

Re formatting PDFs:
bringing digital files up to current
digi standards

Shepherd School of Music
Performance Programs

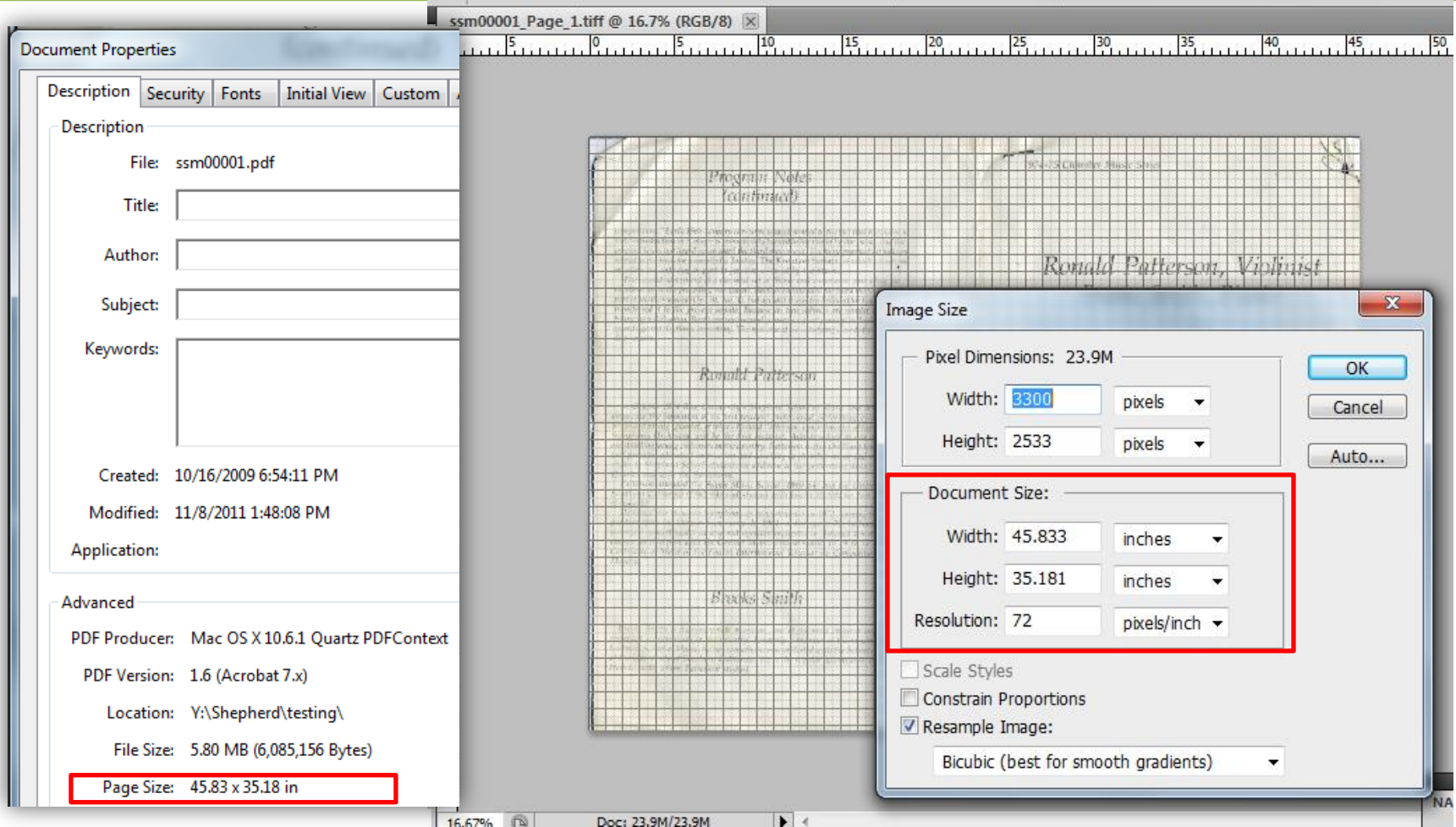
December, 2012

Issues

- PDFs are not OCR'd (non-full text searchable)
- PDFs page dimensions do not match actual physical page dimensions
- Page order in PDF is incorrect

Page dimensions off

Image Parameters do not match actual document



Adobe : file properties

PS : image output from PDF

Confirmation test

Image Size Calculator

[Pixels Explained](#)

Convert pixels to inches (Output to Monitors/Printers)

Formula: Pixels ÷ DPI = Inches

Image Width: Pixels

Image Height: Pixels

[Show me an image this size](#)

Select the resolution of the output device

Image Width: inches

Image Height: inches

If you know the pixel width and height of an image, this section will calculate the physical size (in inches) of the image when it is printed or displayed on various devices.

To use the calculator, enter the width and height, then select the output resolution which will be used.

* Most monitors display images at approximately 75DPI. To check your own monitor, use the [chart](#) below.

** Many printers use a process known as Image Refinement when processing images. This may impact the relation between input and output resolution. Generally, 300DPI is adequate for photo-like prints.

Re-formatting work

Re-sample page images to update PDFs files to current digi-standards:

- Change page dimensions to more closely match actual size
- Up resolution from 72ppi to 300ppi
- Resort page order
- Apply OCR

Conversion Steps

Activity	Tools/Methods
Export PDF pages as tiff images	batch process using Adobe Acrobat
Resample (ie convert all tiffs to 300ppi)	batch process using Adobe Photoshop
Crop tiffs to single pages where necessary	Bridge/Photoshop
Rename files so pages sort naturally in proper order	Bridge/Photoshop
Convert tiffs to PDF format, OCRd & Optimize	batch process using Adobe Acrobat
Combine PDFs to single PDF per item	Batch process: poppler (mac)
Embedded global metadata (description, source, copyright statement)	Bridge (XMP template)

Outcomes

- 347 enhanced PDF files
- Reduced total size from 1.45 GB to 187MB
- Simple OCRd aids with discovery

Next steps

- Re-Ingest : replace PDFs, update minor metadata edits (from last year) by Jan-2013
- Ingest of master files *.wav by 1st quarter 2013 (depending on new storage solution for IR)
- Reconfigure audio mapping (depending on new streaming server solution by Campus IT)