



PDFlib pCOS PDF情報 抽出ツール



PDFlib pCOSとは

PDFlib pCOS は、PDF 文書内のページ内容以外のあらゆる情報を抽出するシンプルでエレガントな機能を提供します。たとえば、PDF メタデータやハイパーテキスト・ページサイズを、pCOS なら容易に取得できます。

pCOS を使えば、さまざまな興味深い項目を抽出し、いろいろな目的に合った出力を作成することが可能です。一回の呼び出しで複数の PDF 文書进行处理することにより、文書情報項目やページ規格・フォント等あらゆるプロパティのサマリーを簡単に作成することができます。これをタブ出力と組み合わせれば、強力な PDF 管理ツールとしての活用が可能です。

PDFlib pCOS には、さまざまな PDF 日常業務上の応用がある一方で、PDF の学習・デバッグ用ツールとしての使い方もできます。次のような典型的シナリオが挙げられます。

- ▶ 受け取った文書が既定の条件を満たしているかチェック
- ▶ 文書を発行前に品質検査
- ▶ 大量ファイル中の問題ファイルを発見
- ▶ プロパティサマリーを作成して文書管理
- ▶ PDF データ構造の詳細学習

PDFlibソフトウェア活用のさまざまな利点

ビクともしない製品です

何万人もの全世界のプログラマーが利用している弊社のソフトウェア。PDFlib は、サーバでの運用に要求されるすべての品質・速度要件を満たしています。すべての PDFlib 製品は、堅牢な 24 × 7 のサーバ運用にも、無人バッチ処理にも適しています。

スピード&シンプル

PDFlib 製品は驚異的に高速です。毎秒数千ページもの処理能力を誇ります。そのプログラミングインタフェースはシンプルで、習得が容易です。

全世界でPDFlib

弊社の製品は、世界のすべての言語と Unicode に対応しています。全世界のお客様にご活用いただいています。

プロフェッショナルサポート

問題が発生すれば、弊社は解決支援に努力します。お客様のビジネスクリティカルなアプリケーションの要求をみため、弊社では有償サポートを提供しています。サポートを追加することにより、最新のバージョンが入手でき、問題発生の際の回答所要時間が保証されます。

PDFlib pCOSの機能

対応入力形式

PDFlib pCOS は入力形式として、以下のすべての PDF 形式に対応しています。

- ▶ PDF 1.6 (Acrobat 7) までの全バージョンの PDF
- ▶ 40 ビット・128 ビット暗号化された PDF
- ▶ 洗練されたセキュリティモデル。パスワードがわからない場合でも、文書作成者の意図に反しないかぎりの情報を抽出可能

情報抽出

PDFlib pCOS は、シンプルな取得インタフェースを提供していますので、低レベルでのパーサプログラミングを行う必要がありません。PDFlib pCOS を使えば、次のようなさまざまな興味深い情報を抽出することができます。

- ▶ 一般情報。PDF の線形化・タグ付きの有無、暗号化の詳細・権限設定、ページ数、フォント
- ▶ 文書情報項目・XMP メタデータ
- ▶ 全フォントの名前・埋め込み状態等
- ▶ Web リンクのリンク先 URL・座標
- ▶ 全しおりとその各リンク先ページ番号。目次作成等が可能
- ▶ フォームフィールドデータ。フルフィールド名・内容・位置等
- ▶ ページサイズ・CropBox・ページ回転
- ▶ PDF/X 準拠ファイルの状態
- ▶ 抽出ファイル添付一覧
- ▶ レイヤー名
- ▶ 注釈の詳細
- ▶ 全コメントとその各校正者名
- ▶ 電子署名の詳細。署名フィールドの名前・署名済 / 未署名・署名者名・署名日・署名理由
- ▶ PDF/X・PDF/A ファイルから ICC 出力_intent プロファイルを抽出
- ▶ PDFlib ブロックプロパティの一覧
- ▶ 文書・ページ・注釈・フィールドレベルの JavaScript

出力形式

PDFlib pCOS は目的に応じ、次のさまざまな出力を生成できます。

- ▶ プレーンテキスト出力
- ▶ タブ出力。表計算ソフト・データベースで処理可能
- ▶ バイナリデータ。再利用可能。例：ICC プロファイル・ファイル添付
- ▶ Unicode テキスト出力。UTF-8・UTF-16 形式
- ▶ ユーザー定義出力形式。カスタム後処理が可能

pCOSパス - PDFオブジェクトのためのシンプルな文法

複雑なツリー構造にわずらわされる必要はありません。シンプルな pCOS パス文法で、しおりやフォームフィールド等、さまざまな PDF オブジェクトに到達できます。この文法は、ページ・フォント・しおり・フォームフィールド等、よく使われる PDF オブジェクトへの便利なショートカットを提供するものです。

pCOSライブラリかコマンドラインツールか

pCOS は、さまざまな開発環境におけるプログラミングライブラリ（コンポーネント）としても利用可能ですし、コマンドラインツールとしてバッチ操作にも利用可能です。どちらも提供する機能はほぼ同じですが、それぞれ異なる実用目的に適しています。

pCOSプログラミングライブラリの応用対象

デスクトップ・サーバアプリケーションへの統合に用いられます。pCOS パッケージには、すべての対応言語バインディングについて、ライブラリの使用例が入っています。

pCOSコマンドラインツールの応用対象

PDF 文書群のバッチ処理に用いられます。プログラミングが完全に不要でありながら、強力なコマンドラインオプションを提供していますので、複雑なワークフローにも統合可能です。pCOS コマンドラインツールは次のようにライブラリ機能を拡張しています。

- ▶ しおり・注釈・メタデータ・フォームフィールド等、よく使われる PDF 要素のシンプルな抽出
- ▶ より複雑なオブジェクトの取得と出力形式のカスタマイズのための拡張モード
- ▶ コンマ区切りまたはユーザー定義形式で情報を出力し、表計算ソフトやデータベースへの読み込みを可能に
- ▶ 辞書・配列等、複雑な PDF オブジェクトをダンプするための再帰機能

対応開発環境

どこでも PDFlib pCOS。事実上、すべてのコンピューティングプラットフォームで動作します。広く使われている Windows・Mac OS・Linux・Unix をはじめ、IBM sServer iSeries・zSeries メインフレームにも、弊社は PDFlib pCOS を提供しています。

pCOS のコアは、高度に最適化された C コードで書かれ、最高のパフォーマンスと小さなオーバーヘッドを実現しています。シンプルな API（アプリケーションプログラミングインタフェース）を用いて、pCOS の機能は次のさまざまな開発環境で利用可能です。

- ▶ COM。VB・ASP・Borland Delphi 等で利用
- ▶ C・C++
- ▶ Java。サーブレット・Java Application Server を含む
- ▶ .NET。C#・VB.NET・ASP.NET 等で利用
- ▶ Perl

ライセンス

サーバライセンス、統合・サイトライセンス、ソースコードライセンスのため、弊社ではさまざまなライセンスプログラムを提供しています。拡張された技術サポートが短い応答時間で無料アップデートとともに提供されるサポート契約も利用可能です。

PDFlib GmbHについて

PDFlib GmbH は PDF 技術に完全に特化しています。PDF 開発は、私たちが 1997 年から手がけてきたことであり、私たちの製品はすべて PDF に関連したものです。私たちには長年の経験があり、PDF に関することならどんなに細かなことでも明らかにする能力があります。私たちの製品は世界中に頒布されており、中でも大きな市場は北米・日本・欧州となっています。



連絡先

完全に機能する評価版を、説明書・サンプルとともに、すべての対応プラットフォーム版について、弊社 Web サイトで入手できます。さらに詳しくは下記宛お問い合わせ下さい。

PDFlib GmbH

Tal 40, 80331 München, Germany

phone +49・89・29 16 46 87

fax +49・89・29 16 46 86

sales@pdflib.com

www.pdflib.com