



PDF Distributor

Revised 12-1-2009

Call 813-298-2474
www.edocfile.com
Sales@edocfile.com

PDF Distributor

Purpose

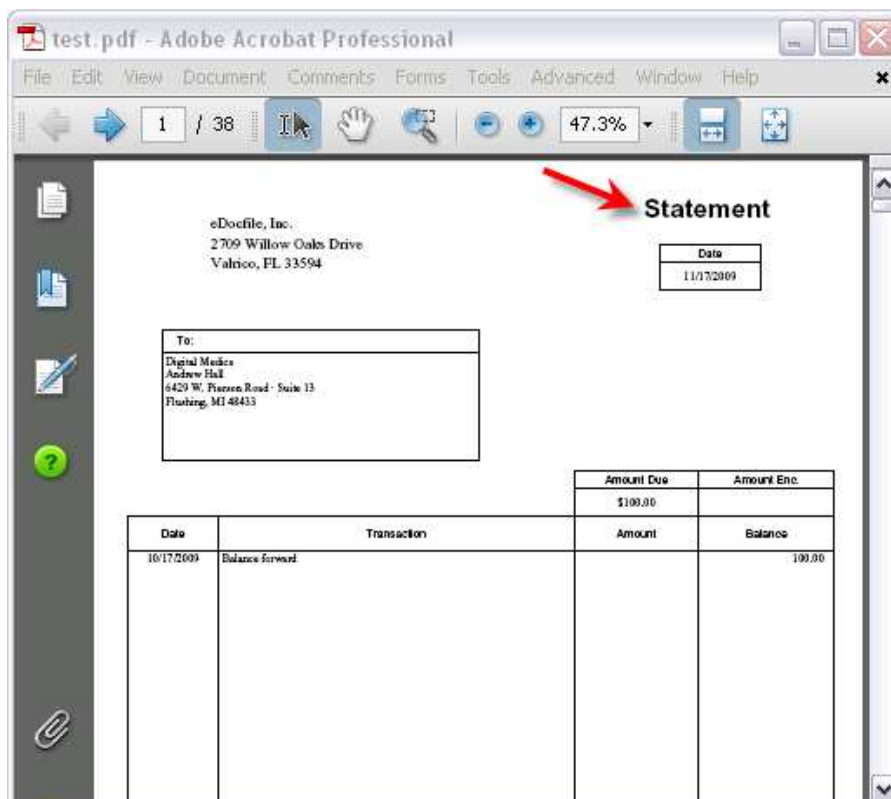
PDF Distributor is a program that will separate a multi-page text searchable PDF and distribute it via email based upon its contents. It is ideally suited for distributing reports and statements without printing and mailing.

How it works

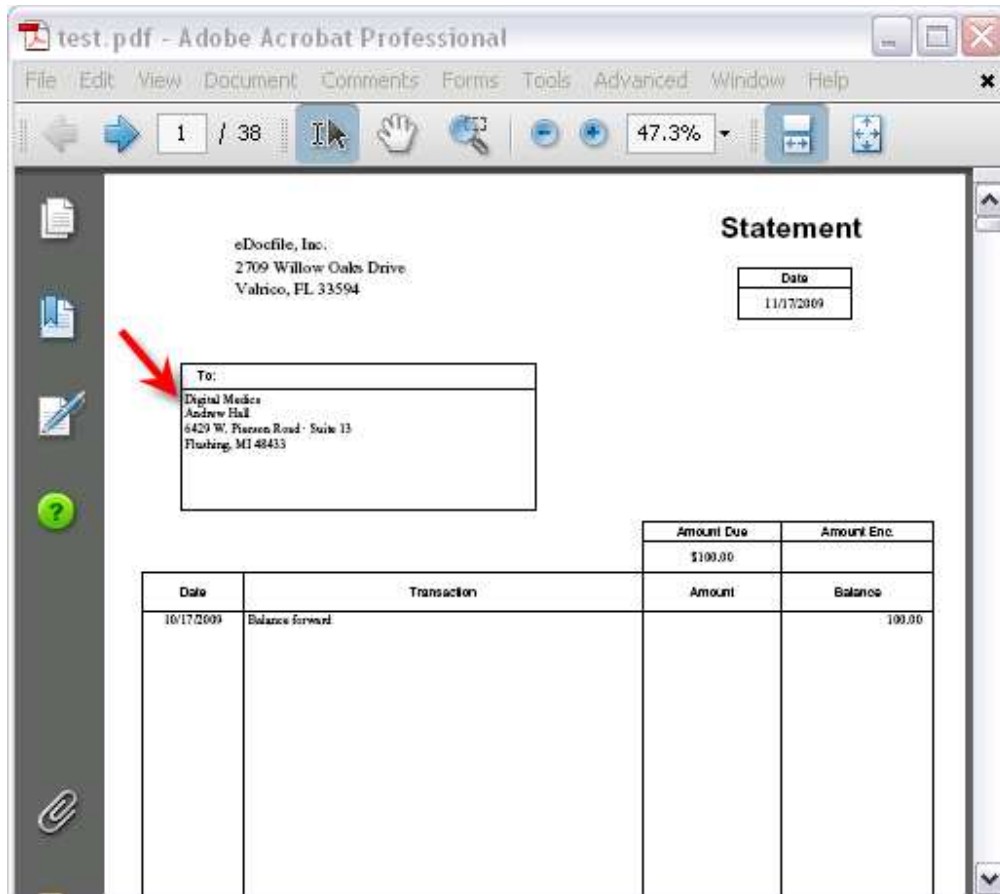
Instead of using a conventional printer, the user prints the statement or report with a program that creates a text searchable PDF. This can be done with the standard version of Acrobat or a freeware program such as PDF Creator. Once created PDF Distributor will split the file into separate documents based upon a unique word that appears on the first page of each document, such as "Statement" or "Page 1". Once the file has been split the program will read a specific area of the document that relates to where it is to be delivered, such as a person's name or department. After the program has this value it will be used to lookup a corresponding email address in Excel or a csv file. When there is a match the file will be delivered via email. If there is no match it will be placed in a file folder for errors and details will be written to a log file.

Understanding the Program

The program will search a *text searchable PDF* and split the file into separate documents based upon a word.

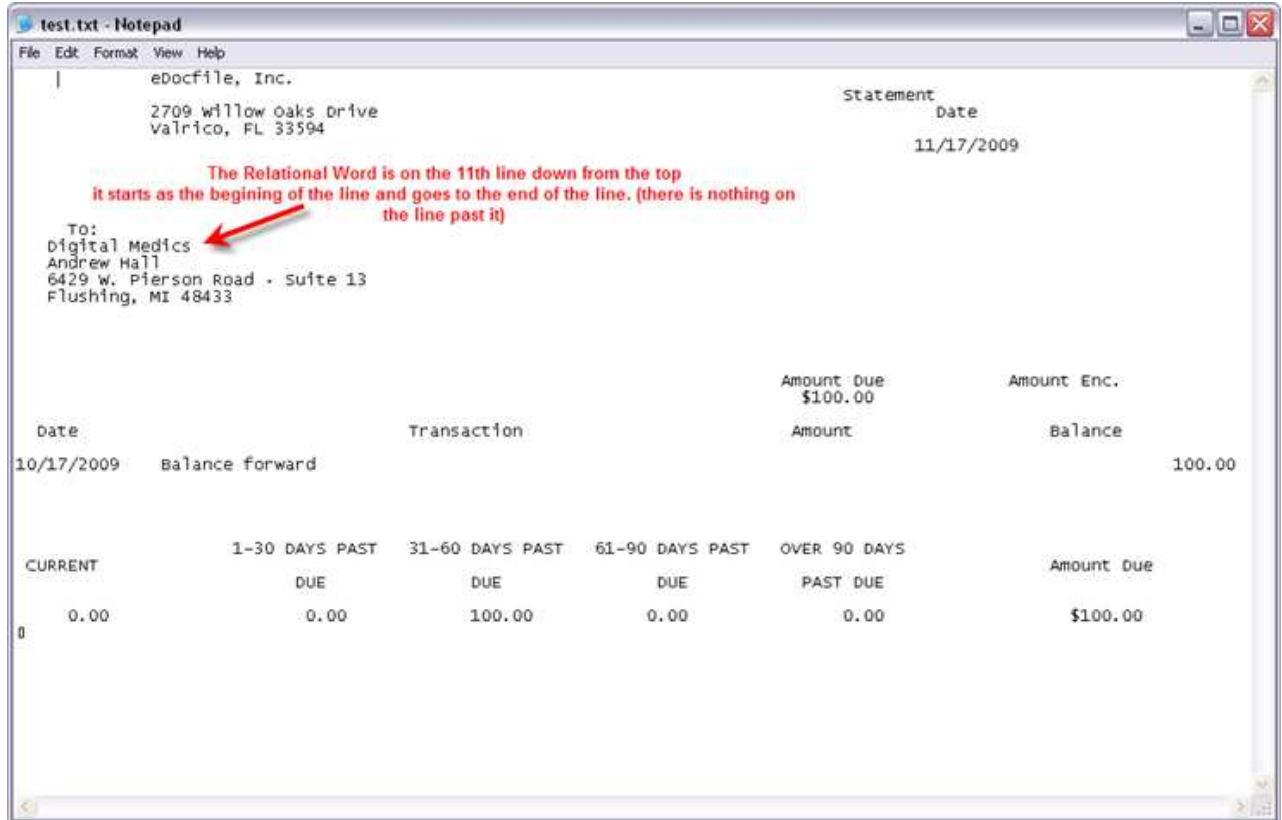


in this example the word "Statement" appears only on the first page of each document so it would be used to separate the file.



The file needs to be delivered to an eMail address at this company. This value will be referred to as the "Relational Word" as it relates the document to the same word (or phrase) in the data table.

Locating the Relational Word



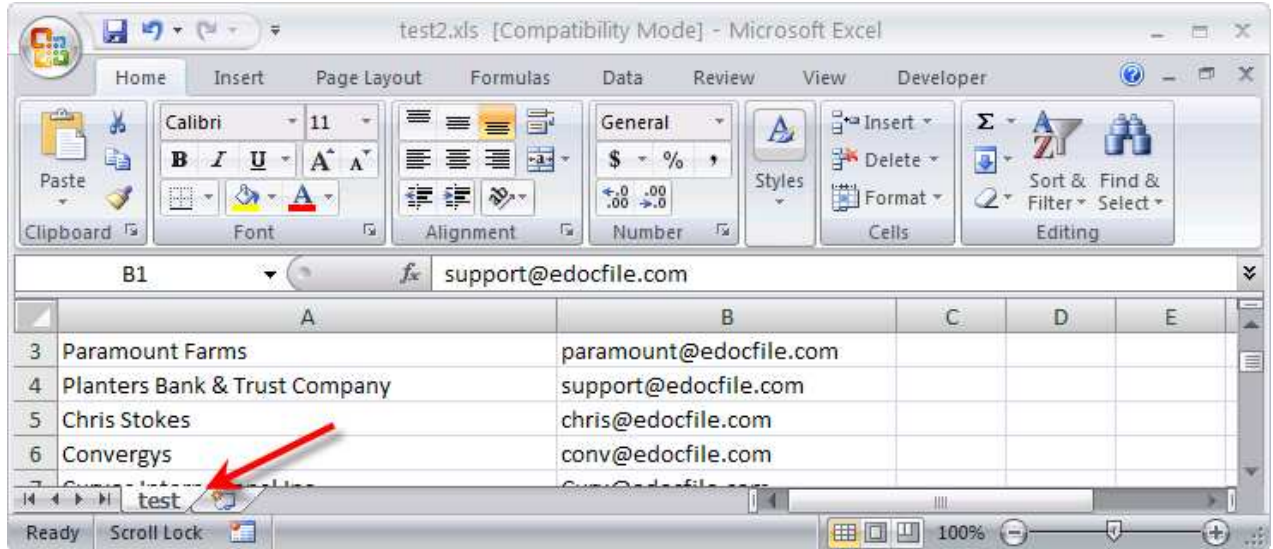
Shown here is the text extracted from the PDF file. The Relational Word is on the 11 line down from the top and continues to the end of the line. The user when in setup would use their arrow keys to count the lines down, the start position and the length of the value to capture. This will be explained in more detail under Setup.

The Data Table Format

Note: If the document is being sent to more than one recipient separate the email addresses with a semicolon ";"



Shown here is a csv file where the values are contained in double quotes. This is the proper format for a csv file. It will also work without double quotes; however, if a comma is found in a company name it will throw off the columns causing an error when sending the eMail.

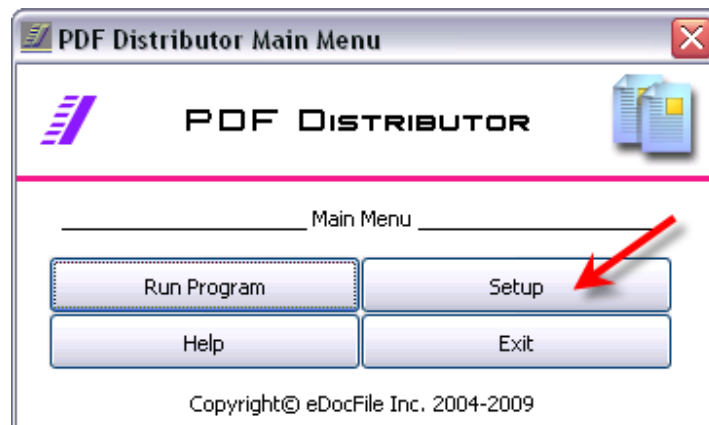


Shown here is a Data Table in an Excel Spreadsheet format. When using this format it is necessary to enter the Data Table name. In this example it would be "test"

Creating the Text Searchable PDF

From your application print your document using the print driver that comes with Acrobat. If Acrobat is not available search on line for "PDF Creator" which is an open source program (freeware) that creates PDF Files. When creating the file the user should ideally make them less than 1000 pages.

Setup



Click on Setup

Setup PDF Distributor

Mail Setup

Mail Server 1

From Address 2

From Name 3

SMTP User ID 4

SMTP Password 5

Port 6

7 SMTP Authentication Required

eMail Subject Line 8

9

Program Settings

Select Input File - Must be a searchable PDF

10

11 12 Separate the PDF on this Word

Set Relational Word Location - this word will relate the PDF to the data table

13 Line Number 14 Start Postion 15 Data Length

Use EOL for end of line

Select Data Table - Excel of CSV File

16

17 Relational Word Column 18 eMail Column 19 Data Table Name - Excel Only

20

21 Show Processing Dialog

22 23 24

Copyright© eDocFile Inc. 2004-2009

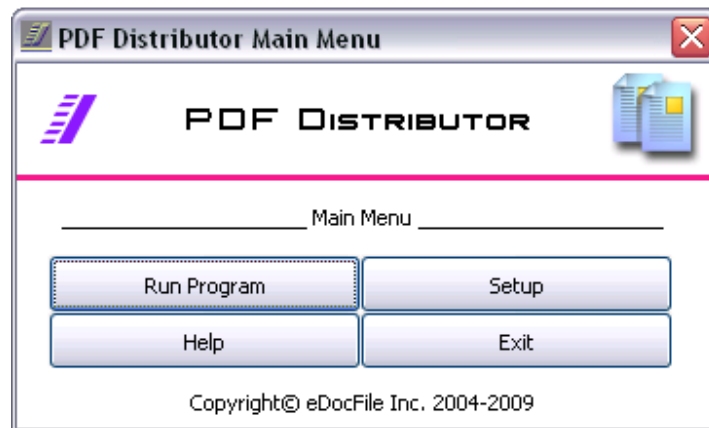
1. Enter a Mail Server
2. Enter the From Address
3. Enter the From Name
4. Enter the User ID for the Mail Server
5. Enter the Password for the Mail Server

6. Enter the Port to use for outgoing mail
7. Place a check mark in this box if Authentication is Required
8. Enter the Subject Line in the eMail
9. Create a plain text file for the body of the eMail and insert the path and name of the file. Please note: All comma's will be removed from the eMail body
10. Enter the name and path of the PDF File to be Split.
11. Click on this button to extract the first page of text from the file selected in number 10. Information from this file will be used in entering 12,13,14 and 15.
12. Enter a word or phrase that appears on the first page of each document, such as "Statement", "Page 1" etc.
13. Insert the line number that contains the Relational Word. To determine the line number place the cursor on the top line of the text file and press the down arrow key while counting until the cursor is on the same line as the Relational Word. The Relational Word is a word that relates the document to a value in a data table. An example would be a Company Name, it is in the document and in the data table. When the program runs it will look in the data table it will pull the eMail address associated with the company.
14. Enter the start position of the Relational Word. To determine this place the cursor on the left of the text file on the same line as the cursor and use the press the right arrow key while counting until it is at the start of the Relational Word.
15. Enter the maximum length of the Relational Word. It is necessary to put the length of the longest value in this box. Begin by placing the cursor at the beginning of the value and press the right arrow key until the next value appears on the same line. If there is nothing else on the same line "EOL" can be used for End of Line.
16. Enter the path and file name of the Data Table. The table needs to be either a csv file or an Excel spreadsheet. If a csv file is being used it should be formatted with double quotes to allow for a comma in the Relational Value. For instance eDocfile, Inc.,info@edocfile.com would not be separated correctly as there is a comma after eDocfile. This would work correctly "eDocfile, Inc.", "info@edocfile.com"
17. Enter the Column in the Data Table that contains the Relational word.
18. Enter the Column in the Data Table that contains the eMail Address.
19. If an Excel file is being used as a Data Table, Enter the Data Table (Sheet) Name
20. Enter an Output Folder. PDF Distributor will save each separated file in this folder along with a log file. It will also create a subfolder called "Errors" which will contain a copy of each PDF it failed to send and a log file of why it failed to send it.
21. Place a check mark in this box to show a dialog of what is being processed.
22. [Click here to test the mail settings](#)

23. Click here to open this Help file

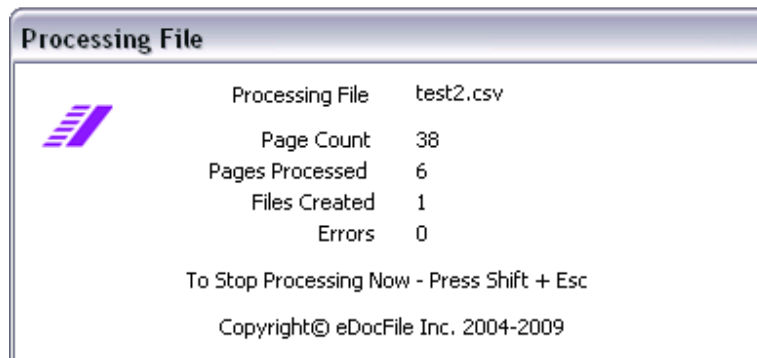
24. Click here to Save and Exit the setup.

Running the Program

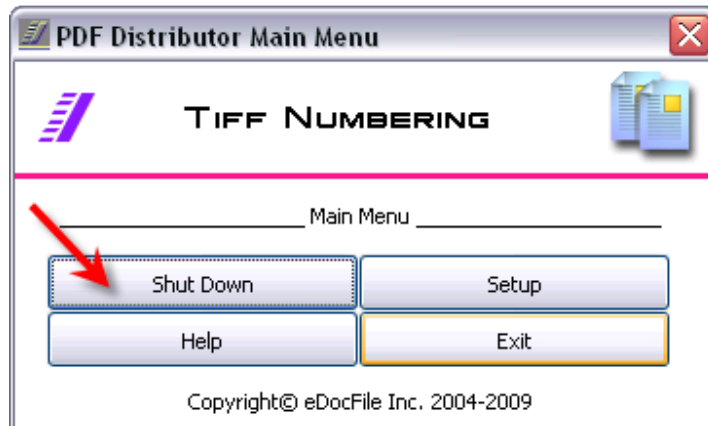


Click on Run Program to start processing

If show progress was selected from the setup this dialog will appear



To shut down the program before it is finished running press Shift + Esc or reopen the application and click on "Shut Down"



Log Files

In the output folder there will be a log file that shows the program's results for each file and a subfolder that will contain another log file that shows only Errors.

The format of the log file is as follows:

```
2009-11-30:08:29:32:000 - File Name = Created File wholedoc20.pdf Extracted Company = Drexel University
2009-11-30:08:29:32:046 -   wholedoc20.pdf is being sent to Drex@edocfile.com at Drexel University
2009-11-30:08:29:34:187 - The server accepted the email address Drex@edocfile.com with the attachment C:\Documents and Settings\Keith Passaur\Application Data\eDocFile\PDF Distributor\output\wholedoc20.pdf
2009-11-30:08:29:34:234 - File Name = Created File wholedoc21.pdf Extracted Company = Selective HR Solutions, Inc.
2009-11-30:08:29:34:281 -

!!!!!!!!!!
Selective HR Solutions, Inc. was not found in the list
!!!!!!!!!!

2009-11-30:08:29:34:625 - File Name = Created File wholedoc24.pdf Extracted Company = Pawtucket School Department
2009-11-30:08:29:34:671 -   wholedoc24.pdf is being sent to PSD@edocfile.com at Pawtucket School Department
```

The format of the error file log is as follows:

```
2009-11-30:08:29:34:281 - Selective HR Solutions, Inc. was not found in the list
```

Data Connection String

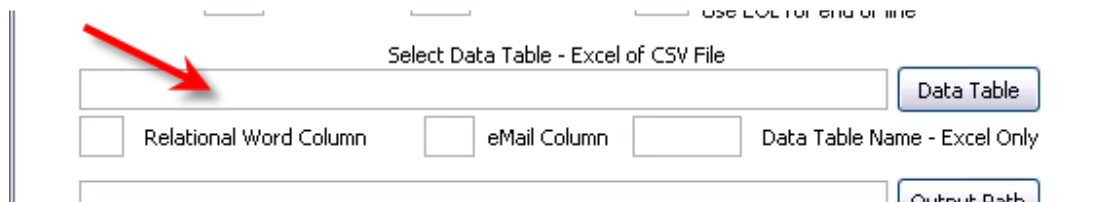
The user can change the Data Connection String and connect directly to their database. However, technical support from eDocfile is not available for this function as there are too many different formats to support. If the user is comfortable with data connection strings the one used by PDF Distributor for processing the csv file can be found here.

C:\Program Files\edocfile\PDF Distributor\connectionstr.ini

and it contains:

```
Provider=Microsoft.Jet.OLEDB.4.0;Data Source=%datapath%;Extended
Properties='text;HDR=NO;FMT=Delimited;Option=1'
```

In connecting directly to a database edit this file. The value above "%datapath%" is the value pulled from the setup that contains the Data Table.



Program Removal

To remove all files created and used by PDF Distributor, go to the Control Panel and under add and remove programs remove the application. Once this is done delete the two following file folders:

C:\Program Files\edocfile\PDF Distributor

C:\Documents and Settings\the user's name\Application Data\edocfile\PDF Distributor

Support

For assistance with either the trial version or full version of PDF Distributor please contact us at:

eDocfile Inc.
2709 Willow Oaks Drive
Valrico, FL 33594
Phone 813-413-5599
sales@edocfile.com