KHANH LE

ISYS 520: FINAL PROJECT

PDF-TO-EXCEL PROGRAM

I. *Executive Summary*

With the development of technology, the online money-transfer systems are widely available for investment banks' clients to instruct transactions. However, many tradition clients still prefer the old way of transferring money via bank tellers. Normally, when clients want to transfer money among accounts, they will email their banks the letter of authorizations (attached as a .pdf file) instructing how much money will be transferred and where the money will be transferred to. After receiving that email, a bank employee will manually enter the necessary information to the bank's system and process that transaction. This manual process is time consuming and costly for the bank. Therefore, this "pdf-to-excel" program will automatically read through the digital .pdf file and extract all needed information and fill it into a excel form. After processing the transaction, users have the option to save the transaction information for future reference.
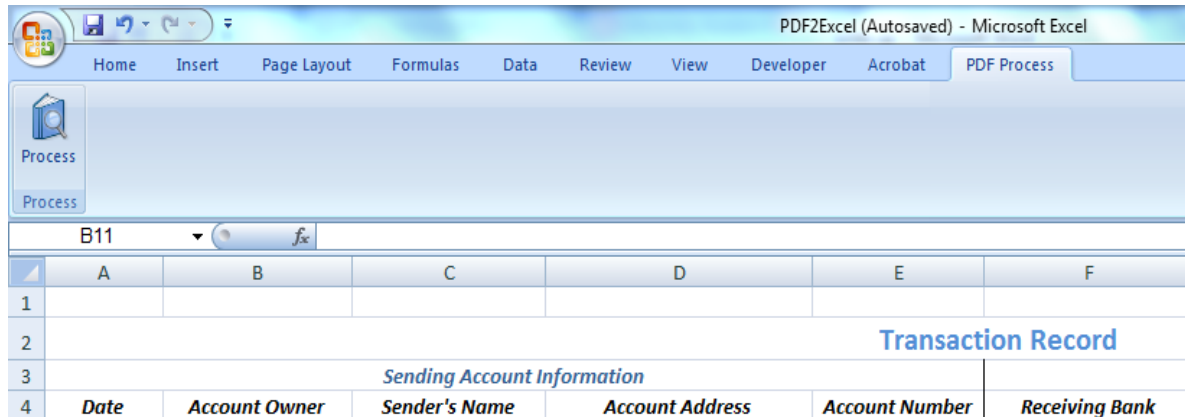
II. *Program Limitation*

1. This program will copy all the words available in a specific .pdf letter of authorization form and store them in an array. The codes will loop through that array and select the only necessary information and paste this information into Excel. For this reason, this particular program can only read and process information from the sample form attached along with this project. If users need to process the information of another form, they need to modify the codes.

2. To successfully read and extract the data from the .pdf file, users need to install Adobe Acrobat to their computer. This software can be easily purchased from Adobe website or at local software stores.

3. The developer of this program does not have access to the bank system, so when users click the "Process Transaction" button, the transaction is not actually processed. Users need to link their bank systems to this program in order to process transactions.

III. Program User Instruction

This program includes one Excel file.
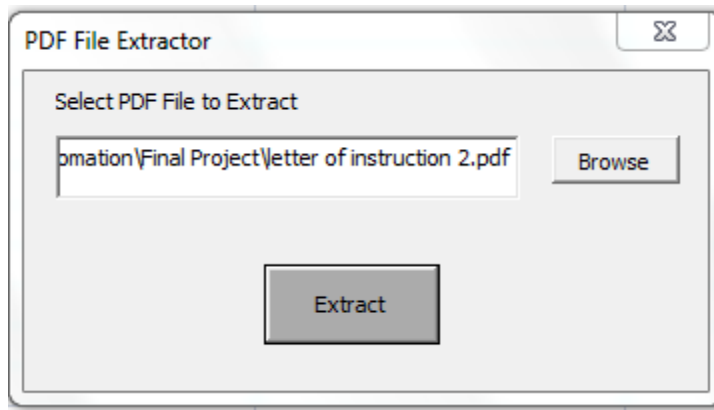
1. *When open this Excel program, users will see this screen:*

Under the "PDF Process" Tab, users can click on the "Process" button to process the file.

2. *After clicking on the Process button, this window will appear:*

Users can use the "browse" button to find the .pdf file that they want to extract.

After successfully locating the .pdf file, users can hit the "Extract" button.

3. *The program will process that .pdf file and export the information into an Excel form like this:*

Transfer Information   ☒

Sending Account   **Receiving Account**

Account Number: 012345678   Bank Name: Bank of United People

Account Owner: John Smith Junior   Bank Address: Seattle Washington

Name of Sender: Sarah Smith Senior   Routing Number: 333444r343

Address: 1234N 11W Salt Lake City   Account Owner: Mary Woods Junior

Amount Wired: $500,000,000   Account Number: 01122334455334

Date to Wire   12 / 15 / 2013   [Process Transaction]
(mm/dd/yyyy)

Users still have the ability to change the incorrect information in the text box before they process the transaction.

To process the transaction, users click on the "Process Transaction" button (Users need to link their bank system with this program in order to actually process the transaction).

4. *After processing the transaction, the "Save data" window will appear and ask if the users want to save the previously-processed transaction data for future references:*

Save Data   ☒

Transaction has been processed. Do you want
to save a record of this transaction?

[Yes]   [No]

If users click "No", this window will disappear.

If users click "Yes", the data will be saved in the current Excel file, under the "Records" tab:

Transaction Record

		Sending Account Information				
Date	**Account Owner**	**Sender's Name**	**Account Address**	**Account Number**	**Receiving Bank**	**Bank**
15/12/2013	John Smith Junior	Sarah Smith Senior	1234N 11W Salt Lake City	012345678	Bank of United People	Seatt

▶ ▶ | Records

IV. Concept Difficulties Encountered

1. When starting this project, the developer of this software does not notice that Adobe Acrobat needs to be installed in order for VBA to read through the .pdf file.

2. After being able to get the text from the .pdf file, the developer recognizes that the text string cannot separate the last word of one line and the first word of the next line by a space. Instead, this text string puts a carriage return symbol in between the two words. The carriage return symbol indicates the end of a line of text, and the first character on the right of the carriage return symbol starts at the next line. This carriage return symbol is equivalent with character # 13 in the ASC II system. Therefore, the developer can replace this symbol by a space (using replace function: String=replace(string, chr(13)," ")), thereby, being able to split each word in the text string into an array.

V. Assistance

This program was written with the instruction of Professor Gove Allen.

The developer of this program also receive helps from the following websites:

http://social.msdn.microsoft.com/Forums/da/isvvba/thread/b2264226-da8e-4453-8f02-3fa3773d090b

http://www.khk.net/wordpress/2009/03/04/adobe-acrobat-and-vba-an-introduction/