



Secrets from the Support Vaults
---Episode #2-10---



Good afternoon from Berlin

Many things can go wrong with text and fonts in PDF

A support case



staatl. anerk. Podologe
med. Fußpfleger

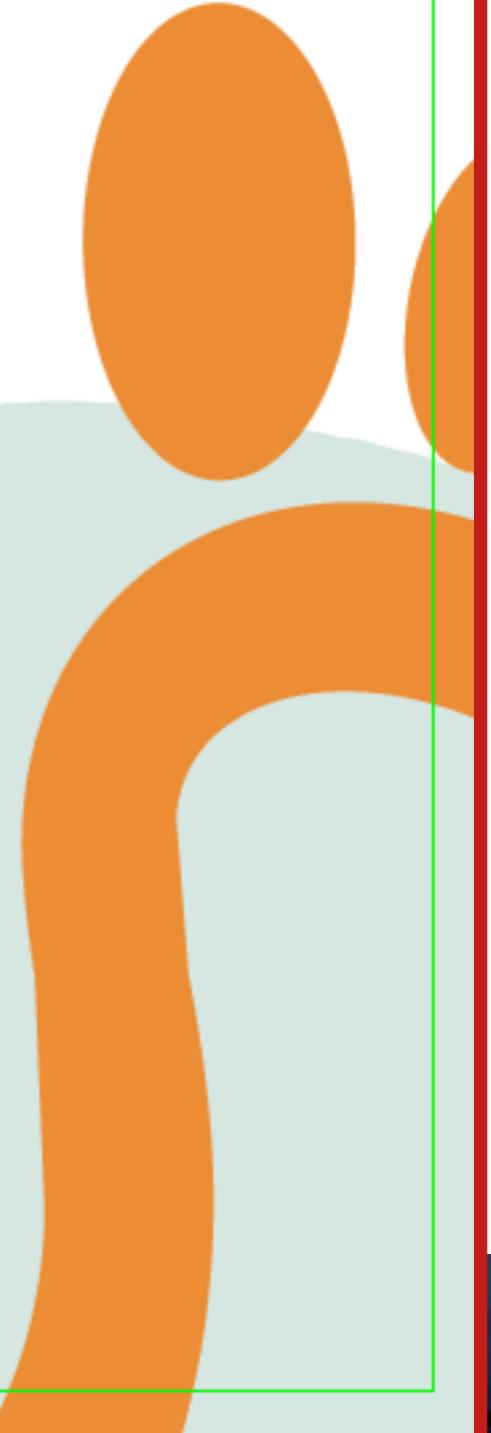
Telefon

Mobil

E-Mail

Web

✉️ 📞 📩 📧 📨 📩 📧 📩 📧 📩 📧 📩



Profile: “List potential font problems”

The screenshot shows a comparison between a PDF document and its analysis results.

Left Side (PDF Document): A page from a PDF titled "Kundendatei_1.pdf". The page contains the logo "Fußmobil" (with a stylized orange foot icon), the name "Patrick" (with a redacted address bar), and contact information: "staatl. anerk. Podologe", "med. Fußpfleger", "Telefon" (redacted), "Mobil" (redacted), "E-Mail" (redacted), "Web" (redacted), and the address "Remseder Str. 5, 49196 Bad". The page is labeled "1/1" and has dimensions "9.10001 x 6.10002 cm" and "PDF/X-3 FOGRA39".

Right Side (Analysis Results): The "callas pdfToolbox" interface. The title bar says "callas pdfToolbox" and "Result". The main area displays a warning message: "callas pdfToolbox profile "List potential font problems" found the following warnings:". Below this, a tree view shows:

- Page 1 from "Kundendatei_1.pdf"
- List potential font problems
 - Encoding entry prohibited for symbolic TrueType font (2 matches on 1 page)
PDF/A prohibits that symbolic TrueType fonts use the Encoding entry.
 - Text cannot be mapped to Unicode (2 matches on 1 page)
- Overview
- callas pdfToolbox information

At the bottom are buttons for "Show in Snap", "Certify", and "Create Report...".

Page Footer: "Licensed to Dietrich (callas)"

Overview

- Font tools in pdfToolbox
- Fonts in PDF
- Issues with glyph selection
- More font tools in pdfToolbox



Font tools in pdfToolbox

- Profiles: Validate for PDF/X-4 or PDF/A-2
- Profile: List potential font problems



Checks in PDF/X-4 and PDF/A-2 and in “List potential font problems”

- In PDF/X-4 and PDF/A-2
 - Character references .notdef glyph
 - Encoding entry prohibited for symbolic TrueType font
 - Font not embedded
 - Glyph in subset font missing
 - Glyph width info in PDF does not match width info in embedded font
 - More than one encoding in symbolic TrueType font’s cmap
 - Wrong encoding for non-symbolic TrueType font
- In addition in PDF/A-2
 - Text cannot be mapped to Unicode (only conformance levels A and U)
 - Type 2 CID font: CIDToGIDMap invalid or missing
- In addition in “List potential font problems”
 - Character is not white space but glyph is empty
 - Font name is not unique
 - Font reverts to .notdef glyph
 - No ‘cmap’ entry in Simple TrueType font



Checks in PDF/X-4 and PDF/A-2 and in “List potential font problems”

- In PDF/X-4 and PDF/A-2
 - Character references .notdef glyph
 - Encoding entry prohibited for symbolic TrueType font
 - Font not embedded
 - Glyph in subset font missing
 - Glyph width info in PDF does not match width in font
 - More than one encoding in symbolic TrueType font
 - Wrong encoding for non-symbolic TrueType font
- In addition in PDF/A-2
 - Text cannot be mapped to Unicode (only conforming fonts)
 - Type 2 CID font: CIDToGIDMap invalid or missing
- In addition in “List potential font problems”
 - Character is not white space but glyph is empty
 - Font name is not unique
 - Font reverts to .notdef glyph
 - No ‘cmap’ entry in Simple TrueType font

- What is
 - a glyph?
 - a .notdef glyph?
 - an empty glyph?
- What is a subset font?
- What is a font encoding?
- Which font types can be used in a PDF?
- What is a “symbolic” font?
- What means “Mapping to Unicode”?

First answers (using pdfToolbox' “Explore Fonts” feature)

- What is
 - a ‘glyph’ is the shape of a character in a font
 - a ‘.notdef’ glyph is a particular character that is used when glyph selection fails
 - an ‘empty glyph’ is a glyph without contour (shape), e.g. a blank

The image displays three separate windows of the 'callas pdfToolbox: Browse Internal Structure of All Document Fonts' tool, each showing a different font's internal structure.

- Font 1 (Left):** Shows a large blue outline of the letter 'C'. The left panel lists GID values from 50 to 72. The right panel shows a tree view of font resources, including 'Table Directory', 'cmap', 'cvt', 'head', 'hhea', 'hmtx', 'loca', 'maxp', 'name', 'post', 'fpgm', 'prep', and 'glyf'. The 'glyf' section shows a tag 'Tag: 0x676C796' and entries for 'GID: 0 ".notdef"', 'Empty Glyphs: [1]', 'GID: 2 "quotedbl"', 'GID: 3 "dollar"', and 'GID: 4 "quoteright"'.
- Font 2 (Middle):** Shows a large blue square with a white center. The left panel lists GID values from 50 to 72. The right panel shows a tree view of font resources, including 'Table Directory', 'cmap', 'cvt', 'head', 'hhea', 'hmtx', 'loca', 'maxp', 'name', 'post', 'fpgm', 'prep', and 'glyf'. The 'glyf' section shows a tag 'Tag: 0x676C796' and entries for 'GID: 0 ".notdef"', 'Empty Glyphs: [1]', 'GID: 2 "quotedbl"', 'GID: 3 "dollar"', and 'GID: 4 "quoteright"'.
- Font 3 (Right):** Shows a large blue square with a white center. The left panel lists GID values from 50 to 72. The right panel shows a tree view of font resources, including 'Table Directory', 'cmap', 'cvt', 'head', 'hhea', 'hmtx', 'loca', 'maxp', 'name', 'post', 'fpgm', 'prep', and 'glyf'. The 'glyf' section shows a tag 'Tag: 0x676C796' and entries for 'GID: 0 ".notdef"', 'Empty Glyphs: [1]', 'GID: 2 "quotedbl"', 'GID: 3 "dollar"', and 'GID: 4 "quoteright"'.

What is a subset font?

- Fully embedded means ALL characters are contained in the PDF
- Subsetting reduces the font to used glyphs

Questions to answer

- ✓ What is
 - ✓ a glyph?
 - ✓ a .notdef glyph?
 - ✓ an empty glyph?
- ✓ What is a subset font?
- What is a font encoding?
- Which font types can be used in a PDF?
- What is a “symbolic” font?
- What means "Mapping to Unicode"?

Overview

- Font tools in pdfToolbox
- Fonts in PDF
- Issues with glyph selection
- More font tools in pdfToolbox



How a page description in PDF makes use of fonts

- Page description (content stream) selects font through an internal name:
`/F25 12 Tf`
- Internal name points to font resource object (for current page):
`Resources << Font << /F25 12 0 R >> >>`
- Font resource object contains:
 - Information about the font, for example
 - Base name (example: /Courier)
 - Font type (example: /TrueType)
 - Character widths (example: [593 615 572 ...])
 - Optionally: the actual font itself (= “font is embedded”)
 - Data stream (for example in /FontFile ...)
 - Optionally: ToUnicode table for mapping to Unicode

pdfToolbox’ “Explore PDF”

- Internal name
- Font resource object
- Font resource object contains:
 - Information about the font
 - Base name
 - Font type
 - Character widths
 - Optionally: the actual font itself
 - Optionally: ToUnicode table

Font - a four letter word beginning with "F".pdf - Internal PDF structure

The screenshot shows a hierarchical tree view of a PDF font resource object. The root node is 'Ty1: (8) [18 0 R] /T:Font /S:Type1 (WSYHHT+QuadraatSans-Bold) Type1'. A red arrow points from the 'Internal name' bullet point to this node. Another red arrow points from the 'Font resource object' bullet point to the same node. A blue arrow points from the 'Information about the font' bullet point to the 'FontDescriptor' node, which is highlighted with a green background. A red arrow points from the 'Optional: the actual font itself' bullet point to the 'FontName: WSYHHT+QuadraatSans-Bold' entry under the 'FontDescriptor' node. A blue arrow points from the 'Optional: ToUnicode table' bullet point to the 'Widths: (185)' entry at the bottom of the list. The tree view also includes nodes for 'BaseFont', 'Encoding', 'FirstChar', 'FontFile', 'Filter', 'Length', 'Length1', 'Length2', 'Length3', 'Font stream', 'ItalicAngle', 'Leading', 'MaxWidth', 'StemH', 'StemV', 'Type', and 'Subtype'.

- Ty1: (8) [18 0 R] /T:Font /S:Type1 (WSYHHT+QuadraatSans-Bold) Type1
 - BaseFont: WSYHHT+QuadraatSans-Bold
 - Encoding: MacRomanEncoding
 - FirstChar: 32
- FontDescriptor: (14) [428 0 R] /T:FontDescriptor
 - Ascent: 880
 - CapHeight: 655
 - Descent: -224
 - Flags: 32
 - FontBBox: (4)
 - FontFile: (5) [429 0 R]
 - Filter: FlateDecode
 - Length: 9639 [430 0 R]
 - Length1: 1140 [431 0 R]
 - Length2: 8844 [432 0 R]
 - Length3: 532 [433 0 R]
 - Font stream
 - Type1 (PostScript) - based font
 - FontName: WSYHHT+QuadraatSans-Bold
 - ItalicAngle: 0
 - Leading: 199
 - MaxWidth: 1406
 - StemH: 82
 - StemV: 137
 - Type: FontDescriptor
 - XHeight: 472
 - LastChar: 216
 - Subtype: Type1
 - Type: Font
 - Widths: (185)

How a page description in PDF refers to glyphs (encoding, glyph selection, glyph mapping)

1. Character code in the page description (a number, one or several bytes)
2. Explicit or implicit mapping, possibly in multiple steps
 - If present via “Encoding” entry in PDF font object
 - Or via encoding defined in the embedded font (built-in encoding)
3. Glyph in actual font by its glyph name or glyph ID
(the outline that actually gets painted)



How a page description in PDF refers to glyphs (encoding, glyph selection, glyph mapping)

1. Character code in the page description (a number, one or several bytes)
2. Explicit or implicit mapping, possibly in multiple steps
 - If present via “Encoding” entry in PDF font object
 - Or via encoding defined in the embedded font (built-in encoding)
3. Glyph in actual font by its glyph name or glyph ID
(the outline that actually gets painted)

Examples:

- Type 1 character code → MacRomanEncoding → PostScript glyph name →
 glyph name to glyph in Type 1 font
- TrueType character code → custom encoding → PostScript glyph name →
 Adobe glyph list (maps glyph names to *Unicode*) → Unicode cmap in TrueType font →
 glyph ID
- CID Type 2 character code → Identity-H → CIDToGIDmap → glyph ID

Font types

- **Simple fonts**

1 byte character codes can address max. 256 glyphs

- Type 1
 - Type 1 (classic)
 - CFF Type 1 (compact font format)
- TrueType (developed by Adobe and Microsoft in the 80th as a competitor to Type 1)
- Type 3 (not a real font – contains PDF code for each glyph)
- OpenType (internally either Type 1/CFF flavor or TrueType flavor)

- **Composite fonts**

1 or 2 byte character codes can address max. 65535 glyphs

- Type 0 (Type 1 or OpenType/Type 1/CFF based)
- Type 2 (TrueType or OpenType/TrueType based)



TrueType

- All font types are either Type 1 (PostScript) like or TrueType like
- TrueType has advanced features (which sometimes cause font issues)
 - Internally more than one encoding possible via “cmaps” (different from “CMaps” in CID fonts)
 - An embedding flag specifies whether a font may be embedded
 - A *symbolic flag* specifies whether it merely contains symbolic characters (Zapf Dingbats, Windings etc.)
 - If non-symbolic the font should specify a standard encoding (MacRoman or WinAnsi) in the PDF
 - If symbolic all encoding information in the PDF part is ignored
 - The .notdef is always on CID 0 (because there are no names)



What is a "symbolic" font?

- A ‘symbolic’ fonts contains symbols
(and no “regular” characters)
- Typical ‘symbolic’ fonts are:
 - Symbol abc → αβχ
 - Zapf Dingbats abc → ☼☀☼
 - Wingdings abc → ☽❀☽
- ▶ A ‘symbolic’ TrueType font can’t reasonably use standard encoding like MacRoman or WinAnsi (because those contain only “normal” characters)

Mapping to Unicode (encoding for semantics)

- Text search, copying of text
- Optional – *not* required for painting glyphs
- Options of mapping to Unicode:
 - Implicitly by standard (glyph selection) encodings in the PDF (example: MacRomanEncoding)
 - Encoding information in the actual font (“ToUnicode” cmap in TrueType fonts)
 - By glyph names in the actual font (in Type 1 fonts) via the Adobe Glyph List (maps glyph names to Unicode)
 - Explicitly by a ToUnicode table in the PDF data structure



Mapping to Unicode in pdfToolbox' font report

- Text search, copying of text
- Optional – not required for painting glyphs
- Options of mapping to Unicode:
 - Implicitly by standard (glyph selection
(example: [MacRomanEncoding](#))
 - Encoding ('ToUnicode')
 - By glyph name
 Glyph List
 - Explicitly by

Inventory report for Document:"Font - a four letter word beginning with "F".pdf"



Font 1 (out of 11)
on pages 2-3, 5-7, 11-12, 14-15, 19, 25-26, 29-32, 34-45

Glyphs used in PDF:

.	1	2	3	•	—
U+002E CID 46 period	U+0031 CID 49 one	U+0032 CID 50 two	U+0033 CID 51 three	U+2022 CID 165 bullet	U+2013 CID 208 endash

Font properties:
TrueType font: "ArialMT"
PostScript name: ArialMT
Encoding: MacRomanEncoding
Font is a subset
Embedded: True
Font descriptors
 Maximum height above the baseline: 905
 y-coordinate of the top of flat capital letters: 716
 Maximum depth below the baseline: -212
 Italic angle in degrees counterclockwise: 0
 Width of the dominant vertical stems of glyphs: 95
 Average width of characters: 441
 Desired spacing between lines: 33
 Maximum width of characters: 2000
 Width to use for un-encoded characters.: 0
 Vertical width of the dominant horizontal stems of glyphs: 84
 y-coordinate of the top of flat non-ascending lowercase letters: 519
 PostScript name: NMNAGD+ArialMT
Various properties
 Fixed-width font: False
 Serif font: False
 Symbolic font: False

Broken mapping to Unicode (encoding for semantics)

#

U+0023
CID 35
NUMBER
SIGN

(6)

U+0024
CID 36
DOLLAR
SIGN

bb

U+0025
CID 37
PERCENT
SIGN

Questions to answer

- ✓ What is
 - ✓ a glyph?
 - ✓ a .notdef glyph?
 - ✓ an empty glyph?
- ✓ What is a subset font?
- ✓ What is a font encoding?
- ✓ Which font types can be used in a PDF?
- ✓ What is a “symbolic” font?
- ✓ What means "Mapping to Unicode"?

Overview

- Font tools in pdfToolbox
- Fonts in PDF
- Issues with glyph selection
- More font tools in pdfToolbox

Missing glyph

- Page description references a non-existing glyph in a font
- Possible reasons:
 - old vs. new font version (example: only new version has the Euro sign "€")
 - incorrect subset font: subsetting algorithm stripped too much, or stripped the wrong glyphs

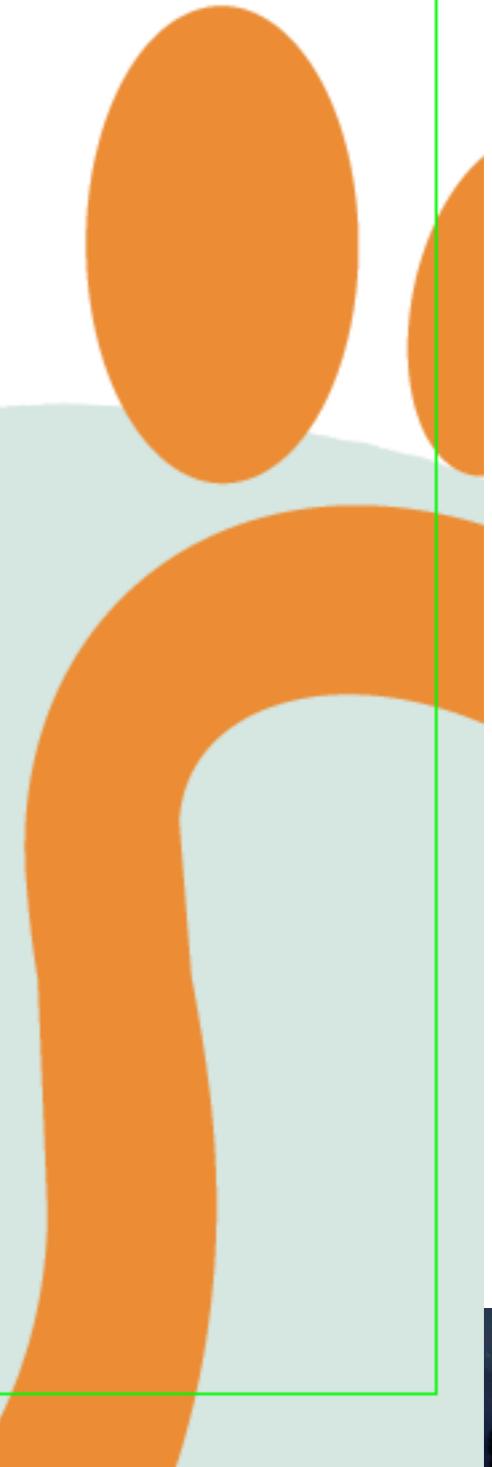


Back to the support case



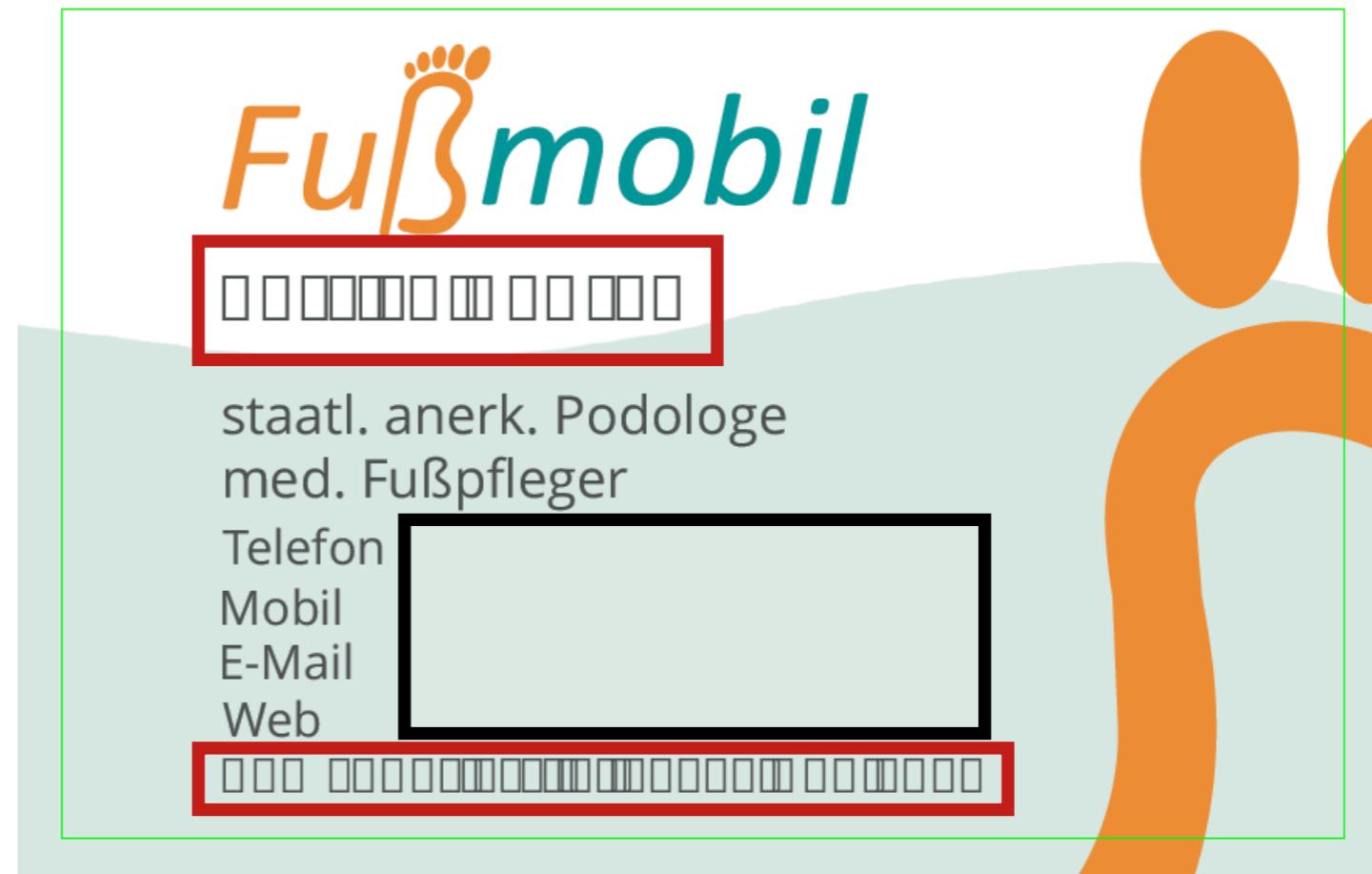
staatl. anerk. Podologe
med. Fußpfleger

Telefon
Mobil
E-Mail
Web



“OpenSans-Bold” TrueType font
Indicated as symbolic font but PDF defines WinAnsiEncoding

- PDF specification:
 - PDF encoding entries are ignored for symbolic fonts
 - If no proper encoding present results are application dependent
- We adjusted fixups and font engine



ncona AVIGNE
arbedor Baskerville Bodoni BROADW
Caslon Century Old Style Cheltenham Clearface COPPER

Wir haben etwas gegen teure Schriften

infinityType

für Mac & PC

downloaden
20 Schriften

www.infinityType.com
Infos unter 09329 00 00 00

3

- 9.999 TrueType-Schriften für PC
- 9.999 PostScript-Schriften für Mac und PC
- Riesige Auswahl bekannter Schriften
- Höchste Qualität für Design, Web und Print
- Individuelle €uro-Zeichen in allen Schriften

Plus: alle Schriften einbettbar in PDF

Plus: alle Schriften sofort einsetzbar ohne Freischaltung

Plus: alle Schriften systemweit verwendbar auf Mac und PC

Plus: alle Schriften vollständig portabel zwischen Mac und PC

Plus: viele Schriften mit osteuropäischen und türkischen Zeichensätzen

Einführungspreis: nur € 229,- statt € 429,-

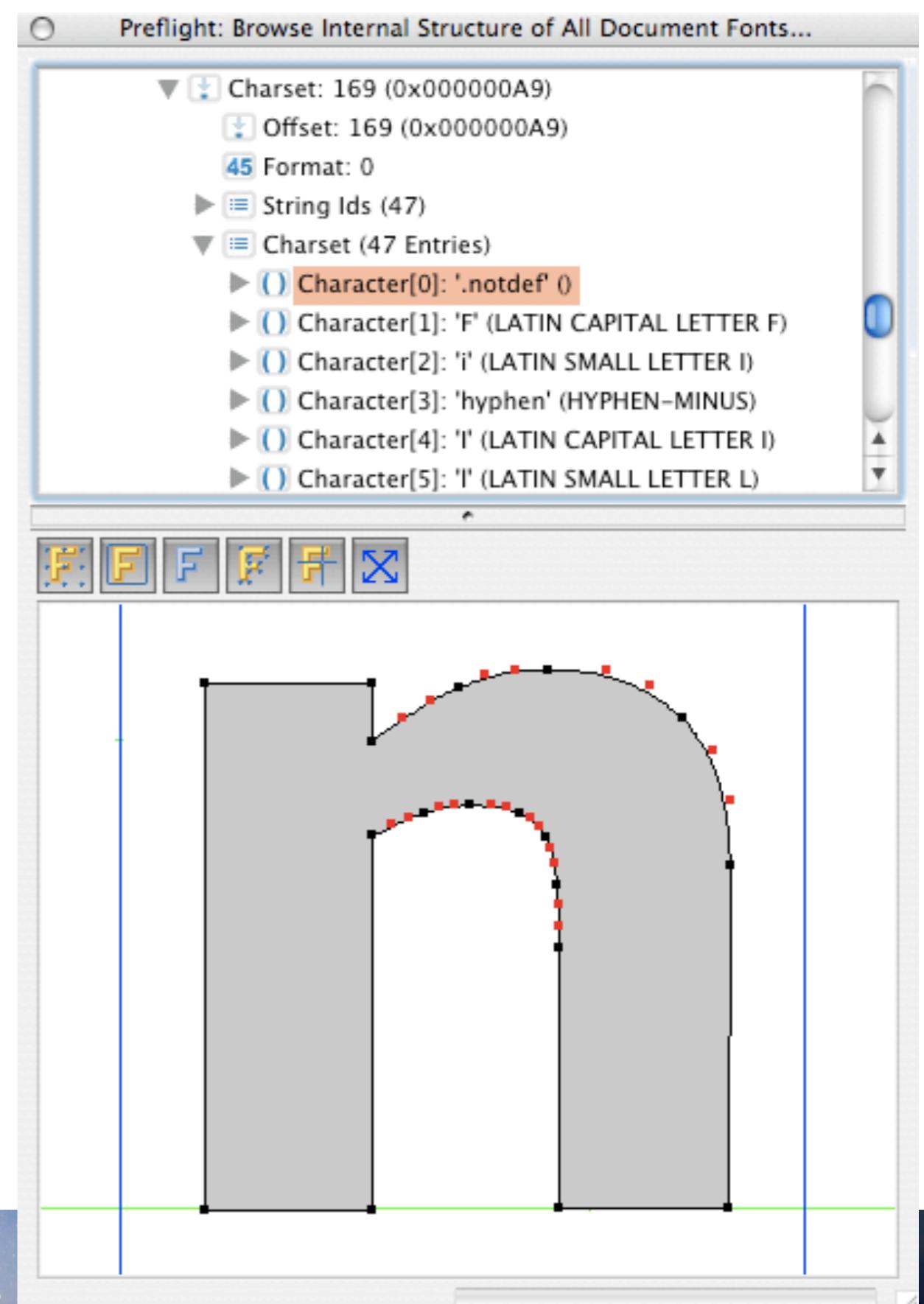


Open Source Schriften

www.infinityType.com
Infos unter 99XXXX99

Einführungspr

.notdef is not a notdef



Glyph substitution

- Character code and encoding do not match each other properly
- Example: a PDF creation software “thinks” in WinAnsiEncoding, but a font is embedded using MacRomanEncoding
 - Umlaut and accented character show up as some other character
 - Glyphs ‘swap places’





DEPARTMENT OF THE TREASURY
UNITED STATES CUSTOMS SERVICE

Zollerklärung

19 CFR 122.27, 148.12, 148.13, 148.110, 148.111, 1498; 31 CFR 5316

FORMULAR GENEHMIGT
OMB-NR. 1515-0041

Jeder ankommende Reisende oder das zuständige Familienmitglied muss die nachfolgenden Angaben machen (es ist lediglich EINE schriftliche Zollerklärung pro Familie erforderlich):

1. Familienname

Vorname (Rufname)

Zweiter Vorname

2. Geburtsdatum Tag

Monat

Jahr

3. Anzahl der mit Ihnen reisenden Familienmitglieder

4. (a) Anschrift in den USA (Name des Hotels/Zielort)

(b) Ort

(c) Bundesstaat

5. Ausstellung des Reisepasses durch (Land)

6. Reisepassnummer

7. Land des Wohnsitzes

8. Folgende Länder wurden auf dieser Reise
vor der Ankunft in den USA besucht

9. Fluggesellschaft/Flugnummer oder Name der Schiffe

10. Der Hauptzweck dieser Reise ist geschäftlicher Natur:

Ja Nein

11. Ich führe (wir führen) Folgendes ein

(a) Früchte, Pflanzen, Lebensmittel, Insekten:

Ja Nein

(b) Fleisch, Tiere, Tier- oder Wildprodukte:

Ja Nein

(c) Krankheitserreger, Zellkulturen, Schnecken:

Ja Nein

(d) Erde oder ich war (wir waren) auf einem
landwirtschaftlichen Betrieb/einer Weide:

Ja Nein

12. Ich habe mich (wir haben uns) in unmittelbarer Nähe eines
Viehbestands aufgehalten (darunter fällt z.B. das Berühren von
und der Umgang mit Tieren):

Ja Nein

13. Ich führe (wir führen) **Geld, Zahlungsmittel oder ausländische
Geldmittel** im Wert von mehr als US\$ 10.000 mit mir (uns):
(Eine Definition von Zahlungsmittel ist auf der Rückseite des
Formulars zu finden)

Ja Nein

14. Ich führe (wir führen) **Handelswaren** mit mir (uns):

Ja Nein

(Zum Verkauf bestimmte Waren, Warenmuster zur Kundenan-
werbung oder andere Güter, die nicht als persönliche Gebrauchsgegenstände gelten.)

15. **Einwohner** – Der **Gesamtwert aller Güter**, einschließlich Handelswaren, die ich (wir) im
Ausland gekauft oder erworben habe (haben) (darunter fallen auch Geschenke für jemand
anders, jedoch nicht in die USA versandte Gegenstände) und in die USA einführe
(einführen), beläuft sich auf: \$

Besucher – Der **Gesamtwert aller Gegenstände**, die in den USA verbleiben, einschließlich
Handelswaren, beläuft sich auf: \$

Lesen Sie bitte die Anleitungen auf der Rückseite dieses Formulars. Dort ist auch Platz zur
Auflistung aller von Ihnen anzumeldenden Gegenstände vorgesehen.

**ICH HABE DIE WICHTIGEN HINWEISE AUF DER RCKSEITE DIESES
FORMULARS GELESEN UND WAHRHEITSGETREUE ANGABEN GEMACHT.**

X

(Unterschrift)

Nur für amtliche Vermerke

Datum (Tag/Monat/Jahr)



DEPARTMENT OF THE TREASURY
UNITED STATES CUSTOMS SERVICE

Zollerklärung

19 CFR 122.27, 148.12, 148.13, 148.110, 148.111, 1498; 31 CFR 5316

FORMULAR GENEHMIGT
OMB-NR. 1515-0041

Jeder ankommende Reisende oder das zuständige Familienmitglied muss die nachfolgenden Angaben machen (es ist lediglich EINE schriftliche Zollerklärung pro Familie erforderlich):

1. Familienname

Vorname (Rufname)

Zweiter Vorname

2. Geburtsdatum Tag

Monat

Jahr

3. Anzahl der mit Ihnen reisenden Familienmitglieder

4. (a) Anschrift in den USA (Name des Hotels/Zielort)

(b) Ort

(c) Bundesstaat

5. Ausstellung des Reisepasses durch (Land)

11. Ich habe (wir haben) Folgendes ein:

(a) Früchte, Pflanzen, Lebensmittel, Insekten:

Ja Nein

(b) Fleisch, Tiere, Tier- oder Wildprodukte:

Ja Nein

(c) Krankheitserreger, Zellkulturen, Schnecken:

Ja Nein

(d) Erde oder ich war (wir waren) auf einem
landwirtschaftlichen Betrieb/einer Weide:

Ja Nein

12. Ich habe mich (wir haben uns) in unmittelbarer Nähe eines
Viehbestands aufgehalten (darunter fällt z.B. das Berühren von
und der Umgang mit Tieren):

Ja Nein

13. Ich führe (wir führen) **Geld, Zahlungsmittel oder ausländische
Geldmittel** im Wert von mehr als US\$ 10.000 mit mir (uns):
(Eine Definition von Zahlungsmittel ist auf der Rückseite des
Formulars zu finden)

Ja Nein

14. Ich führe (wir führen) **Handelswaren** mit mir (uns):

Ja Nein

(Zum Verkauf bestimmte Waren, Warenmuster zur Kundenan-
werbung oder andere Güter, die nicht als persönliche Gebrauchsgegenstände gelten.)

15. **Einwohner** – Der **Gesamtwert aller Güter**, einschließlich Handelswaren, die ich (wir) im
Ausland gekauft oder erworben habe (haben) (darunter fallen auch Geschenke für jemand
anders, jedoch nicht in die USA versandte Gegenstände) und in die USA einführe
(einführen), beläuft sich auf: \$

Besucher – Der **Gesamtwert aller Gegenstände**, die in den USA verbleiben, einschließlich
Handelswaren, beläuft sich auf: \$

Lesen Sie bitte die Anleitungen auf der Rückseite dieses Formulars. Dort ist auch Platz zur
Auflistung aller von Ihnen anzumeldenden Gegenstände vorgesehen.

**ICH HABE DIE WICHTIGEN HINWEISE AUF DER RÜCKSEITE DIESES
FORMULARS GELESEN UND WAHRHEITSGETREUE ANGABEN GEMACHT.**

X

(Unterschrift)

Datum (Tag/Monat/Jahr)

Nur für amtliche Vermerke

Folgen Sie dem Sommer ans Mittelmeer !

Kroatien - Rabac

30.07.-08.08.07

10 Tg. HP ab **699,-**

Italienische Blumenriviera

23.07.-29.07./31.08.-06.09.07

7 Tg. HP nur **419,-**

Spanien - Costa Brava

z.B. Lloret de Mar am 03.08. und 10.08.07 10 Tg. (7xHP) ab **562,-**

17 Tg. (14xHP) ab **910,-**

Kinderfestpreis Spanien nur 20,- EUR

pro Woche Aufenthalt (gilt für Kinder von 3-12 Jahren bei Unterbringung mit 2 Erwachsenen im DZ mit Zustellbett)

Folgen Sie dem Sommer ans Mittelmeer !

KroatieK - Rabac

30.07.-08.08.07

10 Tg. HP ab 699,-

ItalieKische BlumeKriviera

23.07.-29.07./31.08.-06.09.07

7 Tg. HP Kur 419,-

SpaKieK - Costa Brava

z.B. Lloret de Mar am 03.08. uKd 10.08.07 10 Tg. (7xHP) ab **562,-**

17 Tg. (14xHP) ab 910,-

KiKderfestpreis SpaKieK Kur 20,- EUR

pro Woche Aufenthalt (gilt für Kinder von 3-12 Jahren bei Unterbringung mit 2 Erwachsenen im DZ mit Zustellbett)

Reasons for glyph substitution

1. The PDF is the problem

- Encodings used incorrectly when producing PDF
- Encodings used incorrectly when rewriting PDF
(optimization, editing, merging fonts, ...)
- .notdef is used and contains a needed glyph



Reasons for glyph substitution

2. The PDF is not the problem

- Broken caching of fonts in PDF viewers, output systems (RIPs)
- Multiple instances of same original font
 - as different subsets
 - as non-identical subsets with the same name
(ABCDEF+BaseFontName)
 - Sometimes happens when two or more documents have been merged
- Processor establishes one instance, when ‘seeing’ the next instance it assumes it is identical to the first instance and uses first instance
- pdfToolbox’ “Fix potential font problems” includes a Fixup that adds a unique prefix to each font name to avoid this

Overview

- Font tools in pdfToolbox
- Fonts in PDF
- Issues with glyph selection
- More font tools in pdfToolbox

Font tools in pdfToolbox

- ✓ Profiles: Validate for PDF/X-4 or PDF/A-2
- ✓ Profile: List potential font problems
 - Subset fonts
 - Font report



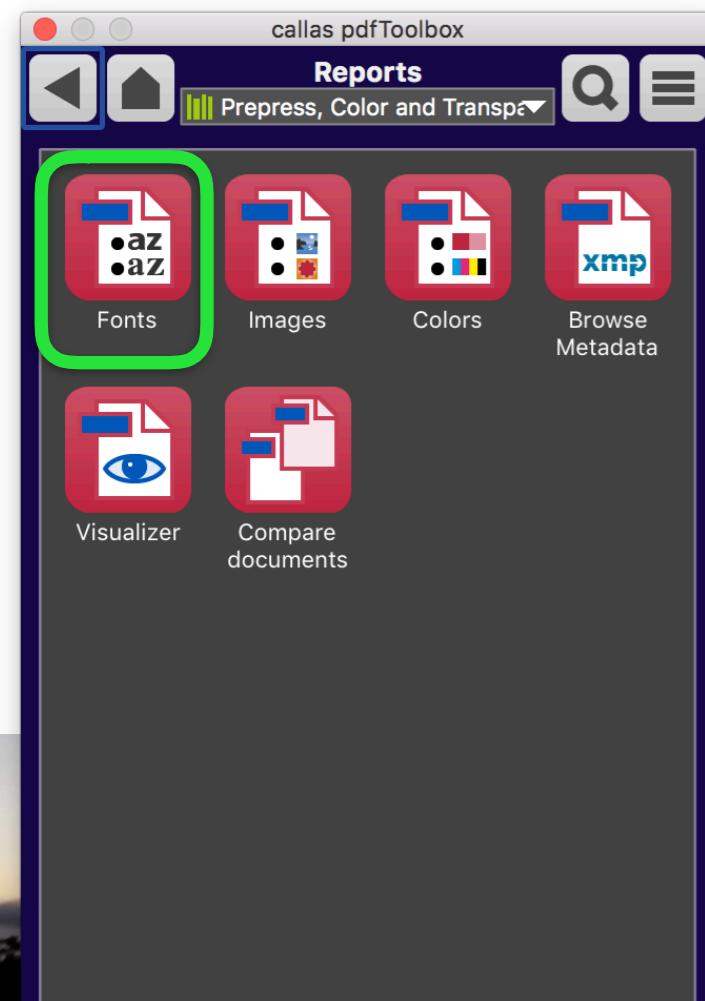
pdfToolbox' Fixup "Subset fonts" performs a rewrite of embedded fonts

- Even for existing subset fonts
- The font will be completely rewritten by pdfToolbox' font engine



Font report

- Can be found in the Switchboard (Reports)
- Creates an overview over all glyphs used in a font and lists a lot of details



Inventory report for Document:"CCW90000171028194822"

Font 1 (out of 2)
on page 1

Glyphs used in PDF:

()	-	.	/	0	1	
U+0020 CID 32 space	U+0028 CID 40 parenleft	U+002D CID 41 hyphen	U+002E CID 46 period	U+002F CID 47 slash	U+0030 CID 48 zero	U+0031 CID 49 one	
2	5	7	A	B	D	F	
U+0032 CID 50 two	U+0035 CID 53 five	U+0037 CID 55 seven	U+0041 CID 65 A	U+0042 CID 66 B	U+0044 CID 68 D	U+0046 CID 70 F	
P	S	T	V	a	b	c	d
U+0050 CID 80 P	U+0053 CID 83 S	U+0054 CID 94 T	U+0058 CID 98 V	U+0061 CID 97 a	U+0062 CID 98 b	U+0063 CID 99 c	U+0064 CID 100 d
e	f	g	h	i	l	m	n
U+0065 CID 101 e	U+0068 CID 102 f	U+0067 CID 103 g	U+0066 CID 104 h	U+0069 CID 105 i	U+0060 CID 108 l	U+0060 CID 109 m	U+0066 CID 110 n
o	p	r	s	t	u	v	w
U+006F CID 111 o	U+0070 CID 112 p	U+0073 CID 114 r	U+0070 CID 115 s	U+0074 CID 116 t	U+0075 CID 117 u	U+0076 CID 118 v	U+0077 CID 119 w
y	©	.					
U+0079 CID 121 y	U+0046 CID 169 ©	U+0057 CID 225 .					

Font properties:
TrueType font: "Verdana"
PostScript name: Verdana
Encoding: MacRomanEncoding
Font is a subset
Embedded: True
Font descriptors
Maximum height above the baseline: 1005
y-coordinate of the top of flat capital letters: 727
Maximum depth below the baseline: -210
Italic angle in degrees counterclockwise: 0
Width of the dominant vertical stems of glyphs: 97
Average width of characters: 508
Desired spacing between lines: 0
Maximum width of characters: 1521
Width to use for un-encoded characters.: 0
Vertical width of the dominant horizontal stems of glyphs: 86
y-coordinate of the top of flat non-ascending lowercase letters: 545
PostScript name: VDPMAI+Verdana
Various properties
Fixed-width font: False
Serif font: False
Symbolic font: False
Script font: False

Glyphs used in PDF:

()	-	.	/	0	1	
U+0020 CID 32 space	U+0028 CID 40 parenleft	U+002D CID 41 hyphen	U+002E CID 46 period	U+002F CID 47 slash	U+0030 CID 48 zero	U+0031 CID 49 one	
2	5	7	A	B	D	F	
U+0032 CID 50 two	U+0035 CID 53 five	U+0037 CID 55 seven	U+0041 CID 65 A	U+0042 CID 66 B	U+0044 CID 68 D	U+0046 CID 70 F	
P	S	T	V	a	b	c	d
U+0050 CID 80 P	U+0053 CID 83 S	U+0054 CID 94 T	U+0058 CID 98 V	U+0061 CID 97 a	U+0062 CID 98 b	U+0063 CID 99 c	U+0064 CID 100 d
e	f	g	h	i	l	m	n
U+0065 CID 101 e	U+0068 CID 102 f	U+0067 CID 103 g	U+0066 CID 104 h	U+0069 CID 105 i	U+0060 CID 108 l	U+0060 CID 109 m	U+0066 CID 110 n
o	p	r	s	t	u	v	w
U+006F CID 111 o	U+0070 CID 112 p	U+0073 CID 114 r	U+0070 CID 115 s	U+0074 CID 116 t	U+0075 CID 117 u	U+0076 CID 118 v	U+0077 CID 119 w
y	©	.					
U+0079 CID 121 y	U+0046 CID 169 ©	U+0057 CID 225 .					



Checking for and analyzing font issues

- ✓ Profiles: Validate for PDF/X-4 or PDF/A-2
- ✓ Profile: List potential font problems
- ✓ Subset fonts
- ✓ Font report
- There are more pdfToolbox tools for in depth analysis of inner PDF and font structures (mentioned before):
 - Explore Fonts
 - Explore PDF -> Resource View



Recap

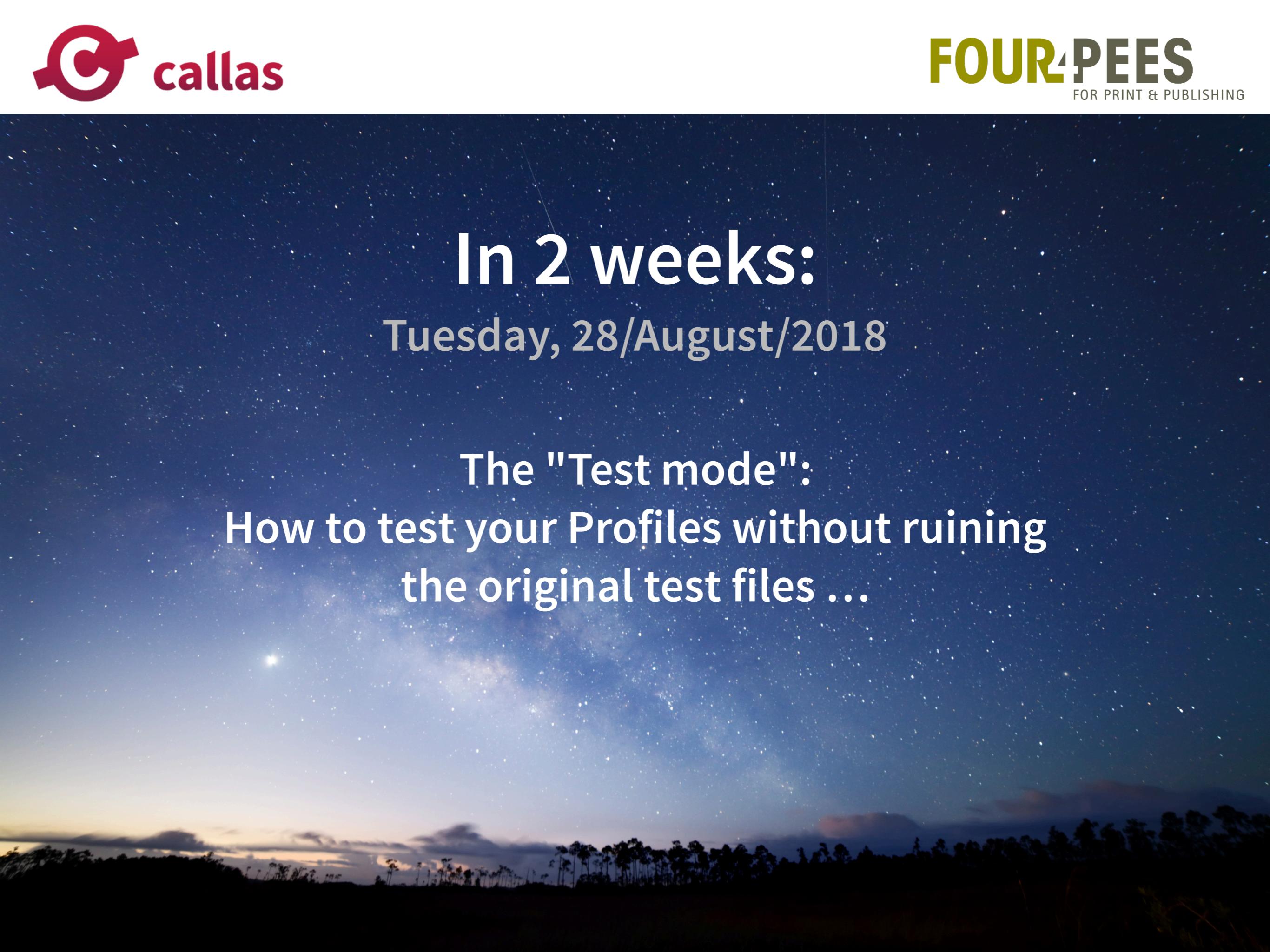
- Font tools in pdfToolbox
- Fonts in PDF
- Issues with glyph selection
- More font tools in pdfToolbox





That's it!

Questions?



In 2 weeks:

Tuesday, 28/August/2018

The "Test mode":
How to test your Profiles without ruining
the original test files ...

Thank you!

support@callassoftware.com

help.callassoftware.com

blog.callassoftware.com