

23.12.1 Klasse PdfDocument (gb.poppler)

Diese Klasse repräsentiert ein PDF-Dokument. Sie ermöglicht es, seinen Inhalt zu lesen und zu rendern. Die Klasse verhält sich wie ein Nur-Lesen-Array. So erzeugen Sie aus einer PDF-Datei in Gambas ein neues PDF-Dokument vom Typ PdfDocument:

```
Dim hPdfDocument As PdfDocument
hPdfDocument = New PdfDocument ( Path As String [ , Owner As String, Password As String ] )
```

- Path ist der Pfad zur PDF-Datei.
- Owner ist der Eigentümer des Dokuments.
- Password ist das Kennwort. Die letzten beiden Parameter sind optional.

Mit Hilfe der Klasse PdfPage können Sie eine bestimmte Seite eines PDF-Dokumentes anhand ihres Indexes als virtuelles Objekt zurückgeben. Index ist dabei ein Wert zwischen 0 und dem Wert der Eigenschaft PdfDocument.Max = PdfDocument.Count -1.

```
Dim hPdfDocument As PdfDocument
Dim hPdfPage As PdfPage
hPdfPage = hPdfDocument [ Index As Integer ]
```

23.12.1.1 Eigenschaften

Die Klasse PdfDocument verfügt über die folgenden Eigenschaften:

Eigenschaft	Datentyp	Beschreibung
Antialiasing	Boolean	Gibt den Wert zurück oder setzt, ob Antialiasing für das Zeichnen von Seiten verwendet wird oder nicht.
Author	String	Gibt den Autor des PDF-Dokuments zurück.
Count	Integer	Gibt die Anzahl der Seiten des PDF-Dokuments zurück.
Max	Integer	Gibt die Anzahl der Seiten des Dokuments minus eins zurück; Max = Count -1.
CreationDate	Date	Gibt das Datum der Erzeugung des PDF-Dokuments zurück.
Creator	String	Gibt den Ersteller des PDF-Dokuments zurück.
Producer	String	Gibt den Hersteller des PDF-Dokuments zurück.
Index	.PdfDocumentIndex	Gibt den Index (Inhaltsverzeichnis) eines PDF-Dokuments als virtuelles Objekt vom Typ .PdfDocumentIndex zurück.
Keywords	String	Gibt eine komma-separierte Liste als String zurück.
Linearized	Boolean	Gibt an, ob das Dokument linearisiert ist oder nicht. Die Linearisierung von PDF-Inhalt ermöglicht einen effizienten inkrementellen Zugriff auf die PDF-Datei in einer Netzwerkumgebung.
ModificationDate	Date	Gibt das Datum der letzten Änderung des PDF-Dokuments zurück.
Resolution	Float	Gibt die Auflösung des PDF-Dokuments in DPI zurück oder legt sie fest.
Rotation	Integer	Liefert oder setzt die globale Seitendrehung in Grad. Der Wert muss ein Vielfaches von 90 sein: 0, 90, 180 oder 270. Der Drehwinkel kann auch negativ wie zum Beispiel -90 sein.
Subject	String	Gibt den Betreff des PDF-Dokuments zurück.
TextAntialiasing	Boolean	Gibt den Wert zurück oder setzt, ob der Text beim Zeichnen einer Seite anti-aliased ist.
TextHinting	Boolean	Gibt den Wert zurück oder setzt, ob beim Zeichnen einer Seite Texthinweise verwendet wurden.
Title	String	Gibt den Titel des PDF-Dokuments zurück.
Version	String	Gibt die Version des PDF-Dokuments als Zeichenfolge zurück.

Tabelle 23.12.1.1.1 : Eigenschaften der Klasse PdfDocument

Hinweise

- **Creator:** Wenn das Dokument aus einem anderen Format konvertiert wurde, ist der Ersteller (C) der Name des Produkts (P), das das ursprüngliche Dokument erzeugt hat, aus dem es konvertiert wurde. Quelle: <https://poppler.freedesktop.org/api/glib/PopplerDocument.html#poppler-document-get-creator>. Beispiele: C = Writer, P = LibreOffice 6.4; C = dvips(k) 5.96 Copyright 2007 Radical Eye Software, P = GPL Ghostscript 8.60.
- **Producer:** Gibt den Hersteller des Dokuments zurück. Wenn das Dokument aus einem anderen Format konvertiert wurde, ist der Hersteller der Name des Produkts, das es in PDF konvertiert hat. Quelle: <https://poppler.freedesktop.org/api/glib/PopplerDocument.html#poppler-document-get-producer>.
- **Keywords:** Beachten Sie, dass nur eine intern komma-separierte Zeichenkette zurückgegeben wird – wie zum Beispiel "gb.media, gb.media.form, GStreamer".
- **ModificationDate:** Wenn ein PDF-Dokument beispielsweise mit OpenLibre neu erzeugt wurde, dann hat der ausgelesene Wert stets ein Datum mit der Jahreszahl 1969 oder 1970 – ist somit inkorrekt. Der folgende Quelltext verzichtet deshalb auf ModificationDate und verwendet die Eigenschaft, die beschreibt, wann das PDF-Dokument zuletzt geändert wurde:

```
sMessage = Format(Stat(FMain.DocumentPath).LastChange, "dd.mm.yyyy hh:nn:ss")
Print sMessage
```

23.12.1.2 Methoden

Die Klasse PdfDocument verfügt nur über eine Methode:

Methode	Rückgabotyp	Beschreibung
Find (Label As String)	Integer	Gib den Seitenindex der Seite zurück, die den Bezeichner 'Label' hat. Wurde keine Seite gefunden, so wird -1 zurückgegeben.

Tabelle 23.12.1.2.1 : Methode der Klasse PdfDocument

23.12.1.3 Exkurs

Auf der Seite <https://wiki.ubuntuusers.de/poppler-utils/> finden Sie die kurze Beschreibung von Kommandozeilen-Programmen der Sammlung poppler-utils, mit denen Sie u.a. Informationen zu PDF-Dokumenten gewinnen oder diese zum Beispiel in eine XML-Datei konvertieren können.

Hier folgt ein Beispiel für das Auslesen der in einem PDF-Dokument verwendeten Schriftarten (Fonts), das so mit Hilfe der Gambas-Komponente gb.poppler nicht möglich ist:

```
hans@pc-a-mint20:~/Dokumente$ pdffonts k23.12.pdf
name                type                encoding            emb sub uni object ID
-----
BAAAAA+LiberationSans  TrueType           WinAnsi            yes yes yes      48  0
CAAAAA+DejaVuSansMono  TrueType           WinAnsi            yes yes yes      33  0
DAAAAA+LiberationSans-Italic  TrueType           WinAnsi            yes yes yes      43  0
EAAAAA+LiberationSans-Bold    TrueType           WinAnsi            yes yes yes      38  0
```