

Was ist neu in PDFlib/PDFlib+PDI/PPS 7?

PDFlib 7 bietet zahlreiche Neuerungen und Verbesserungen, die in folgender Übersicht zusammengefasst werden.

Tabellenformatierung. PDFlib enthält einen neuen Tabellenformatierer, mit dem sich Tabellenzeilen und -spalten automatisch nach Benutzervorgaben platzieren und Tabellen über mehrere Textspalten oder Seiten verteilen lassen. Tabellenzellen können ein- oder mehrzeiligen Text, Bilder oder PDF-Seiten enthalten und anhand zahlreicher Optionen formatiert werden, etwa Rahmenfarbe, Hintergrund oder Zellenlineatur. Die Größe der Tabellenzeilen und -spalten wird automatisch berechnet, wobei verschiedene konfigurierbare Einstellungen berücksichtigt werden.

Textflow-Erweiterungen. Der Textflow-Formatierer für Fließtext wurde verbessert:

- ▶ Links und andere interaktive Elemente können automatisch aus den Textfragmenten innerhalb eines Textflows angelegt werden.
- ▶ In einer Textspalte kann Platz für Bilder reserviert werden.
- ▶ Die Textflow-Formatierung unterstützt Führungszeichen, zum Beispiel wenn in einem Inhaltsverzeichnis die Einträge durch Punkte mit den entsprechenden Seitenzahlen verbunden werden.
- ▶ Textinhalte und Formatierungsoptionen können schrittweise in mehreren Portionen übergeben werden. Es ist kein Puffer mehr erforderlich, der den gesamten Text inklusive der Inline-Optionen enthält.
- ▶ Zeichen können anderen Formatierungsklassen zugeordnet werden, um zum Beispiel festzulegen, ob das Zeichen »/« bei der Formatierung wie ein Buchstabe oder wie ein Satzzeichen behandelt wird.

Weitere Formatierungsfunktionen.

- ▶ Dezimaltabulatoren und Führungszeichen werden jetzt auch in einzeiligem Text, unterstützt (bisher nur in mehrzeiligen Textflows).
- ▶ Die neue Stempelfunktion berechnet die optimale Größe und Position eines Textstempels für ein Rechteck.
- ▶ Verbesserte Abfragefunktionen für die Textposition und andere geometrische Eigenschaften.

Fontbehandlung und Unicode. Die Fontverarbeitung wurde wie folgt erweitert:

- ▶ Benutzer können Eigenschaften eines geladenen Fonts in allen Einzelheiten abfragen, einschließlich der typografischen Einträge in TrueType/OpenType-Fonts, der Anzahl der verfügbaren Glyphen, u.v.m.
- ▶ Für Type-3-Fonts können Untergruppen gebildet werden.
- ▶ Unicode-Encoding wird für alle Fonttypen einschließlich Type 3 unterstützt.
- ▶ Text kann im Format UTF-32 übergeben werden; für Unicode-Werte jenseits der Basic Multilingual Plane (d.h. größer U+FFFF) werden Surrogate-Paare unterstützt.
- ▶ Die Ersetzung für im Font fehlende Glyphen kann vom Benutzer gesteuert werden, zum Beispiel könnte der griechische Buchstabe *Omega* statt einer fehlenden Glyphe *Ohm* verwendet werden. Ist auch die Ersatzglyphe nicht verfügbar, wird eine entsprechend definierbare Standardersatzglyphe genommen. Alternativ kann PDFlib zur Ausgabe einer Warnung veranlasst werden, wenn eine benötigte Glyphe im Font nicht vorhanden ist.

- ▶ Wenn in einer Schriftfamilie eine fette, kursive oder fett-kursive Variante fehlt, kann PDFlib einen künstlichen Schriftstil erzeugen.
- ▶ Neben numerischen Referenzen können Glyphen in einem Font über ihren Namen adressiert werden, was zum Beispiel bei Ligaturen oder Stilvarianten hilfreich ist.
- ▶ Sequenzen aus Gegenschrägstrichen im Stil von C und Java werden im Text erkannt. Dies erleichtert die Behandlung von Steuerzeichen und nicht-ASCII-Zeichen unabhängig von den Anforderungen und Einschränkungen der jeweiligen Programmiersprache.

Verbesserte Verarbeitung von chinesischem, japanischem und koreanischem Text. In PDFlib 7 entfallen eine Reihe von Einschränkungen bei der CJK-Textbehandlung:

- ▶ PDFlib bietet vollständige Unterstützung für die Zeichenmetriken bei allen Standard-CMaps; so kann zum Beispiel Shift-JIS-Text jetzt mittels Textflow formatiert werden.
- ▶ Vertikale Schreibrichtung wird für alle TrueType- und OpenType-Schriften unterstützt.
- ▶ Chinesische, japanische und koreanische Codepages (etwa Codepage 932 für Shift-JIS) werden nicht nur unter Windows, sondern auf allen Plattformen unterstützt.
- ▶ CJK-CMaps werden nicht nur für den Seiteninhalt, sondern auch für interaktive Features wie Lesezeichen unterstützt.
- ▶ Die in Acrobat vordefinierten CJK-Fonts können mit Unicode-Encoding verwendet werden.
- ▶ OpenType-CID-Fonts mit einer der vordefinierten CMaps werden nicht zwangsläufig eingebettet, was die Dateigröße erheblich verringern kann.

Matchboxen. Das Matchbox-Konzept, das in verschiedenen Text- und Bildfunktionen unterstützt wird, bietet einfachen Zugriff auf die Koordinaten formatierter Text- oder Bildobjekte. Durch einfaches Markieren (statt einer Berechnung über Koordinaten) lassen sich zum Beispiel automatisch Anmerkungen oder Auszeichnungen erstellen oder Links in Textflows anlegen, Textabschnitte mit Rahmen versehen oder Text innerhalb eines formatierten Absatzes hervorheben.

pCOS-Schnittstelle in PDI integriert. PDFlib enthält das PDF-Abfragemodul pCOS 2.0. Über eine einfache Pfadsyntax lassen sich beliebige Eigenschaften eines vorhandenen PDF-Dokuments abfragen. Damit lassen sich zum Beispiel alle vorhandenen Schriften, Bilder oder Farbräume auflisten, Seiten- und Dokumenteigenschaften ermitteln oder Informationen wie der PDF/A- oder PDF/X-Status, Dokumentinfelder oder XMP-Metadaten abfragen. Die seit 2005 in pCOS 1.0 enthaltenen Kernfunktionen wurden erheblich erweitert, etwa durch Bild- und Farbraumeigenschaften, symbolische Seitennamen (page labels), Ressourcen und vieles mehr.

PDF-Import (PDI). PDI enthält einen neuen Reparaturmodus für beschädigte PDF-Eingabedokumente. Ein neuer Optimierungsschritt entfernt redundante Objekte, die sich beim Importieren mehrerer PDF-Dokumente ergeben können. Enthalten die Dokumente zum Beispiel mehrmals denselben Font, werden die überflüssigen Fonts nicht mehr in die Ausgabe übernommen, sondern entfernt.

AES-Verschlüsselung. PDFlib unterstützt 128-Bit-Verschlüsselung mit dem in Acrobat 7 eingeführten AES-Verfahren (Advanced Encryption Standard). AES-Verschlüsselung gilt als wesentlich sicherer als ältere Chiffrierverfahren.

Weitere Funktionen von PDF 1.6 (Acrobat 7). *UserUnits* bieten eine bessere Dokumentskalierung und ein größeres Spektrum an Seitengrößen. Es gibt neue Öffnen-Modi für das Dokument (zum Beispiel Navigationsregisterkarte *Anhänge* sichtbar) oder die Möglichkeit, eine Standardskalierung des Dokuments für den Ausdruck zu bestimmen.

PDF/A zur Archivierung. PDFlib kann Ausgabe in den Formaten PDF/A-1a und PDF/A-1b (standardisiert als ISO 19005-1) erzeugen. PDF/A legt eine standardisierte Teilmenge von PDF fest, die auf die langfristige Aufbewahrung und Archivierung von PDF-Dokumenten ausgerichtet ist. Vorhandene PDF/A-Dokumente können importiert und zusammengefasst oder aufgeteilt werden; Bilder (in jedem Farbraum) können nach PDF/A konvertiert werden. Während PDF/A-1b das visuelle Erscheinungsbild von PDF-Dokumenten erhält, kann mit PDFlib auch die erweiterte Variante PDF/A-1a erstellt werden, die zudem die Semantik der Dokumente garantiert.

PDFlib Personalization Server und Block-Plugin. Mehrere Textflow-Blöcke können so verknüpft werden, dass ein Block den übrigen Text des Vorgängerblocks aufnimmt. Dies ermöglicht flexiblere Layouts bei der Personalisierung von Dokumenten mit variablen Daten. Mit der neuen pCOS-Schnittstelle können Sie verschiedenste Arten von Blockdaten aus einem PDF abfragen.

Interaktive Elemente und 3D-Animationen. Anmerkungen (Weblinks) können auf eine separate Ebene gelegt werden, so dass sie nur sichtbar sind, wenn auch die zugehörige Ebene sichtbar ist. In die PDF-Ausgabe können 3D-Animationen im U3D-Format eingebettet werden, wobei zahlreiche Optionen zur Steuerung bereitstehen. Aktionen dienen zur Steuerung der 3D-Animationen

Schmuckfarben. Es werden alle Pantone-Schmuckfarben unterstützt, inklusive der 2006 von Pantone, Inc. veröffentlichten neuen Farbskalen. Dazu gehören die neue Pantone Color Bridge sowie neue Farben in den Metallic- und Pastel-Skalen. Die Pantone-Farbnamen wurden in das PDFlib-Block-Plugin integriert und können direkt in der Benutzeroberfläche für Blockeigenschaften ausgewählt werden.

Bildbehandlung. In TIFF- und JPEG-Bildern wird der Freistellungspfad berücksichtigt, so dass bei der Platzierung von Bildern die Trennung von Vorder- und Hintergrund automatisch erhalten bleibt, ohne dass weitere Beschneidungs- oder Transparenzoperationen erforderlich wären.

XMP-Metadaten. PDFlib erstellt aus Dokumentinfocfeldern automatisch XMP-Metadaten. Benutzer können separat erzeugte XMP-Daten für das Dokument sowie andere Objekte wie Seite, Font, Bild, importierte PDF-Seite, Template oder ICC-Profil einbinden. Für client-spezifische Metadaten werden XMP-Schemata unterstützt.

Sprachbindungen. Es gibt diverse Verbesserungen in den Sprachbindungen, insbesondere die Unterstützung neuer Versionen (etwa Python 2.5) und Unicode-Unterstützung im Python-Wrapper.

Dokumentation. Die Dokumentation wurde in zwei Manuale aufgeteilt; zusätzlich gibt es das PDFlib-Kochbuch mit Codebeispielen und Erläuterungen.