

Daily steps for post-exam form scanning

Konrad Herbst

[2021-03-08 Mon], version 0.1

1 Start computer and scanner

- start scanner (right round button)
- start computer (upper right key on keyboard)
- log-in for
 - user: `scanner-station`
 - pw: `virusfinder`

2 Set up and start Scanning pipeline

- use short-cut on desktop to enter directory `2021-PostExamScanning`
- make new working directory for respective scanning day and enter it
- copy-paste file `00-init.sh` into newly created directory and execute the file by double-clicking
 - several new files and directories will appear
- start scanning pipeline by executing file `01-start_scan_processor.sh`
 - a black terminal window will appear, which you need to leave running in the back
 - if it stops working/crashed, rename the directory `Scans_processed` (e.g. to `Scans_processed_beforeCrash`) and restart pipeline by double-clicking again on `01-start_scan_processor.sh`
- start the program `xsane` (brown 'S' with blue sphere on the left of the desktop) which will be used to control the scanner
- in the window `xsane multipage project` create a new project by
 - clicking the page icon (upper left, besides path entry text field) navigating to the directory `Scanning` within your newly created working directory
 - before you click OK create a new folder called `Batch1` (`Batch2`, `Batch3` . . . , see below) with the upper right icon `Create Folder` and hitting the `Enter` key
 - confirm that the path in this window now points to your `Batch1` folder you just created
 - click `Create project`

3 Scanning

- feed a stack of straitened questionnaires into the scanner
 - max. 30 pages (otherwise **xsane** might crash because scanning is too fast for the saving speed)
 - make sure the feeder guides (black plastic on either side of the stack) are pushed together, otherwise the paper might be scanned with too much rotation
 - note that the side with the filled questionnaire must point towards the back (only that is scanned)
- the scanning pipeline will process the scanned pages and automatically moves the raw scans out of the **Scanning** directories into the result directory **Scans_processed**
- repeat feeding in a stack of paper one after the other
- from time to time (after approx. 200-300 scanned pages) one should close **xsane** and restart it because if it runs for too long it also tends to crash and the last couple of scans will be lost.
 - If you restarted purposely, increment to a new subdirectory (e.g. **Batch2**) and continue scanning as before.
 - If you needed to restart **xsane** after a crash, also increment to a new subdirectory and repeat scanning of the last stack of paper you processed to ensure no page is missing. This will mean that a couple of pages were scanned twice and has to be taken into consideration later during analysis
- once you are done with scanning and the scanning pipeline indicates it is done with processing (0 files found to process) close that terminal window

4 Aggregate data

- the processed data is in the subdirectories within **Scans_processed** within your working directory
 - successfully processed scans are in the **processed** directory which also contains a directory **crops** with cropped scans of the filled textbox answers
 - successfully processed scans are in the **processed_failed** directory; if everything worked this will remain empty
 - the directory **processing** should be empty if the pipeline is done processing the **Scanning** directory
- execute answer aggregation by double clicking on the file **02-collect_scan_results.sh**
 - a terminal window will temporarily open indicating collection of the answers
 - within each **Batch...** subdirectory within **Scans_processed** new files appear of which the file **aggregated_results.tsv** is the relevant file containing the collection of answers

In principle, all the **aggregated_results.tsv** and the **crops** directory within each **processed** directories contain all the automatically detected answers of the scans. One could collect those in a separate folder depending on the subsequent analysis.

5 Computer and scanner shut-down

- once analysis is done one can shut-down the computer (clicking the arrow in the upper right) and turn off the scanner