

# Church History Project

Timothy Nass

December 8, 2009

## Description

It was my goal to develop and use a process to digitally reproduce rare and/or out-of-print books which have significant value in the history and teaching of the Lutheran Church. I focused on the early years of *Theologische Quartalschrift*, and was able to complete the first 26 volumes, 1904-1929, a total of 7,636 pages. These are now recorded as PDF files on a CD.

## Need

Our seminary library has a valuable collection of Latin and German books, especially in our rare book room and in our collection of WELS publications. The problem is that there is limited access to these resources, especially for parish pastors. My goal is to make some of the most important Latin and German works available to pastors who might make use of them. Once produced, digital reproductions of books would require very little cost for distribution and would not require space on a pastor's bookshelves—many volumes will fit on one CD or DVD.

## Process

The process consists of two steps: 1) making a digital image of each page, and 2) processing the images into a usable format. While step one can be completed on many documents with a scanner, books present a special problem because it is difficult to make them lie completely flat and because someone needs to turn each page. Both of these problems can be solved by chopping the spines off the books and processing the loose pages, but that isn't an option for our books. Expensive machines exist for digitizing books, usually the combination of a high-quality digital camera with specialized software for processing the photos taken of each page. Some even include robotic arms with suction cups that automatically flip the pages. (Such machines are used by Libronix and Google Books.) I can produce similar results with my digital camera at home and batch photo editing software.

Most digitizing projects include OCR software, which converts images of words to searchable text. OCR software exists that can read Fraktur, but it is expensive—\$.10 - \$.20 per page. The 17<sup>th</sup>-century typesets of the rare books mixed with Fraktur, Greek, and Hebrew are even more of an OCR nightmare. I did not run OCR on the Fraktur text of these images, although the images are hopefully useable for that if someone in the future wants to make the text searchable. Rather than searchable Libronix-type files, the *Quartalschrift* volumes are saved as PDF files made of images rather than searchable text.

## My Process

My setup could not be called professional, and I have no doubt that others could improve on what I've done. In a few years, new technology will make much of this obsolete. Here's what I used:

### Camera

I used my personal digital camera, an 8 megapixel Canon PowerShot SD1100. It was made to be small and to fit in a pocket and to take snapshots, not really to take thousands of pictures of book pages. But it worked. I bought a machine screw that matched a tripod screw's threading. I then drilled a hole in a 1x4 piece of wood so that I could screw the camera to the edge of the board using the machine screw. I then clamped the board to a hutch above my desk so that the camera pointed down toward the desk surface.

### Camera settings:

Exposure: ISO 80

White balance: Evaluated from the yellowed book page

Colors: Black and White

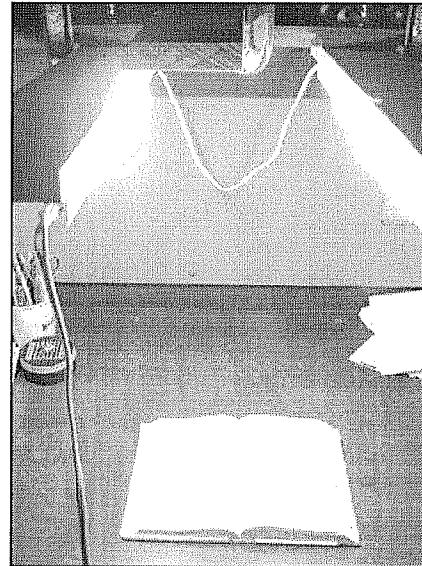
Size: 3264 x 2448 pixels

Ideal features my camera didn't have include a cord to plug the camera directly into an outlet. The lithium battery of my camera was good, lasting two hours before needing recharging, but it would have

been nice to have had a constant power supply. Even more helpful would have been a remote control, or better still, a foot pedal to take the pictures so that I could have had both hands free to position the book.

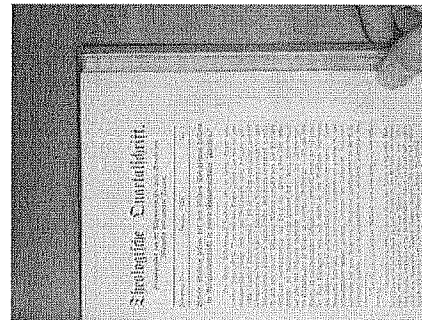
### Consistent Lighting

The key to the whole project is consistency over thousands of pages. The hardest part was perfectly even lighting. At first I tried using regular light bulbs and then diffusing the light by putting sheets of white tag board around the book and shining the lights indirectly off the tag board. It was impossible to make adjustments other than trial and error, however, and there were always hotspots and dark patches because the lighting wasn't perfect. (A computer scanner bed has perfectly even lighting.) Finally I purchased two short fluorescent lights, attached them to 1x4 boards, and clamped them to the desk hutch on either side of the camera. Even then the light wasn't perfect—it was brighter in the middle of the page because the lights were slightly closer there and darker toward the page edges. I put layers of tissue paper over the middle of the fluorescent bulbs until it was balanced. Even so, there is a consistent problem with a dark spot appearing on the bottom left corner of even-numbered pages. Part of the problem with lighting is getting the book to lie perfectly flat. If parts of the page are curved (e.g. going down into the spine), the result is a dark spot in the picture.



### Processing

Processing all the photos individually would take an incredible amount of time, so some kind of batch editing software is essential. The seminary library has Adobe Photoshop, but I found what I needed in a less expensive program called BatchPhoto. I used the program in a three steps in processing the photos.



Picture 1 to the right is how an original photo looked. I took all the odd pages in a row (1, 3, 5, etc.) and then all the even pages so that I wouldn't have to move the whole book for each page. When I took the pictures off the camera, I put all the odd numbers pages in one folder and all the evens in another. In the first step, I cropped all the pictures (taking them in exactly the same position each time was imperative), and rotated them correctly (picture 2). At the same time I renumbered them 001, 002, 003, etc. BatchPhoto could do all that for hundreds of pictures in a matter of minutes. I then did the same thing to the even pages, but I numbered them 001a, 002a, 003a, etc., so that when the pictures were combined, the odds and evens would be in order again (001 = p.1, 001a = p.2, 002 = p.3, 002a = p.4, etc.).

1

Step two was to rename all the photos according to their actual page numbers (WLQ\_026\_001.jpg; WLQ\_026\_002.jpg; etc.).



2

3

In step three I used a contrast filter set at 95 and a brightness filter with lightness set at 100 and gamma set at 100 to improve the contrast. At the same time I converted the JPEGs to TIFF files (picture 3 to the right). The advantage of a TIFF is that it uses only black and white pixels (monochrome) rather than shades of gray like a JPEG, so that the file size is much smaller—less than 20 Mb for 300 pictures. The trouble is that poor lighting can easily become black spots that are illegible. When I became comfortable with the whole process, it took a little less than an hour to photograph a 300-page volume of *Quartalschrift*, plus about half an hour to process it. Any mistakes however, (e.g. one missing page) extended that time significantly.

### Conversion to PDF

Using Adobe Acrobat 7.0, I could convert the TIFF files into a 300-page PDF in a matter of minutes. It was also easy to alter the page numbering of the PDF so that the table of contents is numbered with Roman numerals. That way, when you type in page 200, it finds actual page 200 (which is really page 203 or 204 in the PDF document because of the table of contents). When you view the PDFs, I recommend going to the "View" menu, selecting "Page Display," and then selecting "Two-Up." The pages then appear somewhat like an issue of *Theologische Quartalschrift*. In the same way, print two pages to one 8.5 x 11 sheet for what looks like a photocopy.

### Personal Evaluation

I was pleased that the results were fairly good, even without professional equipment. The text is very legible, and a whole shelf of books fits on a single CD. If I would do it all over again, however, I wouldn't use this process on the *Quartalschrift* issues. I would obtain a set to chop the spines off, run them through an automatic scanner, and then rebind the issues into books. They aren't that rare to merit the extra care in keeping them perfectly intact. That probably would have been much quicker, and the quality would have been better because the lighting would have been consistent. The process I used, or something similar to it, would, however, be the best way to inexpensively reproduce rare books that cannot be cut apart.