





**What is PDFlib pCOS?**

PDFlib pCOS provides a simple and elegant facility for retrieving information from a PDF document which is not part of the page contents. For example, PDF metadata, reference streams, links etc. Long page dimensions can easily be queried with pCOS.

With pCOS you can extract a variety of interesting items and create output for alternative purposes by processing multiple PDF documents with a single call or you can simply create some items and documents into HTML page formats, titles or any other property. Combined with HTML output this provides a powerful PDF administration tool.

There are many every-day pCOS applications for PDF practitioners, but you can also use PDFlib pCOS as a tool for homepage debugging PDF. Here are some typical examples:

- Check opening documents for preflight errors
- Check PDFs for security problems and active content (Java, Scripts)
- Check documents for quality assurance before publication
- Identify problem files in a large collection
- Create property summaries for document management
- Learn details of PDF data structures

**PDFlib pCOS Features**

**Supported Input**

- PDFlib pCOS supports all relevant flavors of PDF input:
- All PDF versions up to PDF 1.7 (Acrobat 8)
- ICC and RGB embedded (embedded fonts are required)
- Sophisticated security models even if you don't know the password, you can query certain parts of information as long as this doesn't violate the document author's intentions
- Damaged PDF input documents will be repaired if possible

**Information Retrieval**

PDFlib pCOS offers a simple query interface, without the need for low-level device programming. With PDFlib pCOS you can extract a variety of interesting items, such as:

- Document information and XMP metadata

- General information (title, author and tagged PDF status, encryption details and permission settings, number of pages and links)
- All fonts with their name, embedding status, etc.
- Images with size, bit depth, color space, compression, etc.
- Color space details for all PDF color spaces
- Target fonts and coordinates of their links
- All bookmarks along with the corresponding page numbers, e.g. to create a table of contents
- Form field data: field names, contents, position, etc.
- Page size, CropBox, page rotation
- Tables of contents and tables of contents files
- List or extract file attachments
- List or extract file attachments
- Rotation details
- List all comments along with the document's name
- Digital signature details, name of signature field(s), signed/unsigned, name of signer, date and reason of signature
- Extract CC output intent profiles from PDF/A files
- List PDFlib block properties
- Inspect or document page, annotation, or field level

**Output Formats**

- PDFlib pCOS can create output for different purposes:
- Plain text output
- Tabular output for processing with a spreadsheet/databse
- Binary data files (e.g. ICC profiles or file attachments)
- Unicode text output in UTF-8 or UTF-16 formats
- Non-unicode output formats for custom post-processing

**pCOS Paths - Simple Queries for PDF Objects**

Instead of getting bogged down by complex tree structures, e.g. for bookmark or font fields, you can easily access PDF objects by using the simple pCOS path syntax. It offers common 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 various development environments, and as a command-line tool for batch operations. Both offer similar features, but are suitable for different deployment tasks.

**The pCOS 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 pCOS package. A variety of additional examples is available in the pCOS Cookbook on the PDFlib Web site.

**The pCOS command-line tool is used...**

- ...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.
- Link information as comma-separated values or user-defined format for import into a spreadsheet or database
- Recursive feature for generating complete PDF objects, such as dictionaries and arrays

**Supported Development Environments**

PDFlib pCOS is environment-independent, practically all computing platforms are supported for all common flavors of Windows, Mac OS X, Linux and Unix.

The pCOS code is written in highly optimized C code for maximum performance and thread overhead. As a sample API (Application Programming Interface) the pCOS functionality is available from a variety of development environments:

- COM for use with VB, ASP, and many other languages
- C and C++
- Java, including servlets and Java Application Server
- NET for use with C#, VB.NET, ASP.NET, etc.
- Perl
- PHP

**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.

**Benefits of using PDFlib Software**

**Rock-solid Products**

Over 100,000 developers worldwide are working with our software. PDFlib meets all quality and performance requirements for server deployment. All PDFlib products are suitable for robust 24/7 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 is 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.

**About PDFlib GmbH**

PDFlib GmbH is completely focused on PDF technology for software developers. Customer feedback on 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. Our service contacts:  
PDFlib GmbH  
Frankfurt-Bock Weg 5 • 61339 Munchen, Germany  
phone: +49 89 452 133 8-9  
fax: +49 89 452 13 84 99  
sales@pdflib.com  
www.pdflib.com

©1999-2008