retrieve information from PDF documents and use it for indexing and search.
► PDFlib PLOP and PLOP DS

These products include the pCOS programming interface. PLOP and PLOP DS packages also include the pCOS command-line tool.

With pCOS you can extract a variety of interesting items and create output for different purposes. By processing multiple PDF documents with a single call you can easily create summaries of document info entries, page formats, fonts, or any other property. Combined with tabular output this provides a powerful PDF administration tool.

There are many application scenarios for using pCOS within PDF workflows, but you can also use pCOS as a tool for learning or debugging PDF. Here are some typical situations:

► Check incoming documents for predefined criteria
► Identify problem files in a large collection
► Create metadata summaries for document management
► quality assurance before publishing documents
► document retrieval and repository workflows
► summarize the bookmarks
► extract components of PDF documents, e.g. ICC profiles
► Check PDFs for security problems (JavaScript etc.)

pCOS Cookbook

The pCOS Cookbook is a collection of programming examples which demonstrate the use of pCOS for various PDF retrieval tasks. The Cookbook is available on the PDFlib Web site and includes code, input documents and pCOS output.

pCOS Features

Information Retrieval

With pCOS you can extract a variety of interesting items, such as:

► document info entries and XMP metadata
► general information: linearization and tagged PDF status, encryption details and permission settings, number of pages and fonts
► fonts with name, embedding status, etc.
► image data, such as bit depth, color space, compression, XMP
► color space details
► target URLs and coordinates of Web links
► bookmarks and the corresponding page numbers, e.g. to create a table of contents
► form field data: full field names, contents, position, etc.
► page size, CropBox, page rotation
► status of ISO standards: PDF/A, PDF/E, PDF/UA, PDF/VCR, PDF/VT, and PDF/X
► geospatial reference information
► list or extract file attachments
► layer names, page labels, article threads
► annotation details
► list comments along with the reviewer's name
► digital signature details: name of signature field, signed/unsigned, name of signer, PAdES
► extract ICC output intent profiles from PDF/X or PDF/A documents
► Block properties for PDFlib Personalization Server
► JavaScript on document, page, annotation, or field level
► retrieve XML invoice data from ZUGFeRD documents
► properties of PDF Packages/Portfolios

Supported Input

pCOS supports all flavors of PDF input:

► all PDF versions up to Acrobat DC, i.e PDF 1.7 (ISO 32000-1) up to extension level 8, as well as PDF 2.0 (ISO 32000-2)
► encrypted documents (password may be required)
► damaged PDF input documents will be repaired if possible

Output Formats

pCOS can create output for different purposes:

- ▶ plain text output
- ▶ Unicode text output in UTF-8 or UTF-16 formats
- ▶ tabular output for processing with a spreadsheet or database
- ▶ binary data, e.g. ICC profiles or file attachments
- ▶ user-defined output formats for custom post-processing

pCOS Paths: Simple Syntax for PDF Objects

Instead of getting bogged down by complex tree structures, e.g. for bookmarks or form fields, you can easily access PDF objects by using the simple pCOS path syntax. It offers convenient shortcuts for accessing commonly used PDF objects, such as pages, fonts, bookmarks, form fields etc.

pCOS Library or Command-Line Tool?

pCOS is available as a programming library (component) for many development environments, and as a command-line tool for batch operations. Note that the pCOS command-line tool is included only in the PDFlib PLOP and PLOP DS product packages. Programming interface and command-line tool offer similar features, but are suitable for different deployment tasks.

The pCOS Programming Interface is used...

...for integration into desktop or server applications. Simple examples for using the library with all supported language bindings are included in the product packages. Many more examples are available in the pCOS Cookbook.
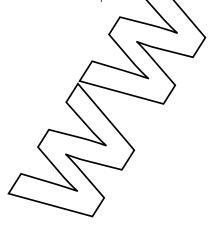
The pCOS Command-Line Tool is suited...

...for batch processing PDF documents. It doesn't require any programming, but offers powerful command-line options which can be used to integrate it into complex workflows. The pCOS command-line tool extends the features of the library.

- ▶ simple retrieval of common PDF elements, such as bookmarks, annotations, metadata, form fields, etc.
- ▶ extended mode for querying more complex objects and customizing the output format
- ▶ extract data items such as file attachments, ICC profiles, etc.
- ▶ emit information as comma-separated values or a user-defined format for import into a spreadsheet or database
- ▶ recursion feature for dumping composite PDF objects such as dictionaries and arrays

Evaluation

All products which contain the pCOS interface are available for evaluation. Fully functional evaluation versions including documentation and samples are available on our Web site.

Benefits of using PDFlib Software

Rock-solid Products

Tens of thousands of programmers worldwide are working with our software. PDFlib meets all quality and performance requirements for server deployment. All PDFlib products are suitable for robust 24x7 server deployment and unattended batch processing.

Speed and Simplicity

PDFlib products are incredibly fast – up to thousands of pages per second. The programming interface is straightforward and easy to learn.

PDFlib all over the World

Our products support all international languages as well as Unicode. They are used by customers in all parts of the world.

Professional Support

If there's a problem, we will try to help. We offer commercial support to meet the requirements of your business-critical applications. By adding support you will have access to the latest versions, and have guaranteed response times should any problems arise.

Licensing

We offer various licensing programs for server licenses, integration and site licenses, and source code licenses. Support contracts for extended technical support with short response times and free updates are also available.

About PDFlib GmbH

PDFlib GmbH is completely focused on PDF technology. Customers worldwide use PDFlib products since 1997. The company closely follows development and market trends, such as ISO standards for PDF. PDFlib GmbH products are distributed all over the world with major markets in North America, Europe, and Japan.

Contact

For more information please contact:

PDFlib

PDFlib GmbH
Franziska-Bilek-Weg 9, 80339 München, Germany
phone +49 • 89 • 452 33 84-0, fax +49 • 89 • 452 33 84-99
sales@pdflib.com
www.pdflib.com