```
Why We Do Not Accept Author-Generated
                                           PDFs
                                                     /FirstChar 44 /LastChar 121
722 667 667 722 611 556 722 722 333 389 722 611 889 722 722 5Volker56
                                                    Schaa R /Encoding 7 0 R
                            Gesellschaft für Schwerionenforschung mbH
                                       Darmstadt, Germany
                                  JACoW Team Meeting,
                                      Shanghai, China
                                          Nov 201 400022316 00000 n 0000022482 00000 n 0000022649 00000 n 0000022143 00000 n
```

```
To ensure special settings like
```

```
To ensure special settings like

    Acrobat 7 compatibility (PDF 1.6)
```

```
To ensure special settings like

    Acrobat 7 compatibility (PDF 1.6)

    Images Downsampled to 300/1200 dpi (color/monochrome)
```

```
To ensure special settings like

    Acrobat 7 compatibility (PDF 1.6)

    Images Downsampled to 300/1200 dpi (color/monochrome)

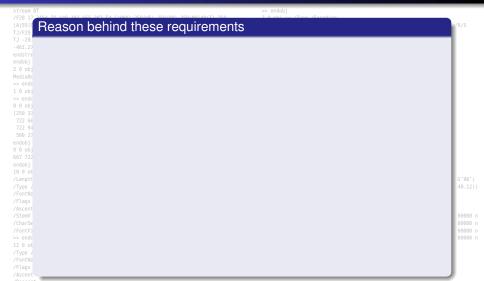
    Font embedding for all fonts

    (even for the 35 of the Core Font Set).
```

```
To ensure special settings like
```

- Acrobat 7 compatibility (PDF 1.6)
- Images Downsampled to 300/1200 dpi (color/monochrome)
- Font embedding for all fonts (even for the 35 of the Core Font Set).
- Ensure JACoW paper size.

Sustainability



Sustainability

Reason behind these requirements

 JACoW wants to ensure PDF/A compatibility which is important for the digital preservation of electronic documents (long-term archival, sustainability).

- JACoW wants to ensure PDF/A compatibility which is important for the digital preservation of electronic documents (long-term archival, sustainability).
- PDF/A differs from PDF by omitting features which are ill-suited to long-term archiving, such as font linking (as opposed to font embedding).

1'00') 40.12))

> 9000 n 9000 n

> 9000 n 9000 n

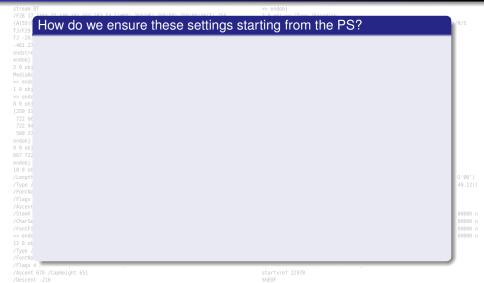
00 n

Volker RW Schaa

- JACoW wants to ensure PDF/A compatibility which is important for the digital preservation of electronic documents (long-term archival, sustainability).
- PDF/A differs from PDF by omitting features which are ill-suited to long-term archiving, such as font linking (as opposed to font embedding).
- Key elements to PDF/A compatibility are:
 - Audio and video content are forbidden.
 - JavaScript and executable file launches are forbidden.
 - All fonts must be embedded. This also applies to the PostScript standard fonts such as Times/Helvetica/etc (Basic 35).
 - Encryption is forbidden.
 - External content references are forbidden (no links!).
 - ...

1'00') 40.12))

10000 n



How do we ensure these settings starting from the PS? We set the PDF version=1.6 when distilling the PS or when the PDF Optimizer is run.

- We set the PDF version=1.6 when distilling the PS or when the PDF Optimizer is run.
- Images are downsampled when distilling the PS or when the PDF Optimizer is run.

- We set the PDF version=1.6 when distilling the PS or when the PDF Optimizer is run.
- Images are downsampled when distilling the PS or when the PDF Optimizer is run.
- Fonts must be embedded in the PS, otherwise we cannot embed them as this is not foreseen by Acrobat Distiller.

- We set the PDF version=1.6 when distilling the PS or when the PDF Optimizer is run.
- Images are downsampled when distilling the PS or when the PDF Optimizer is run.
- Fonts must be embedded in the PS, otherwise we cannot embed them as this is not foreseen by Acrobat Distiller.
 - No chance to fix this in the current setup.





- We set the PDF version=1.6 when distilling the PS or when the PDF Optimizer is run.
- Images are downsampled when distilling the PS or when the PDF Optimizer is run.
- Fonts must be embedded in the PS, otherwise we cannot embed them as this is not foreseen by Acrobat Distiller.
- No chance to fix this in the current setup.
- JACoW paper size is enforced by executing the PitStop Action List or
 - when the PDF Optimizer is run.

The following options can be guaranteed using PDF Optimizer on a PDF file PDF version compatibility Downsampling JACoW paper size

The following options can be guaranteed using PDF Optimizer on a PDF file

- PDF version compatibility
- Downsampling
- JACoW paper size

Only on PS files with the correct font setup (JACoW PS Printer)

E . E . L . I . .

Font Embedding

Motivation

Ensuring the JACoW and PDF/A settings

The following options can be guaranteed using PDF Optimizer on a PDF file

- PDF version compatibility
- Downsampling
- JACoW paper size

Only on PS files with the correct font setup (JACoW PS Printer)

Font Embedding

Do we fulfill PDF/A standard for our conferences?

Not embedding fonts means that we do not produce our PDFs according to PDF/A standards, which makes the portability and long-term archiving problematic or even questionable.

```
The number of papers which do not have fonts embedded
```

IPAC2011 had 515 out of 1245 papers, that means 41.4 %.

But that's much better than all other conferences before, i. e.

EPAC2008 had 830 papers out of 1229, that's 67.5%.

The number of papers which do not have fonts embedded

IPAC2011 had 515 out of 1245 papers, that means 41.4 %.

But that's much better than all other conferences before, i. e.

• EPAC2008 had 830 papers out of 1229, that's 67.5 %.

Why better?

We can see that since Raphael published the JACoW PS driver/printer the numbers are getting better.

or

Has the overall quality of the source/PS files delivered by authors grown so much worse that editors have to do a lot more re-generating PS files (number of initial odots?).

</p

```
So what do we see?
```



```
So what do we see?

    People who deliver directly usable PostScript files AND have the

    fonts embedded is still too low.

    It is hard to enforce the usage of our PS printer.
```

So what do we see?

 People who deliver directly usable PostScript files AND have the fonts embedded is still too low. It is hard to enforce the usage of our PS printer. IPAC2011 had 871 papers produced on Windows (70.0%), where the use of the JACoW PS printer would be possible.

So what do we see?

- People who deliver directly usable PostScript files AND have the fonts embedded is still too low.
- It is hard to enforce the usage of our PS printer.
- IPAC2011 had 871 papers produced on Windows (70.0 %), where the use of the JACoW PS printer would be possible.
- What about the 223 Mac, 145 Linux, 12 Unix papers (30.0%)?

So what do we see?

- People who deliver directly usable PostScript files AND have the fonts embedded is still too low.
- It is hard to enforce the usage of our PS printer.
- IPAC2011 had 871 papers produced on Windows (70.0 %), where the use of the JACoW PS printer would be possible.
- What about the 223 Mac, 145 Linux, 12 Unix papers (30.0%)?

What can we change?

Volker RW Schaa

Why We Do Not Accept Author-Generated PDFs

So what do we see?

- People who deliver directly usable PostScript files AND have the fonts embedded is still too low.
- It is hard to enforce the usage of our PS printer.
- IPAC2011 had 871 papers produced on Windows (70.0 %), where the use of the JACoW PS printer would be possible.
- What about the 223 Mac, 145 Linux, 12 Unix papers (30.0%)?

What can we change?

not much?

not much

CATE1 22370

So what do we see?

- People who deliver directly usable PostScript files AND have the fonts embedded is still too low.
- It is hard to enforce the usage of our PS printer.
- IPAC2011 had 871 papers produced on Windows (70.0 %), where the use of the JACoW PS printer would be possible.
- What about the 223 Mac, 145 Linux, 12 Unix papers (30.0%)?

What can we change?

- not much?
- I think we can! Yes!!

Volker RW Schaa

Why We Do Not Accept Author-Generated PDFs

So what do we see?

- People who deliver directly usable PostScript files AND have the fonts embedded is still too low.
- It is hard to enforce the usage of our PS printer.
- IPAC2011 had 871 papers produced on Windows (70.0 %), where the use of the JACoW PS printer would be possible.
- What about the 223 Mac, 145 Linux, 12 Unix papers (30.0%)?

What can we change?

- not much?
- I think we can! Yes!!
- I'm sure it might get much better if we accept PDF from authors!

So what do we see?

- People who deliver directly usable PostScript files AND have the fonts embedded is still too low.
- It is hard to enforce the usage of our PS printer.
- IPAC2011 had 871 papers produced on Windows (70.0 %), where the use of the JACoW PS printer would be possible.
- What about the 223 Mac, 145 Linux, 12 Unix papers (30.0%)?

What can we change?

- not much?
- I think we can! Yes!!
- I'm sure it might get much better if we accept PDF from authors!
 - so let me change the title of my talk and justify my claim.

Reasoning **Required Steps** Conclusion

Stev-by-Step to PDF/A conforming PDFs

Why Don't We Accept Author-Generated PDFs?

```
722 667 667 722 611 556 722 722 333 389 722 611 889 722 722 5Volker56
```

```
Schaa R /Encoding 7 0 R
Gesellschaft für Schwerionenforschung mbH
```

Darmstadt, Germany

/FirstChar 44 /LastChar 121

JACoW Team Meeting,

```
Shanghai, China
     Nov 201 480022316 00000 n 0000022482 00000 n 0000022649 00000 n 0000022143 00000 n
```

Why Don't We Accept Author-Generated PDFs?

Volker RW Schaa

Reasoning Required Steps Conclusion

Stey-by-Step to PDF/A conforming PDFs

Reasoning pro author PDFs

Reasoning Required Steps Conclusion

Stey-by-Step to PDF/A conforming PDFs

Reasoning pro author PDFs

```
I've tested this approach since 2009 when I told LATEX people to
upload their pdfLTFX generated PDFs as "Other supporting files".
What I found:
```

Reasoning pro author PDFs

I've tested this approach since 2009 when I told LATEX people to upload their pdfLATEX generated PDFs as "Other supporting files". What I found:

 The number of directly usable papers with uploaded PDFs is much higher than the same papers

with PS due to bad PS generation using pswrite which generates full page images, no fonts.

```
40.12)
```

```
Ascent 63 / Aspensario / Ascent 63 / Aspensario / Ascent 64 / Aspensario / Aspensar
```

Reasoning pro author PDFs

I've tested this approach since 2009 when I told LaTeX people to upload their pdfLaTeX generated PDFs as "Other supporting files". What I found:

 The number of directly usable papers with uploaded PDFs is much higher than the same papers

with PS due to bad PS generation using pswrite which generates full page images, no fonts.

w/o PS you have to go to the source to generate PS.

Reasoning pro author PDFs

I've tested this approach since 2009 when I told LATEX people to upload their pdfLATEX generated PDFs as "Other supporting files". What I found:

- The number of directly usable papers with uploaded PDFs is much higher than the same papers
- with PS due to bad PS generation using pswrite which generates full page images, no fonts.
- w/o PS you have to go to the source to generate PS.
- Papers which had errors in the PDF, had them in the PS too, so touching the source was necessary (but that doesn't change the reasoning).

```
40.12)
```

```
| Ascent bas / Asc
```

```
0000000000 65535 f 0000000529 00000 n 000000418 00000 n 000000015 00000
0000022316 00000 n 0000022482 00000 n 000002249 00000 n 0000022143 00000
0000000608 00000 n 0000000937 00000 n 000001055 00000 n 0000022143 00000
0000010338 00000 n 00000021862 00000 n 0000022756 00000 n 0000022756 00000
trailer
```

The steps to use the PDFs

What do we have to do to make the PDFs to conform to JACoW and PDF/A rules?

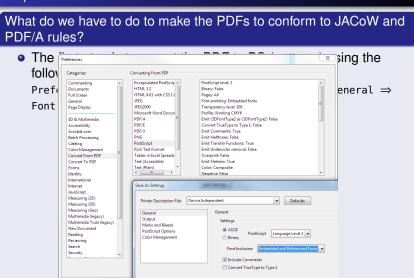
The steps to use the PDFs

What do we have to do to make the PDFs to conform to JACoW and PDF/A rules?

 The first step is to export the PDF to PS (save as) using the following setting (preconfigured)

```
Preferences \Rightarrow Convert From PDF \Rightarrow PostScript \Rightarrow General \Rightarrow
Font Inclusion ⇒ Embedded and Referenced Fonts
```

The steps to use the PDFs



What do we have to do to make the PDFs to conform to JACoW and PDF/A rules?

 The first step is to export the PDF to PS (save as) using the following setting (preconfigured)

```
Preferences \Rightarrow Convert From PDF \Rightarrow PostScript \Rightarrow General \Rightarrow
Font Inclusion ⇒ Embedded and Referenced Fonts
```

 This option controls which fonts should be included in the PostScript file. Embedded fonts are taken directly from the PDF, Referenced fonts are taken from the system in use.

What do we have to do to make the PDFs to conform to JACoW and PDF/A rules?

 The first step is to export the PDF to PS (save as) using the following setting (preconfigured)

```
Preferences ⇒ Convert From PDF ⇒ PostScript ⇒ General ⇒
Font Inclusion ⇒ Embedded and Referenced Fonts
```

- This option controls which fonts should be included in the PostScript file. Embedded fonts are taken directly from the PDF, Referenced fonts are taken from the system in use.
- The same can be done by printing the PDF to file using the JACoW PostScript printer (A4 or Letter has to be respected).

What do we have to do to make the PDFs to conform to JACoW and PDF/A rules?

- The first step is to export the PDF to PS (save as) using the following setting (preconfigured)
 - Preferences \Rightarrow Convert From PDF \Rightarrow PostScript \Rightarrow General \Rightarrow Font Inclusion \Rightarrow Embedded and Referenced Fonts
- This option controls which fonts should be included in the PostScript file. Embedded fonts are taken directly from the PDF, Referenced fonts are taken from the system in use.
- The same can be done by printing the PDF to file using the JACoW PostScript printer (A4 or Letter has to be respected).
- Now we distill the exported PS with the standard JACoW settings and get a file which contains ALL fonts.

1'00') 40.12))

> 0000 n 0000 n 0000 n

What do we have to do to make the PDFs to conform to JACoW and PDF/A rules?

- The first step is to export the PDF to PS (save as) using the following setting (preconfigured)
 - Preferences \Rightarrow Convert From PDF \Rightarrow PostScript \Rightarrow General \Rightarrow Font Inclusion \Rightarrow Embedded and Referenced Fonts
- This option controls which fonts should be included in the PostScript file. Embedded fonts are taken directly from the PDF, Referenced fonts are taken from the system in use.
- The same can be done by printing the PDF to file using the JACoW PostScript printer (A4 or Letter has to be respected).
- Now we distill the exported PS with the standard JACoW settings and get a file which contains ALL fonts.
- Nudging, cropping, shifting or any text corrections should be made before exporting to PS as this will ensure the correct embedding of fonts and placement of objects in the final PDF.

1'00')

10.12))

9000 n 9000 n 9000 n

Reasoning Required Steps Conclusion

Stey-by-Step to PDF/A conforming PDFs

```
We will get font inclusion
```

```
We will get font inclusion

    when the JACoW PS printer is used.
```

```
We will get font inclusion

    when the JACoW PS printer is used.

  o from LATEX (embed fonts is default)
```

```
We will get font inclusion

    when the JACoW PS printer is used.

    from LATEX (embed fonts is default)

 when we regenerate PS from user source files
```

```
We will get font inclusion
```

- when the JACoW PS printer is used.
- o from LATEX (embed fonts is default)
- when we regenerate PS from user source files
- from papers with PDF files when we export to PS and re-distill as described above.

```
endobj
9 o.bj [722 667 722 722 667 611 778 778 389 500 778 667 944 722 778 611 778
667 722 722 1000 722 722]
endobj
10 0.bj 
    //Length1 1626 /Length2 7293
//Length3 0 /Length 8919 >>
//Type /FontNescriptor
//FontName /Times-Boltd
//Flags 4 /FontNeskojth 650 /Descent -200 /ItalicAngle 0
//Scent 140 /Khejnt 461
//CharSet (/A/C/D/E/F/N/I/L/M/N/O/P/R/S/T/U/M/Y)
//FontFile 100 /Khejnt 461
//CharSet (/A/C/D/E/F/N/I/L/M/N/O/P/R/S/T/U/M/Y)
//FontFile 100 /Khejnt 461
//Descent -200 /Length1 1630 /Length2 9796 /Length3 0 /Length 11426 >>
//Type /FontDescriptor
//FontName /TimesRoman
//Flags 4 /FontRebox |-168 -281 1000 924]
//Ascent 678 /Caphelpit 651
//Sescent -216
//Se
```

```
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 00| <\ |
| 0 0
```

We will get font inclusion

- when the JACoW PS printer is used.
- from LATEX (embed fonts is default)
- when we regenerate PS from user source files
- from papers with PDF files when we export to PS and re-distill as described above.

So my conclusion is

We will get font inclusion

- when the JACoW PS printer is used.
- o from LATEX (embed fonts is default)
- o when we regenerate PS from user source files
- from papers with PDF files when we export to PS and re-distill as described above.

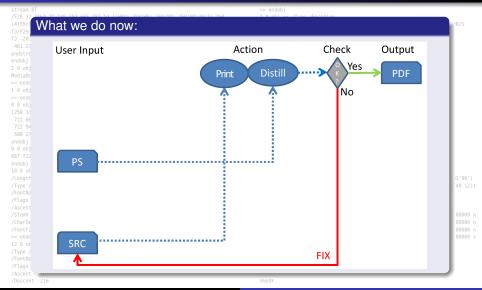
So my conclusion is

we can only win!

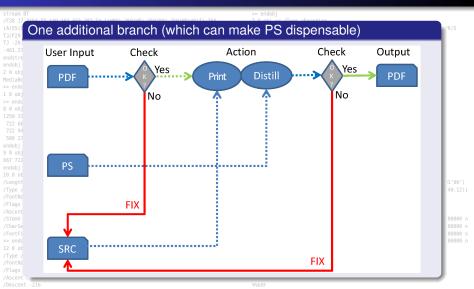
xref 0 16 0000000000 6553 0000022316 0000

<</pre>
<</pre>
/ID (575106298EB869B5538AEED2243D4BB4>)
<575106298EB869B5538AEED2243D4BB4>] >>
startxref 22978

The new workflow



The new workflow



The new workflow

