

Scripting Night 2020: Scanning to PDF fixup

December 3, 2020

John Nash
for
Virtual Meeting of Linux-Ottawa

Issue

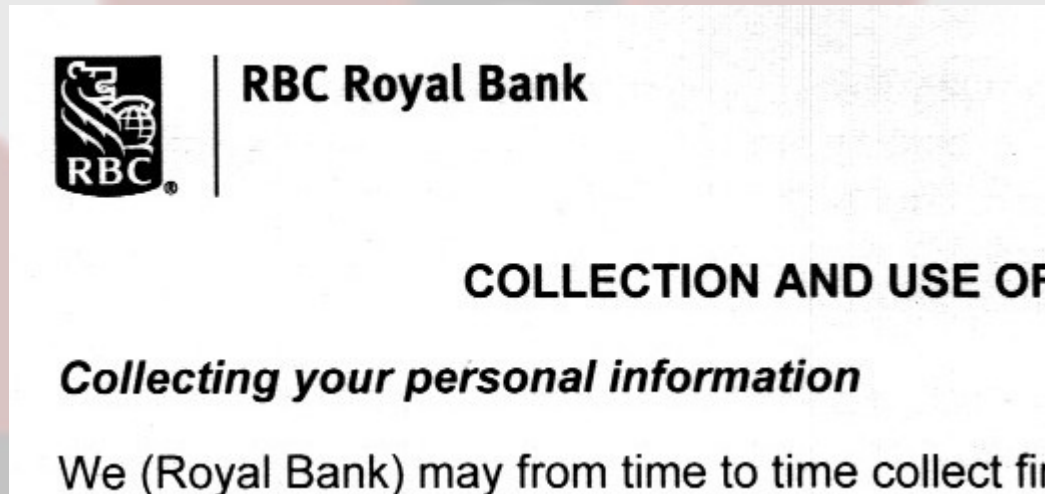
- Canon Lide60 scanner purchased July 2006
- Still works nicely, but ...
- SimpleScan (now Document Scanner) used to give fairly clear pdf output, but now gives muddy grey background
- Recipients of scans have complained of overuse of ink when printing!

Example

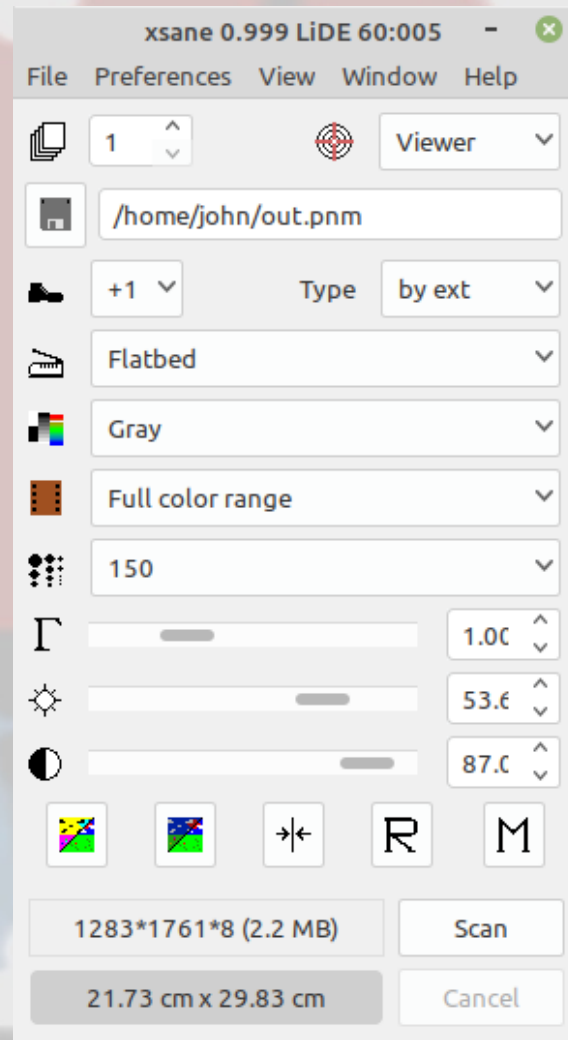
			STAMP C
UNIQUE NO. 0609837	SPEC.	PATIENT'S OFFICE ACCOUNT NO. 009263	I hereby certify that the name is correct
Dr. David Schooley 99 Kakigi Rd. Suite 202 Kanata, Ontario K2L3C8 Fax (USU) 000-0000 PHONE NO. (613) 592-0615			
Us, procedures,		I understand that the fees listed in this claim may not be covered by my plan/benefit. I understand that I am financially responsible. I acknowledge that the total fee of \$_____.	

Can scan OK with effort

Xsane using special settings



Brightness/Contrast settings



Ease of use

- Xsane -> PDF inconvenient
 - Saves as .pnm, need to convert and consolidate
 - Scan each page separately
- SimpleScan makes this “easy”, but ...
 - Also gscan2pdf (tried lots of options, but ...)
- A script can be prepared fairly easily

canonscan20J19

```
#!/bin/bash
# Copyright (C) John C Nash 2009
## To scan multiple pages and increment the names of the image files
## May not perfectly correct images
echo "canonscan20J19 J C Nash 2020 -- scan multiple pages to incremented
files"
echo "You may need to fix the scanner string 'scanimage -L' to detect"
echo "Files currently present"
ls
echo "Enter last index of scan (0 to start anew)"
read ii
echo "Starting at $ii"
## My scanners ## genesys:libusb:003:005 -- Canon LiDE 60 flatbed
scanner
## Note that the quotes from scanimage -L are WRONG!!!
DEV="genesys:libusb:003:005"
END_CONDITION='X'
RES=150
echo "Put document page on scanner; Enter to continue, X to stop"
read var1
```

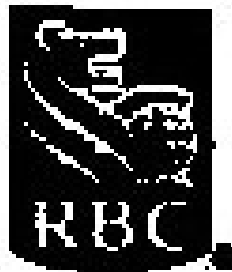
```
until [ "$var1" = "$END_CONDITION" ] # Tests at top of loop.
do
  let ii=ii+1
  jj=$(printf %03d $ii)
  echo "scanning page $jj"
  scanimage --format jpeg --resolution $RES -d $DEV >image.jpg
  echo "cleaning image"
  echo "converting to jpg"
  convert -quality 90 -threshold 55% image.jpg imagex.jpg
  cp imagex.jpg "scanimg$jj.jpg"
  echo "Put document page on scanner; Enter to continue, X to stop"
  read var1
done
echo "Rename images - no entry skips renaming"
read -p newname: newname
if [ "$newname" != "" ]; then
  echo "RENAMING scanimg??.jpg to $newname??.jpg"
  rename s/scanimg/$newname/ *.jpg
fi
echo "Files:"
ls
echo "DONE!"
rm imag*.jpg
convert $newname*.jpg $newname.pdf

rm "$newname??.jpg"
```


Notes

- Need device handle in ***scanimage*** to get correct scanner, and save time with searching
 - Versions of script for different computers
- Could possibly find a way to automate that
 - Suggestions??
- Could auto-launch pdf-arranger (ex. pdf-shuffler)

Output



RBC Royal Bank

COLLECTION

Collecting your personal information

We (Royal Bank) may from time to