

PDF (Auto) Splitter_Renamer

Version 0.2

for

Individual File Naming of Scanned BIR Forms 2316

developed by:

Regional Finance Service Office 8

Philippine National Police Finance Service

Prepared by:

NUP Mark W Marcos
Administrative Aide VI/Data Controller I

Verified by:

PLTCOL MERINE C DE GUZMAN
Asst Chief / Chief, RCS
Date: August 12, 2021

Noted by:

PCOL EMMANUEL R BRAVO
Chief, RF8O8
Date: August 12, 2021



Table of Contents

- I. Introduction/Executive Summary
- II. Review of Other PDF Splitter Tools and Methods
- III. PDF File Splitting and Renaming Automated
 - A. PDF (Auto) Splitter_Renamer version 0.1 or PASR 0.1
 - B. PDF (Auto) Splitter_Renamer version 0.2 or PASR 0.2
- IV. How to Use PASR 0.2
- V. Development Direction and Recommendation



I. Introduction/Executive Summary

Since 2019, the Withholding Tax Section (WTS) of the Revenue and Collection Division of the PNP Finance Service has started scanning and saving into computer hard drives the employer's copy of the Certificate of Compensation Payment/Tax Withheld (BIR Form No. 2316) for CY 2018 filed and submitted by all PNP personnel.

For easy reference, access and retrieval, and in case of requests for copy issuance, each of these scanned forms has to be properly archived and file named. Hence, after the forms are scanned and saved in bulk by unit or by RHQ/Regional Unit/RMFB/Provincial/City Police Office, WTS passes on these forms saved in Portable Document Format (PDF) and delegates to the Regional Finance Service Offices/Finance Service Offices (RFSOs/FSOs) the task of segregating or splitting the scanned forms into single pages with prescribed filenames taken from the alphalist for the given period, being provided by WTS, using the data for the listed personnel's full name and PNPFS/payslip account number so that each separate or single tax form must, for example, be saved with a filename of *'Dela Cruz, Juan Esperas _ 20210000001.pdf'*.

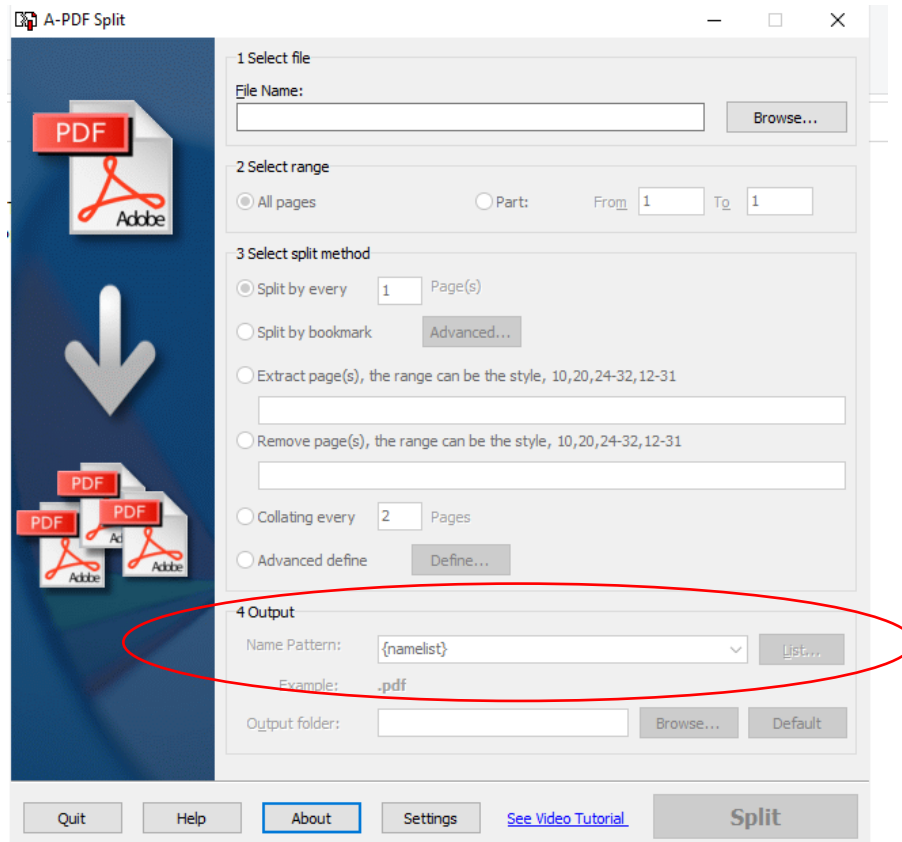
In a memo dated July 28, 2020 signed by then PBGEN FERDINAND B DAWAY, Director, PNPFS, with subject: Submission of File Named BIR Form No 2316 for CY 2019, procedural guidelines were provided for the file naming process. However, following the guidelines entailed utilizing more personnel and spending more work days than necessary since the instruction was for a personnel or operator to extract manually each page and refer one by one from the alphalist for the filename to be given to each page for extraction before saving each as a separate PDF file.

As a workaround, since the alphalist's sequence and the page order of the scanned forms by unit should match, why not look for some other PDF editor that can split up all the pages of a PDF file and at the same time accept as input the list of filenames from which this editor should per line take from this list the filename to be given to each page before saving each as a separate PDF file?

II. Review of Other PDF Splitter Tools and Methods

By trying initially around the month of April 2019 for the processing of the CY 2018 forms, and again about a month before the deadline for the CY 2019 forms (NLT September 30, 2020) set by the memo addressed to the Chief of RFSOs/FSOs, and then after coming across a Lifewire.com review article by Tim Fisher entitled "11 Best PDF Splitter Tools & Methods", the search for PDF editors with the needed function proved unfruitful.

Though the 10th and 11th software tool reviewed by Fisher already had in them some splitting functions, the specific file renaming feature below (under '4 Output') of the 11th tool, A-PDF Split, was nonexistent then and included only lately (Tim Fisher's April 23, 2021 update of same article). Even so, A-PDF Split is not open-source, not free and can only be fully utilized for a fee.



What We Like

- ✓ Simple interface.
- ✓ Multiple ways to split the PDF.
- ✓ Runs on macOS and Windows.

What We Don't Like

- ✗ Isn't free.
- ✗ Doesn't include a built-in PDF viewer to see which pages you're working with.

(Image and commentary taken from "11 Best PDF Splitter Tools & Methods", Tim Fisher, Lifewire.com review article, Updated on April 23, 2021.)



III. PDF File Splitting and Renaming Automated

a. PDF (Auto) Splitter_Renamer version 0.1 or PASR 0.1

After splitting the source PDF file into single pages using Nitro PDF with initially a numbered output file name for the split pages (a limitation in the file renaming function of Nitro PDF), Version 0.1 consisted only of the following Python3 scripts (used for the file renaming of Biliran PPOs' forms):

```
>>> import os

>>> i = 1

>>> directory = os.listdir('E:/2316/RFSO_8 BILIRAN PPO 238')

>>> new_filename = open('E:/2316/BILIRAN.txt')

>>> while i <= len(directory):
    src = 'E:/2316/RFSO_8 BILIRAN PPO 238/0_' + str(i) + '.pdf'
    dst = 'E:/2316/RFSO_8 BILIRAN PPO 238/' +
        new_filename.readline().rstrip('\n') + '.pdf'
    os.rename(src, dst)
    i += 1
```

Ran as Python3 shell commands then edited to point to the next directory of split pages to be renamed and to the preformatted '.txt' file containing the input list of filenames, automating the file renaming routine was first made possible during Regional Finance Service Office 8's processing of the CY 2018 forms.

b. PDF (Auto) Splitter_Renamer version 0.2 or PASR 0.2:

Since the earlier version of PASR, it was observed that relaying instructions on how to execute the scripts properly can be incomprehensible to those who have no coding background. For the less technical users or operators, PASR 0.2 came out of enhancing and wrapping the codes behind a user-friendly graphical user interface or GUI of an executable stand-alone Windows desktop application, with the inclusion of some readme instructions for those new to the task.

After test-running PASR 0.2 and utilizing it in the file naming of the CY 2019 forms, its source code was first published last September 6, 2020 and accessible at <https://github.com/Mak0y-Plaks/PDF-Auto-Splitter-Renamer>.

PASR 0.2 was coded using Python 3.7 with PyPDF2 and package-freezed using PyInstaller as a stand-alone executable application for the Windows 8/10 environment.

It was specifically developed for the later part of Philippine National Police's Finance Service's approach to offline mass archiving of scanned individual BIR Forms No. 2316 tax forms (scanned at PNPFS NHQ) which are filed



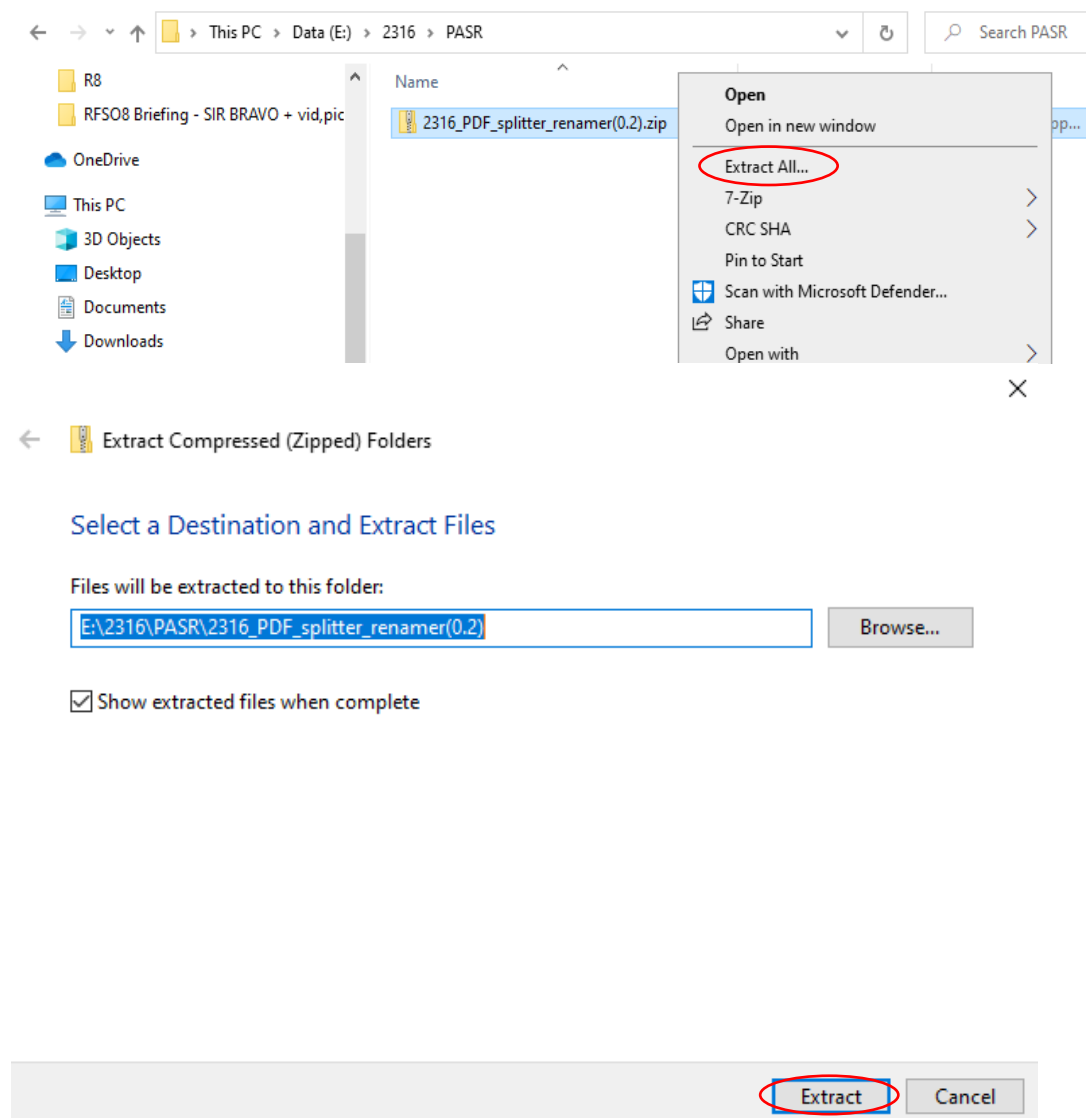
annually by all personnel then collected separately, sorted, bundled and submitted mainly by Regional Headquarter/Regional Unit/Mobile Force Battalion/Provincial/City Police Office Finance Sergeants to Regional Finance Service Offices/Finance Service Offices.

PASR 0.2 is a straight-forward application, retro-fitted to RFSOs/FSOs' process of automating page-by-page splitting of scanned 2316 forms in PDF format and renaming of each page using a preformatted/pre-arranged list of filenames.

Coding was focused only on splitting or copying every page from a (read-access only) PDF containing the scanned tax forms and subsequently saving and renaming each copied page with the use of a list of strings appended from either manual CLI input() or multiple text input() from clipboard.

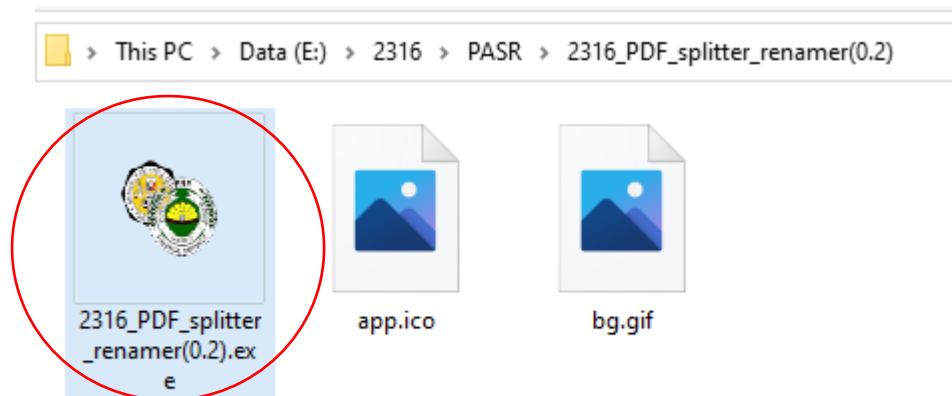
IV. How to Use PASR 0.2

To use the application, (1) first get and extract the files contained in the ('2316_PDF_splitter_renamer(0.2).zip') found at the GitHub site.

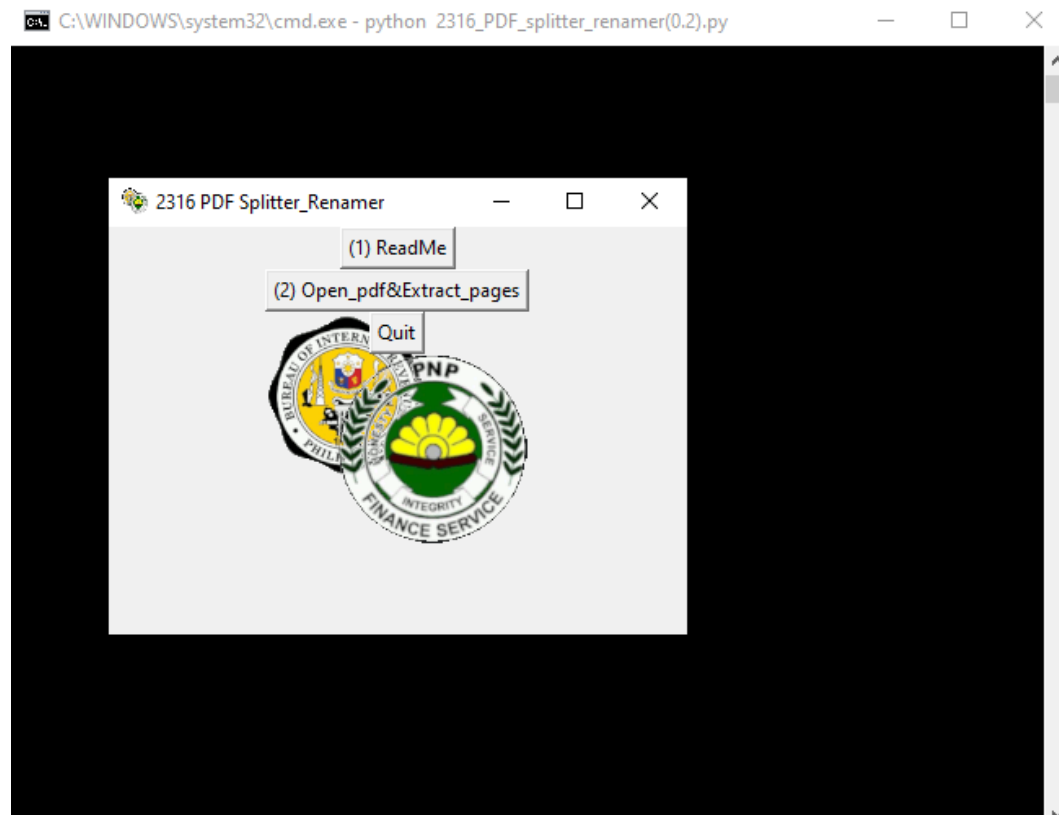




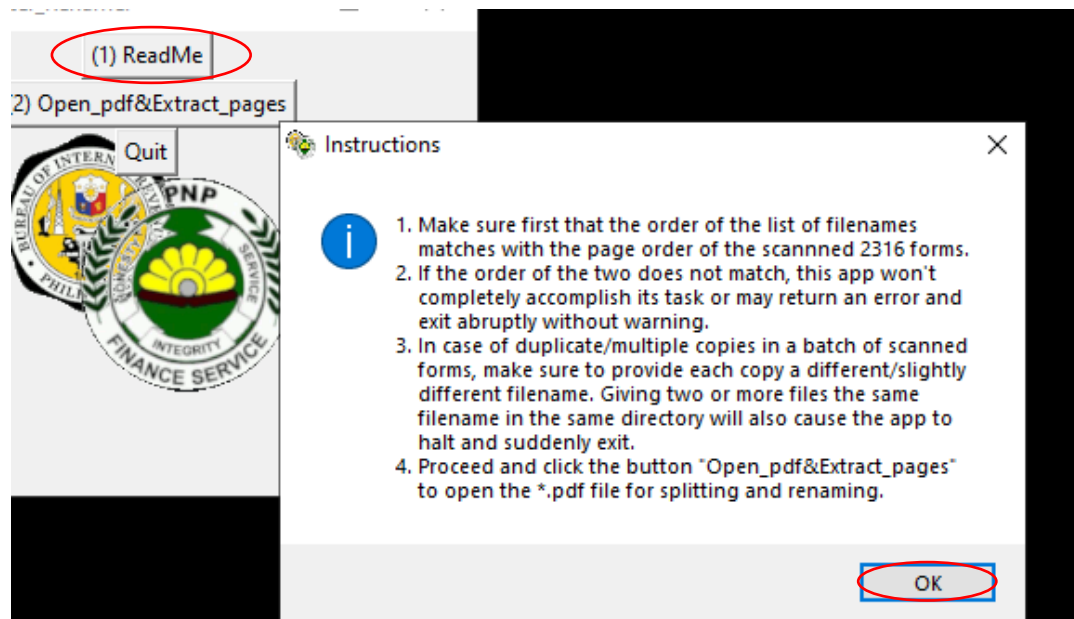
(2) Run/double-click '2316_PDF_splitter_renamer(0.2).exe'.



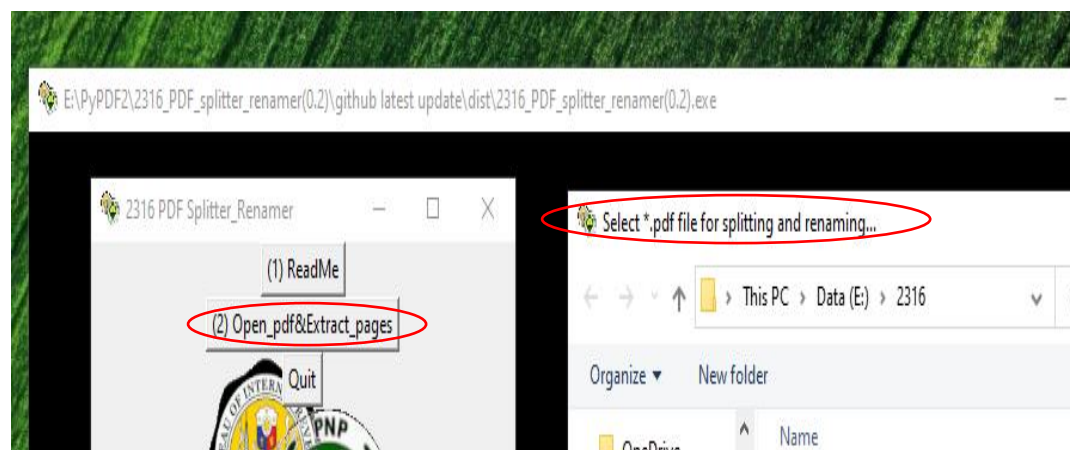
(3) Two (2) windows will appear. The main window with the title bar labeled '2316 PDF Splitter_Renamer' and behind it, the Windows CMD terminal, which will serve as the input terminal for the list of filenames later.



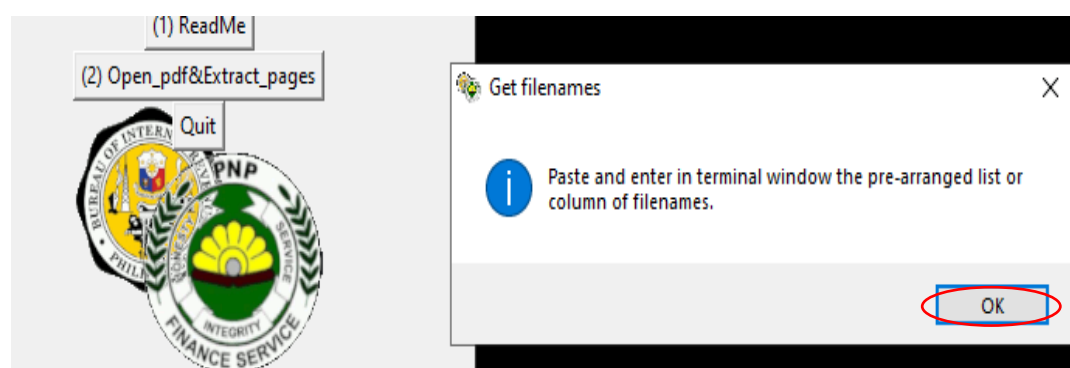
(4) Clicking the 'ReadMe' button will show the following instructions:



(5) Click the 'Open_pdf&Extract_pages' button to get the PDF file for splitting.

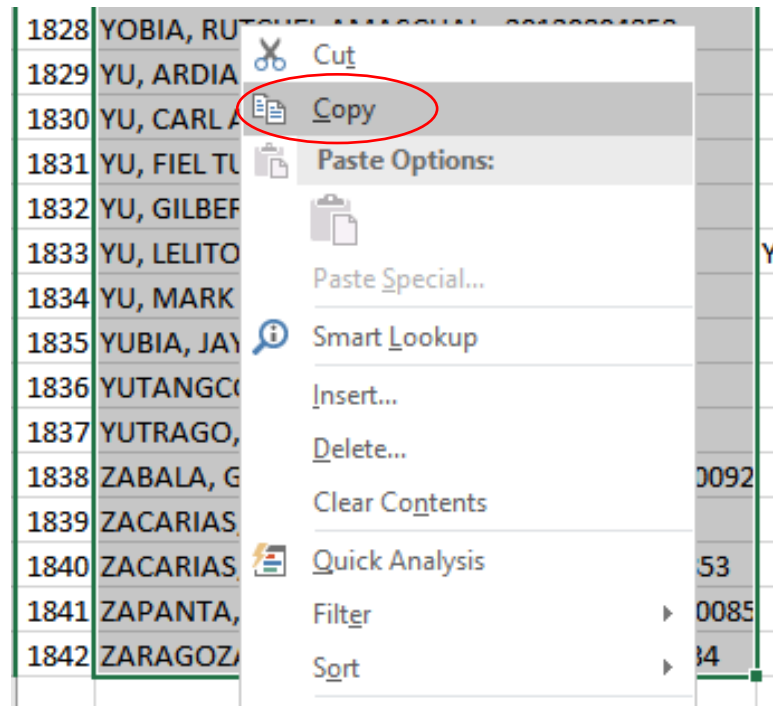


(6) After choosing the PDF file for splitting, click 'OK' on the prompt that appears afterwards.

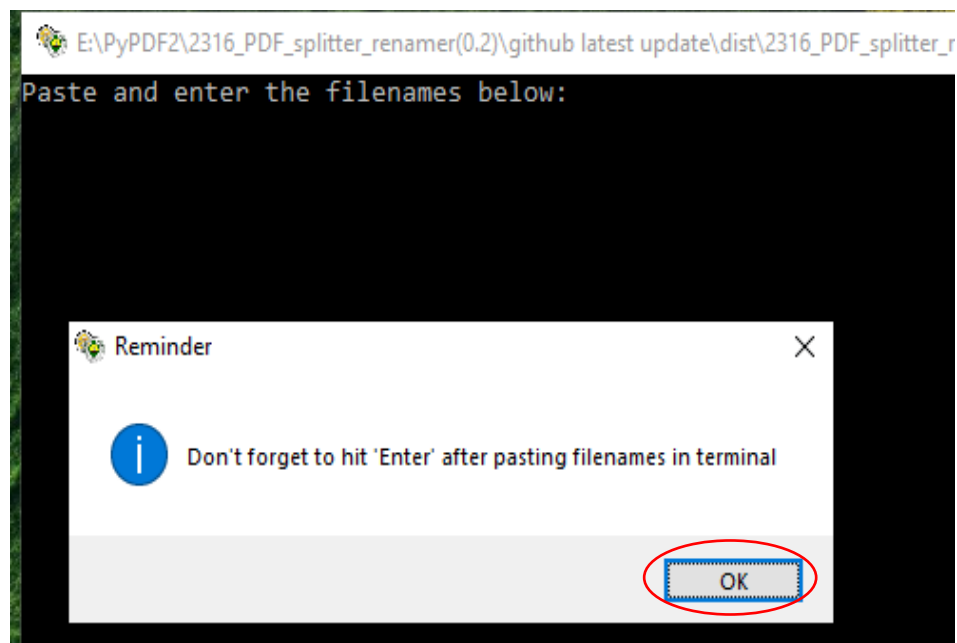




(7) Copy the preformatted and complete list of filenames for the unit from the alphalist.

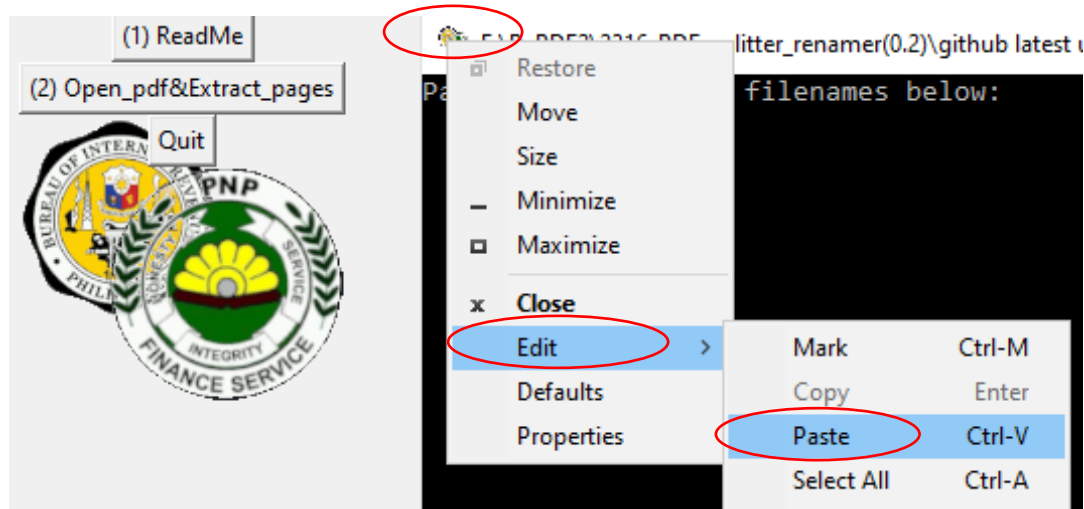


(8) Click 'Ok' first on the window prompt showing the following reminder:

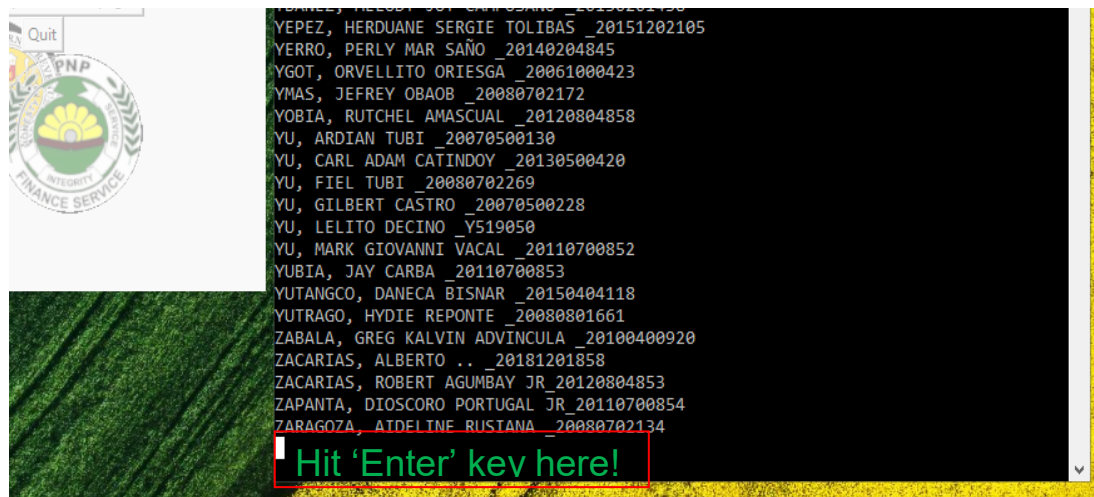




(9) Click on the on the logo on the top-left header of the Windows CMD terminal and go to Edit > Paste to load into PASR 0.2 the input list for the source of filenames to be given to all the single split pages from the source PDF file.



(10) Hit the 'Enter' key at the bottom of the loaded list.



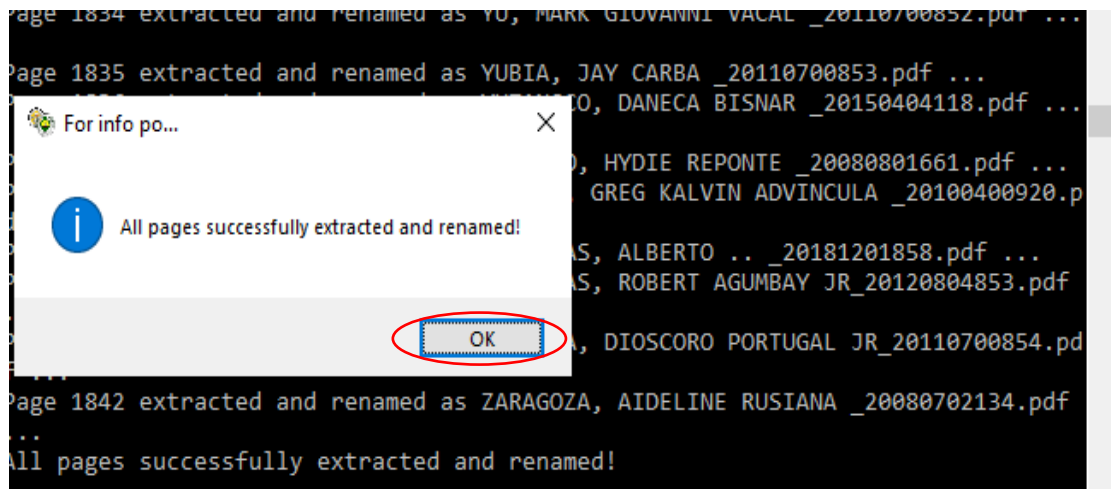
(11) PASR 0.2 begins the automated page-by-page splitting and renaming.

```
Splitting each page of 'E:/2316/R8/RFSO 8 LEYTE NORTE 1842.pdf' and extracting
to
same folder/directory 'E:/2316/R8/' ...

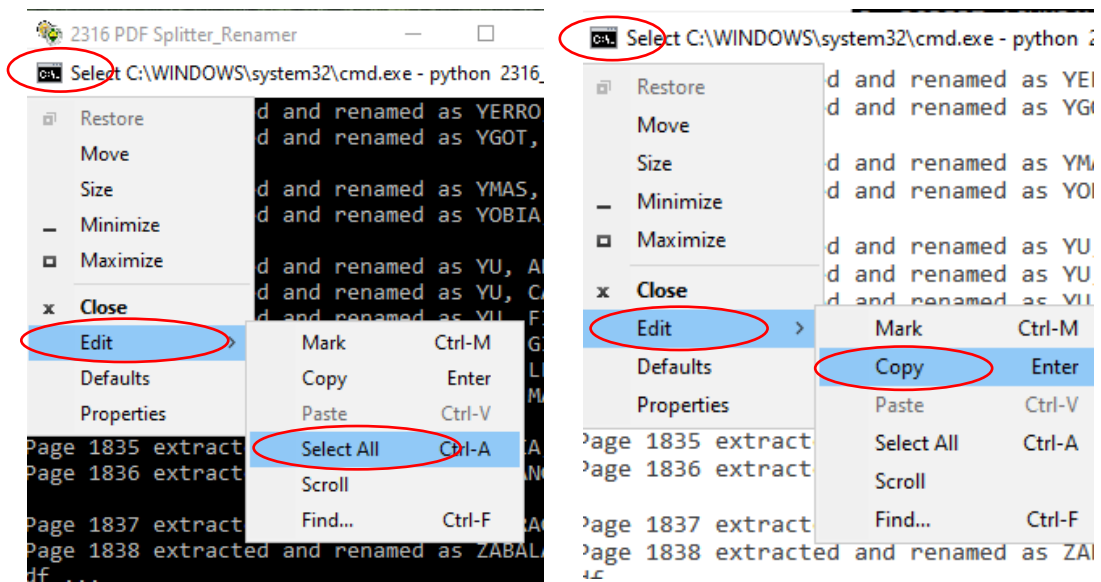
Page 1 extracted and renamed as ABAD, KRISTHINE PEREZ _20111200198.pdf ...
Page 2 extracted and renamed as ABAH, FHAUJIEA NAZ _20140200122.pdf ...
Page 3 extracted and renamed as ABAH, KHALID MILLARES _20110700459.pdf ...
Page 4 extracted and renamed as ABAL, GABRIEL AUSTERO _20151201528.pdf ...
Page 5 extracted and renamed as ABALLA, BRYAN SALGADO _20061000227.pdf ...
Page 6 extracted and renamed as ABALLE, JOMAR MOMO _20160800010.pdf ...
Page 7 extracted and renamed as ABALLE, ROMNICK MAHAIT _20140200376.pdf ...
```



(12) A prompt appears after the last page is renamed. Click 'OK'.



(13) Note that the logs are not saved by PASR 0.2. For reference, you may copy and save to a text file the logs before closing.





V. Development Direction and Recommendation

If deemed viable, other essential functions, add-ons or options that can be added to future versions of PASR may be the following:

- (1) User log-in interface for security;
- (2) PDF page viewing and within-app reconciliation of discrepancies in the sequence of the source PDF's page order and the input list of filenames;
- (3) Alphalist database sourcing, generation and record logging of necessary input/output data; and
- (4) Optical Character Recognition or OCR with the increasing legibility of the scanned forms.



DevInfo



2316_PDF_splitter_renamer is being developed by:

RFSO 8, PNP - Finance Service
Camp Ruperto K. Kangleon,
Barangay Campetic, Palo
Leyte, Philippines

For queries or feedback, contact details:
email: rfsopro8@gmail.com
Office phone no.: (6353) 832-4373

OK