

Version: 1.0

Datum: 17. 10. 2005

Autor: Michael Horvath

michael.horvath@onb.ac.at

Datei: ONB_PDF-Einstellungen_GSview4_1-0

Ghostgum GSview 4.7: Einstellungen zum Erzeugen archivierbarer Dateien im Format PDF

Einführung

Das vorliegende Dokument basiert auf den von der ÖNB separat publizierten „Empfehlungen zum Erzeugen archivierbarer Dateien im Format PDF“. Es dient dem Zweck, die praktische Umsetzung dieser Empfehlungen bei der Verwendung des PDF-Schreibprogramms „pdfwrite“ innerhalb der von der australischen Firma **Ghostgum** entwickelten Software **GSview 4.7**, im Folgenden **GSview** genannt, zu unterstützen.

Voraussetzung für die Nutzung von **GSview** ist die Installation von **AFPL Ghostscript** ab der Version 7.04, einem Programm zum Interpretieren von PostScript-Dateien, da **GSview** eine grafische Benutzeroberfläche zu **Ghostscript** darstellt. Beide Produkte können kostenlos aus dem Internet heruntergeladen werden, z. B. unter <http://www.cs.wisc.edu/~ghost/gsview/>.

Sollten die im Folgenden beschriebenen Einstellungen in einzelnen Fällen dazu führen, dass bei der Erstellung einer PDF-Datei signifikante Eigenschaften der Ursprungsdatei verloren gehen, etwa Bildinformationen aus einer digitalen Druckvorlage, dann können die Einstellungen nach Rücksprache mit der ÖNB modifiziert werden.

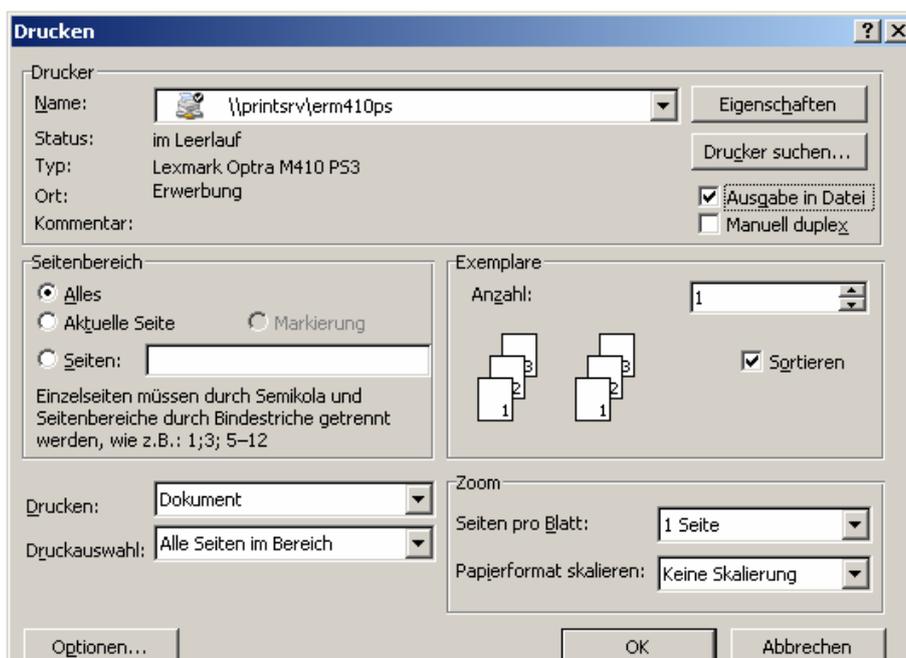
Arbeitsfluss

Beim Erzeugen einer PDF-Datei mithilfe von **GSview** aus einer Datei in einem beliebigen anderen druckbaren Dateiformat muss die Ursprungsdatei zunächst in das von Adobe zur Steuerung von Druckern entwickelte SeitenbeschreibungsfORMAT PostScript, ein Vorgängerformat zu PDF, übersetzt werden. Auf diese Weise entsteht eine Datei, die auf einen ausgewählten Drucker hin optimiert ist.

In einem der Microsoft-Office-Programme Word, Excel oder Powerpoint erzeugt man die benötigte Druckdatei beispielsweise dadurch, dass man unter „Datei – Drucken...“ einen Drucker auswählt, der PostScript verarbeiten kann, was auf die meisten heutigen Laserdrucker zutrifft, und das Kästchen „Ausgabe in Datei (umleiten)“ ankreuzt (siehe Abb. 1). Nach dem Bestätigen durch „OK“ wählt man das Verzeichnis, in dem die Druckdatei mit der Endung „.prn“ angelegt werden soll, und startet durch erneutes Bestätigen den Schreibvorgang.

Hinweis: Unter Umständen ist es notwendig, die unter „Eigenschaften“ (siehe Abb.1) für den ausgewählten Drucker festgelegten Werte zu verändern, um eine PostScript-Datei zu erzeugen, die als brauchbare Vorlage für eine archivierbare PDF-Datei dienen kann. Da aber die einstellbaren Eigenschaften je nach Drucker variieren, können hierzu keine allgemeinen Empfehlungen gegeben werden.

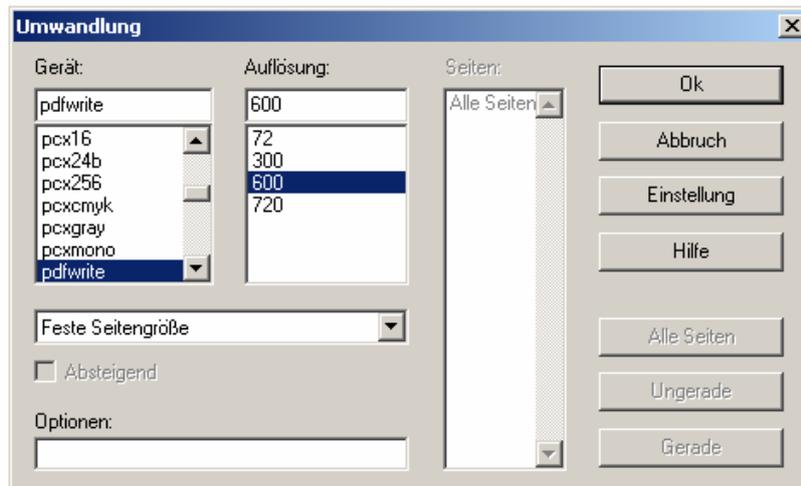
Abb. 1: Microsoft-Office-Programme



Auf Basis der so entstandenen Druckdatei kann man mithilfe von **GSview** in einem weiteren Schritt eine PDF-Datei erzeugen. Nach dem Starten des Programms wählt man zunächst mit „Datei – Öffnen...“ die Druckdatei aus, um sie schließlich mit „Datei – Konvertierung...“ umzuwandeln (siehe Abb. 2). Als „Gerät“ wählt man das bereits erwähnte „pdfwrite“.

Für die „Auflösung“ wird der Wert „600“ empfohlen, was die Standardauflösung für die Konvertierung auf 600 Bildpunkte pro Zoll (2,54 cm) festlegt. Dieser Wert ist im Allgemeinen sowohl für die Darstellung auf einem Monitor als auch für das Drucken ausreichend. Die Einstellung „Feste Seitengröße“ bewirkt, dass alle Seiten der zu erzeugenden PDF-Datei jene Größe haben, die unter „Formate“ eingestellt ist, wobei „A4“ als Standardformat dient.

Abb. 2: „Umwandlung“ in GSview



Grundsätzlich gilt, dass aus einer Ursprungsdatei verschiedenste PDF-Dateien generiert werden können, die zwar auf einem Monitor gleich dargestellt sind, sich in ihren Eigenschaften jedoch stark voneinander unterscheiden. Nach Klicken auf „Einstellung“ (siehe Abb. 2) können einige der Eigenschaften der zu erzeugenden PDF-Datei dahingehend optimiert werden, dass sie deren langfristiger Archivierbarkeit möglichst entgegenkommen.

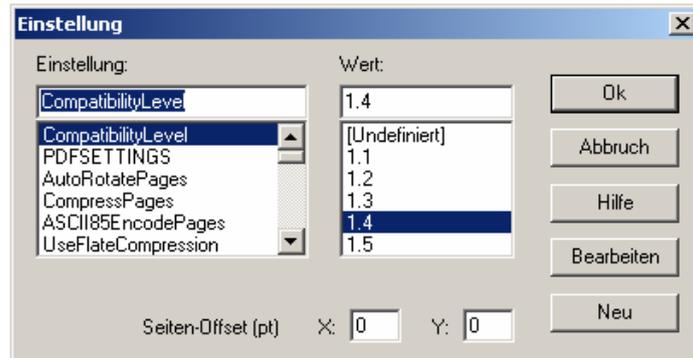
Die ÖNB empfiehlt, die im folgenden Kapitel angeführten Werte einzustellen, bevor eine Datei ins Format PDF umgewandelt wird.

Hinweis: Bei der Eingabeaufforderung für den Namen der resultierenden Datei im Fenster „Name der Ausgabedatei“, das nach Bestätigen der Eingaben für das Fenster „Umwandlung“ erscheint, ist die Endung „.pdf“ manuell einzugeben, damit die Datei von PDF-Leseprogrammen geöffnet werden kann.

Einstellungen

Wenn im Fenster „Umwandlung“ auf die Schaltfläche „Einstellung“ geklickt wird, erscheint das in Abb. 3 dargestellte Fenster in der Anzeige. Der untenstehenden Kurzübersicht entsprechend, können hier Werte für insgesamt 33 Einstellungen (Variablen) gesetzt werden, die die Umwandlung einer Datei in das Format PDF steuern. Dazu werden zuerst die jeweilige Einstellung (links) und dann der dazugehörige Wert (rechts) markiert, bis die Liste abgearbeitet ist. Anschließend kann das Fenster durch Bestätigen mit „OK“ wieder geschlossen werden, und die nun eingestellten Werte sind auch nach einem Neustart von **GSview** verfügbar.

Abb. 3: „Einstellung“ in GSview



Einstellung:	Wert:
CompatibilityLevel	1.4
PDFSETTINGS	[Undefiniert]
AutoRotatePages	/None
CompressPages	true
ASCII85EncodePages	false
UseFlateCompression	true
EmbedAllFonts	true
SubsetFonts	true
MaxSubsetPct	100
ConvertCMYKImagesToRGB	false
AutoFilterColorImages	false
ColorImageFilter	/FlateEncode
EncodeColorImages	true
DownsampleColorImages	false
ColorImageDepth	[Undefiniert]
ColorImageResolution	[Undefiniert]
ColorImageDownsampleThreshold	[Undefiniert]
ColorImageDownsampleType	[Undefiniert]
AutoFilterGrayImages	false
GrayImageFilter	/FlateEncode
EncodeGrayImages	true
DownsampleGrayImages	false
GrayImageDownsampleThreshold	[Undefiniert]
GrayImageDownsampleType	[Undefiniert]
GrayImageDepth	[Undefiniert]
GrayImageResolution	[Undefiniert]
MonolImageFilter	/CCITTFaxEncode
EncodeMonolImages	true
DownsampleMonolImages	false
MonolImageDownsampleThreshold	[Undefiniert]
MonolImageDownsampleType	[Undefiniert]
MonolImageDepth	[Undefiniert]
MonolImageResolution	[Undefiniert]

Erklärungen zu den Einstellungen

CompatibilityLevel: 1.4

Hier wird eingestellt, auf welcher Spezifikation des Formats PDF die zu erzeugende Datei basieren soll. Wie im Entwurf zum PDF/A-Standard gefordert, wird die Version 1.4 verwendet.

PDFSETTINGS: [Undefiniert]

Es sollen keine der von Adobe vordefinierten Einstellungen verwendet werden.

AutoRotatePages: /None

Diese Option ermöglicht es, die Seiten eines Dokuments so zu drehen, dass der überwiegende Teil des Textes lesbar ist, was aber einerseits die ursprüngliche Datei verändert und andererseits zu Problemen beim Editieren führen kann. Die in der Quelldatei gewählte Ausrichtung der Seiten wird daher beibehalten.

CompressPages: true

Das Komprimieren der Texte und Vektorgrafiken auf den Seiten eines Dokuments verringert die Größe der resultierenden PDF-Datei. Wird für diesen Vorgang ein „verlustfreies“ Kompressionsverfahren verwendet, bei dem keine in der Quelldatei vorhandene Information verloren geht (siehe Variable „UseFlateCompression“), dann kann die Variable auf „true“ gesetzt werden.

ASCII85EncodePages: false

Hier wird festgelegt, ob die Daten im Binärformat oder im ASCII-Textformat in die PDF-Datei geschrieben werden. Da aus technischen Gründen die Speicherung der gleichen Daten in einer Binärdatei weniger Speicherplatz braucht als in einer Textdatei, wird ersteres Format bevorzugt.

UseFlateCompression: true

Wenn „CompressPages“, wie oben eingestellt, den Wert „true“ hat, dann wird mit „UseFlateCompression“ entschieden, welches Verfahren beim Komprimieren der Seiten zum Einsatz kommen soll. „FlateCompression“, auch als „ZIP“ bezeichnet, ist ein verlustfrei arbeitendes Kompressionsverfahren, dessen Funktionsweise öffentlich und frei zugänglich vorliegt, so dass es auch im Bereich der Langzeitarchivierung eingesetzt werden kann.

EmbedAllFonts: true

Gemäß den separat publizierten PDF-Empfehlungen soll eine Datei die grafischen Beschreibungen aller darin verwendeten Schriftarten enthalten, damit sie auch dann originalgetreu dargestellt werden kann, wenn auf dem darstellenden Computer nicht alle erforderlichen Schriftarten installiert sind. Diese „Einbettung von Schriften“ wird durch obigen Parameter gewährleistet.

SubsetFonts: true

Werden nicht alle Zeichen einer Schriftart in der Datei verwendet, was meist der Fall ist, dann soll darin eine Untergruppe dieser Schriftart eingebettet werden, die lediglich die grafischen Beschreibungen der tatsächlich vorkommenden Zeichen enthält. Vor allem bei Unicode-Schriften kann diese Option zu einer erheblichen Verringerung der Dateigröße führen.

MaxSubsetPct: 100

Die Variable wird ausgewertet, wenn „SubsetFonts“ den Wert „true“ hat (siehe oben). Von einer verwendeten Schriftart soll nur dann eine Untergruppe gebildet werden, wenn der Prozentsatz der daraus in der Datei tatsächlich vorkommenden Zeichen kleiner ist als der hier angegebene. Der Wert „100“ bedeutet, dass nach Möglichkeit immer eine Untergruppe gebildet werden soll, also dann, wenn nicht 100 % der in der Schriftart enthaltenen Zeichen verwendet werden.

Hinweis: Sollte der Wert „100“ in der entsprechenden Auswahl nicht zur Verfügung stehen, kann er manuell eingegeben werden.

Für die nun folgenden Einstellungen bezüglich der Verarbeitung von Grafiken gilt der Grundsatz, dass diese möglichst unverändert in die PDF-Datei übernommen werden sollen.

ConvertCMYKImagesToRGB: false

CMYK (Cyan – Magenta – Yellow – Keycolor = Black) ist ein Modell zur Beschreibung von Farben, welches diese durch subtraktive Mischung oben genannter vier Grundfarben abbildet. Es findet vor allem im professionellen Druckbereich Verwendung. Bei RGB (Red – Green – Blue) wird dagegen jede Farbe durch additive Farbmischung von Anteilen an Rot, Grün und Blau beschrieben. Monitore arbeiten mit RGB-Farben.

Da beim Konvertieren von CMYK-Bildern in RGB-Bilder die ursprüngliche Farbinformation verloren ginge, wird dieser Vorgang nicht durchgeführt.

AutoFilterColorImages: false

Wenn die Variable „EncodeColorImages“ den Wert „true“ hat (siehe unten), wird mit dieser Einstellung entschieden, ob das PDF-Schreibprogramm das Verfahren zum Komprimieren von Farbbildern jeweils automatisch wählen soll. Um den Einsatz von verlustbehafteten Kompressionsverfahren zu vermeiden, wird dieses automatische Komprimieren abgeschaltet.

ColorImageFilter: /FlateEncode

Zum Komprimieren von Farbbildern soll generell das verlustfreie FlateEncode-Verfahren (ZIP) verwendet werden.

EncodeColorImages: true

Alle in der Quelldatei enthaltenen Farbbilder sollen komprimiert werden.

DownsampleColorImages: false

Die Auflösung, das ist die Anzahl der Bildpunkte, aus denen eine Farbgrafik besteht, soll gegenüber der ursprünglichen Datei nicht verringert werden.

ColorImageDepth: [Undefiniert]

ColorImageResolution: [Undefiniert]

ColorImageDownsampleThreshold: [Undefiniert]

ColorImageDownsampleType: [Undefiniert]

Da sich diese Variablen nur dann auf die Erzeugung einer PDF-Datei auswirken, wenn „DownsampleColorImages“ den Wert „true“ hat, müssen sie in diesem Fall nicht definiert werden.

AutoFilterGrayImages: false

Einstellung für Graustufenbilder, analog zu „AutoFilterColorImages“.

GrayImageFilter: /FlateEncode

Einstellung für Graustufenbilder, analog zu „ColorImageFilter“.

EncodeGrayImages: true

Einstellung für Graustufenbilder, analog zu „EncodeColorImages“.

DownsampleGrayImages: false

Einstellung für Graustufenbilder, analog zu „DownsampleColorImages“.

GrayImageDownsampleThreshold: [Undefiniert]

GrayImageDownsampleType: [Undefiniert]

GrayImageDepth: [Undefiniert]

GrayImageResolution: [Undefiniert]

Einstellungen für Graustufenbilder, analog zu jenen für Farbbilder.

MonolImageFilter: /CCITTFaxEncode

Für die Komprimierung von monochromen Bildern („schwarzweiß“) wird das ursprünglich zur Übertragung von Fax-Nachrichten entwickelte und international gebräuchliche CCITT-Format verwendet, das verlustfrei arbeitet und im Durchschnit die besten Ergebnisse liefert.

EncodeMonolImages: true

Einstellung für Schwarzweißbilder, analog zu „EncodeColorImages“.

DownsampleMonolImages: false

Einstellung für Schwarzweißbilder, analog zu „DownsampleColorImages“.

Referenzen

- PDF Reference — Third Edition — Adobe Portable Document Format — Version 1.4
<http://partners.adobe.com/public/developer/en/pdf/PDFReference.pdf>
- Document management — Electronic document file format for long-term preservation — Part 1: Use of PDF (PDF/A)
[http://www.aiim.org/documents/standards/ISO_19005-1_\(E\).doc](http://www.aiim.org/documents/standards/ISO_19005-1_(E).doc)
- Adobe Acrobat Distiller Parameters – Technical Note #5151 – Version: Acrobat 5.0 (“Distiller Parameter-Handbuch” unter Menüpunkt “Hilfe”)
- How to Create Accessible Adobe PDF Files
<http://www.adobe.com/products/acrobat/pdfs/accessbooklet.pdf>