

PDF SPLIT AND MERGE 0.7B1 TUTORIAL

DEVELOPED BY ANDREA VACONDIO

RELEASE 0.7 BETA, 23 JUNE 2007

<http://www.pdftsam.org/>

INDEX

| | |
|---|----|
| • What is pdfsam?..... | 3 |
| • Why pdfsam? | 3 |
| • Install and run | 3 |
| • Merge | 4 |
| • Split | 9 |
| • Settings..... | 12 |
| • Console | 13 |
| • Tweak&Tips | 14 |
| • Thanks to open source projects involved | 15 |
| • License | 16 |

WHAT IS PDFSAM?

PDF Split and Merge is a very simple, easy to use, free, open source utility to split and merge pdf files. It has a simple graphical interface to let the user choose pdf files, split or merge them. Pdfsam is written with Java Swing that is *"a set of Java class libraries provided as part of [Java 2 Platform, Standard Edition \(J2SE\)](#) to support building graphics user interface (GUI) and graphics functionality for client applications that will run on popular platforms such as Microsoft Windows, Linux, and Mac OSX."*¹

Pdfsam uses some external libraries such as:

- itext-2.0.2
- dom4j-1.6.1
- jaxen-1.1
- jcmdline-1.0.3
- looks-2.1.1
- bcmail-jdk14-135.jar
- bcprov-jdk14-135.jar

It uses itext-2.0.2 library to manipulate pdf files, looks-2.1.1 to have a nice and changeable look and feel, jcmdline-1.0.3 to parse and validate console input arguments, dom4j-1.6.1 and jaxen-1.1 to deal with xml format.

WHY PDFSAM?

I needed to merge together some pdf files a couple of years ago and it happened again since that first time. Guess what? I couldn't find an open source software to make it easily and quickly, even if there are tons of commercial software for that. This is the reason why i decided to write pdfsam, a simple, open source and free software for pdf manipulation. Actually one more thing was the iText library; i knew the main job about the pdf manipulation was already done by iText but people need something easier, simple interface, something "click and run". And here it is, pdfsam.

INSTALL AND RUN

Prerequisites:

pdfsam need a working Java Runtime Environment. It has been developed using Sun Java(TM) 2 SDK, Standard Edition, Version 1.4.2 and tested on many Microsoft Windows version and on Ubuntu using J2SE 5.0 and J2SE 1.4.2 .

[pdfsam-0.7-b1.zip](#)

Unzip the archive in a directory and double click the file pdfsam-0.7b1.jar or type the command "java -jar INSTALL_DIR/pdfsam-0.7b1.jar" in a console.

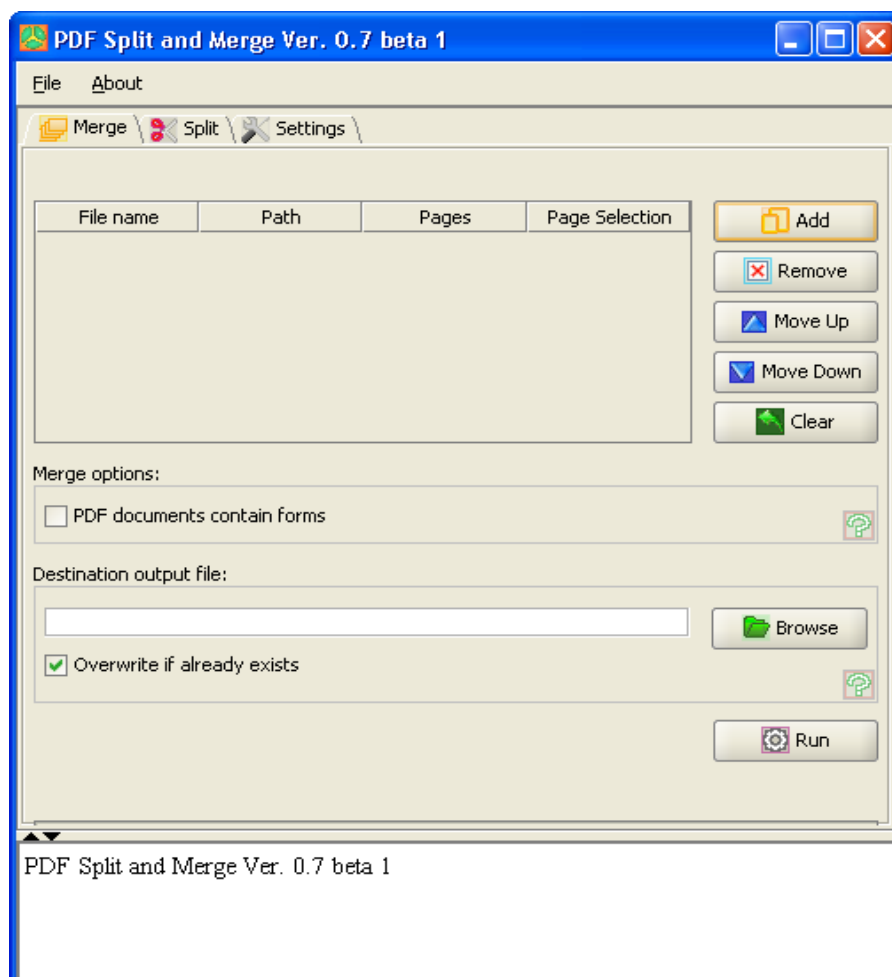
[pdfsam-win32inst-v0 7 b1.exe](#)

This version is for win32 operating systems. Double click it and than, as usual, next -> next -> next ->next... You'll find the pdfsam group in the start menu.

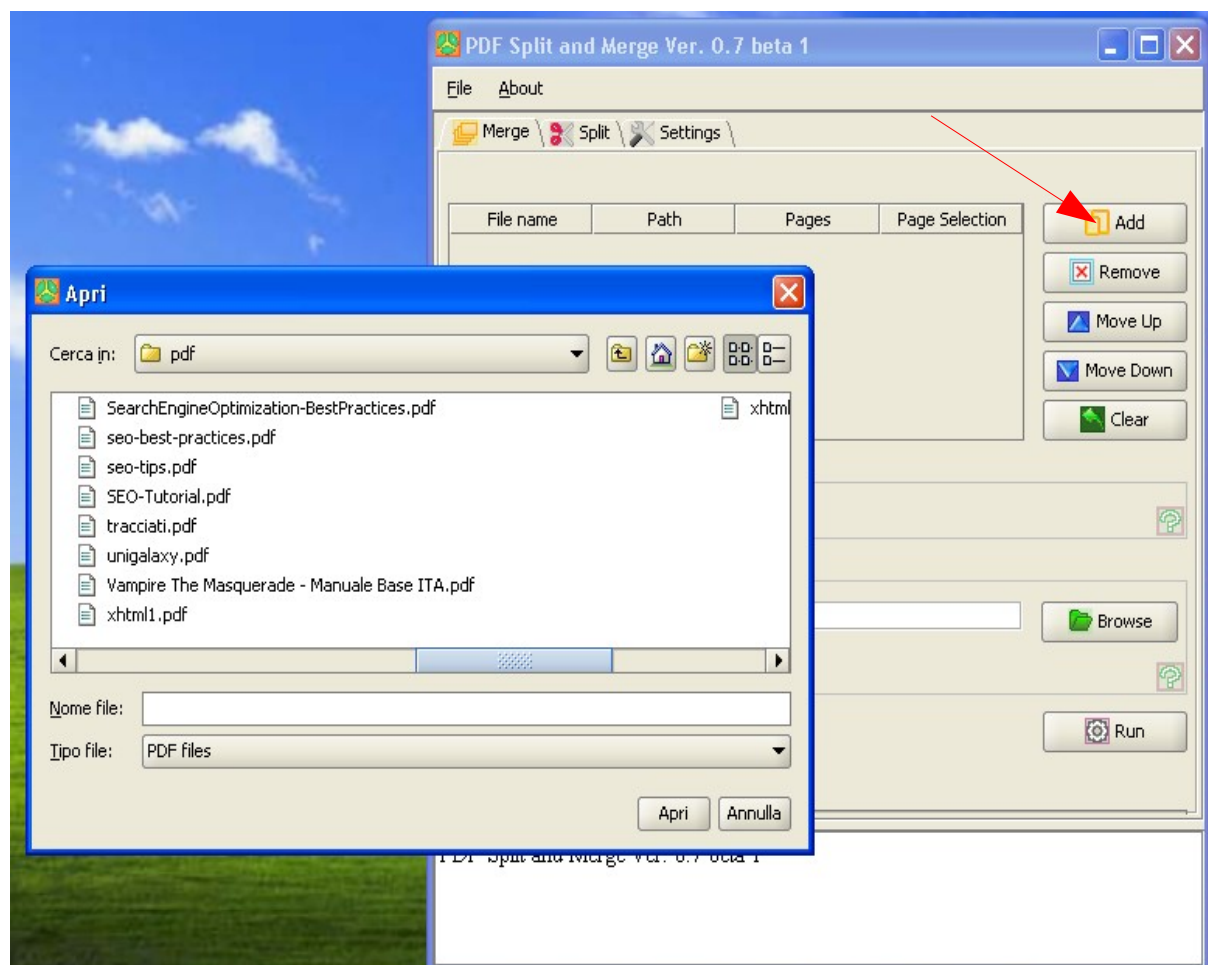
1 From the official <http://java.sun.com>

MERGE

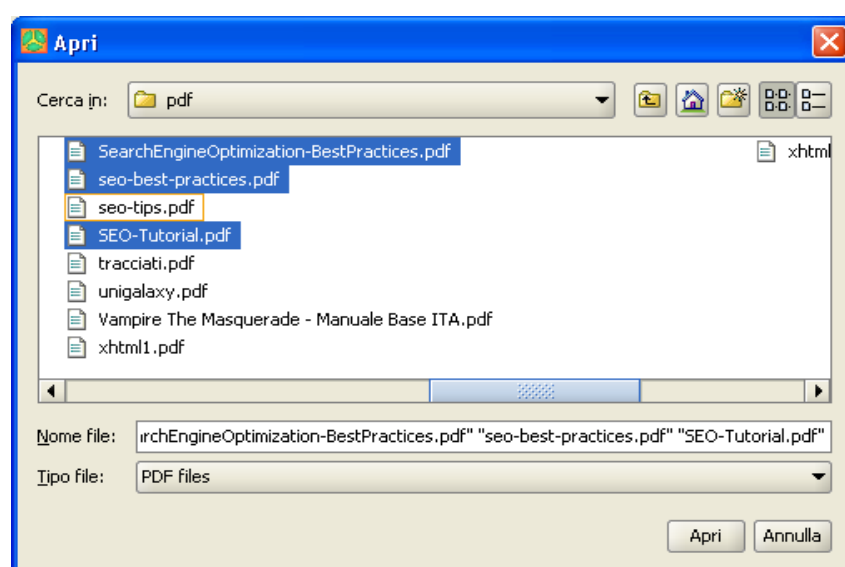
Here is the Merge main window:



Click the "Add" button to choose the pdf files you want to merge together.



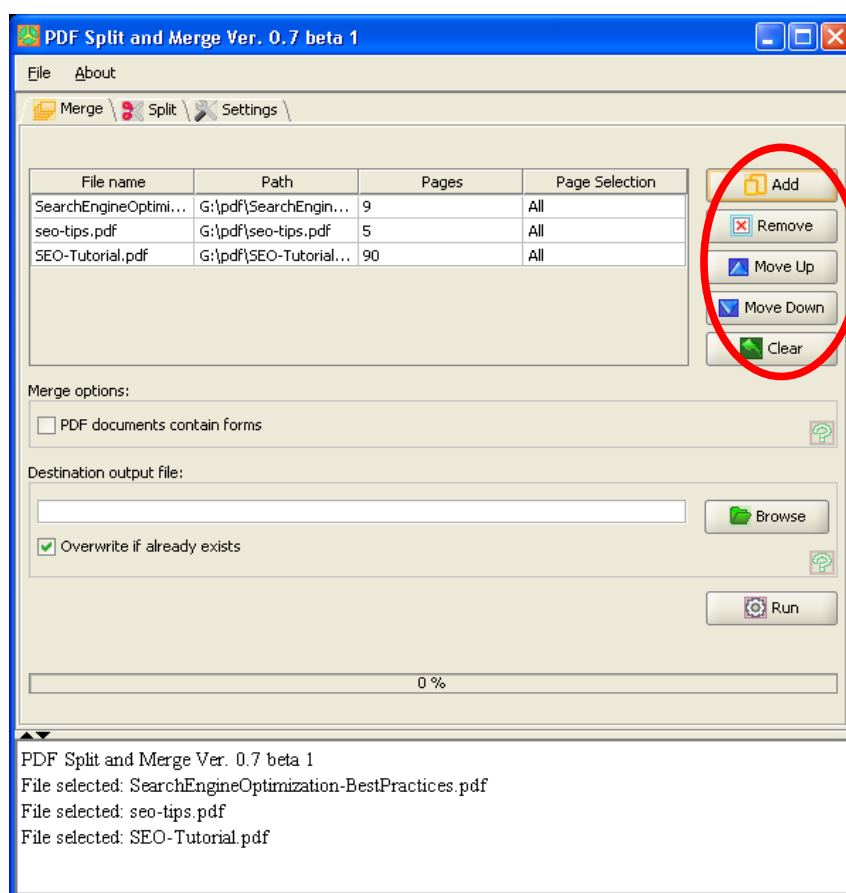
Select pdf files you want to merge; if you want to select more the one file press the "shift" button on you keyboard.



Once your files has been selected you'll see a table containing file ditails:

- File name: The file name
- Path: The file absolute path in the filesystem
- Pages: The total number of pages in the document
- Page Selection: "All" means that you want to merge the whole document, "page-topage" means you want to merge only pages between "page" and "topage" comprehensive, a single page number means that page. Comma separated value are accepted. (ex 4,13-25 if you want to merge page 4 and pages between 13 and 25).

Use the buttons "Move Up", "Move Down" to change the files order or select one or more and drag&drop. Press "Remove" to remove unwanted file from the list.



Double click the Page Selection cell if you want to merge only a subset of the document.

| File name | Path | Pages | Page Selection |
|-----------------------|------------------------|-------|----------------|
| SearchEngineOptimi... | G:\pdf\SearchEngin... | 9 | All |
| seo-tips.pdf | G:\pdf\seo-tips.pdf | 5 | All |
| SEO-Tutorial.pdf | G:\pdf\SEO-Tutorial... | 90 | All |

Fill with the pages subset you want to merge; page numbers are comprehended. You may use print notation like "2,12-16,17" or simply "2" if you want to merge only page number 2.

| File name | Path | Pages | Page Selection |
|-----------------------|------------------------|-------|----------------|
| SearchEngineOptimi... | G:\pdf\SearchEngin... | 9 | All |
| seo-tips.pdf | G:\pdf\seo-tips.pdf | 5 | All |
| SEO-Tutorial.pdf | G:\pdf\SEO-Tutorial... | 90 | 7,23-37 |

Check the box if your documents contain pdf forms. This will tell pdfsam to use a different merge algorithm specific to deal with forms. Setting this options all the documents will be loaded in memory with a very high memory usage.

Merge options:

☐ PDF documents contain forms

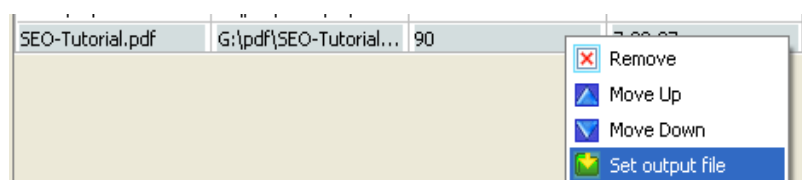
Fill the destination field with the output file name or select it using the "Browse" button. Check or uncheck "Overwrite" if you want overwrite the output pdf document if it already exists.

Destination output file:

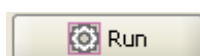
c:\out.pdf Browse

☒ Overwrite if already exists ?

You can automatically set a default output file by clicking with the right button on a file in merge table and selecting "Set output file". This will fill the output field with a default filename in the same directory as the selected file in merge table.

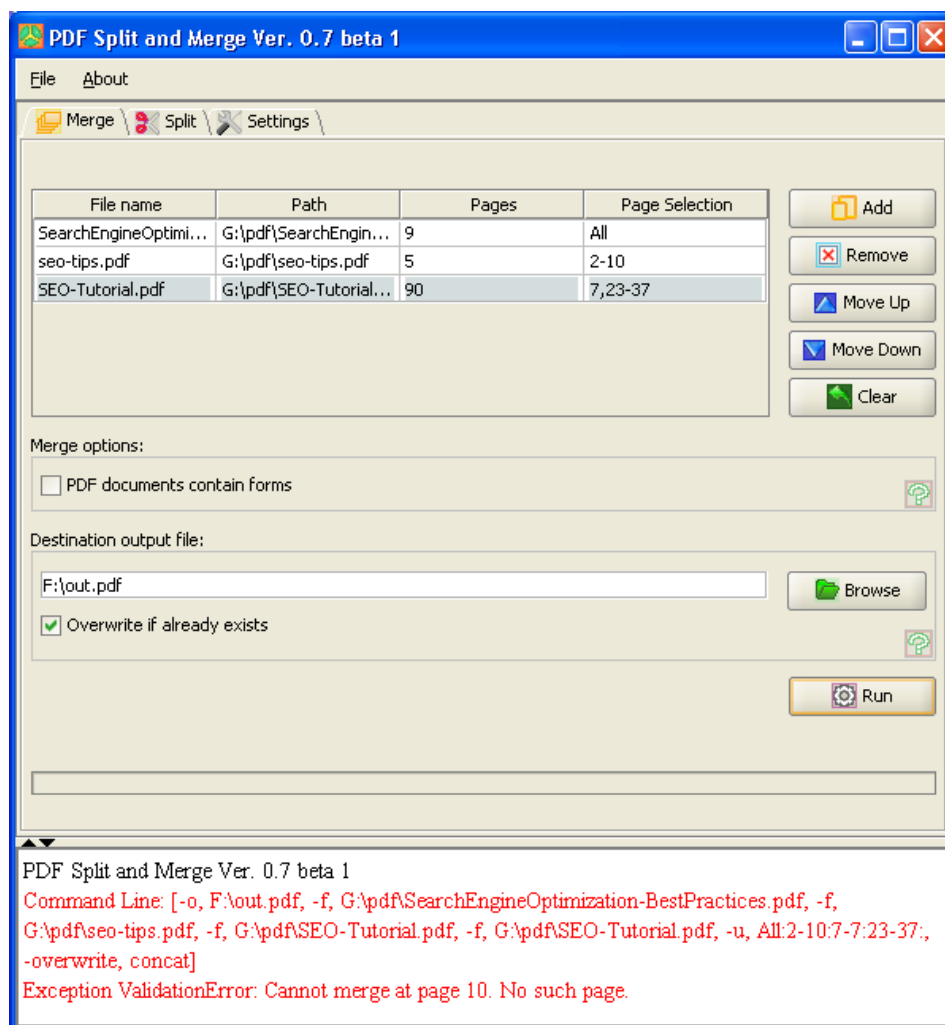


Click the "Run" button.



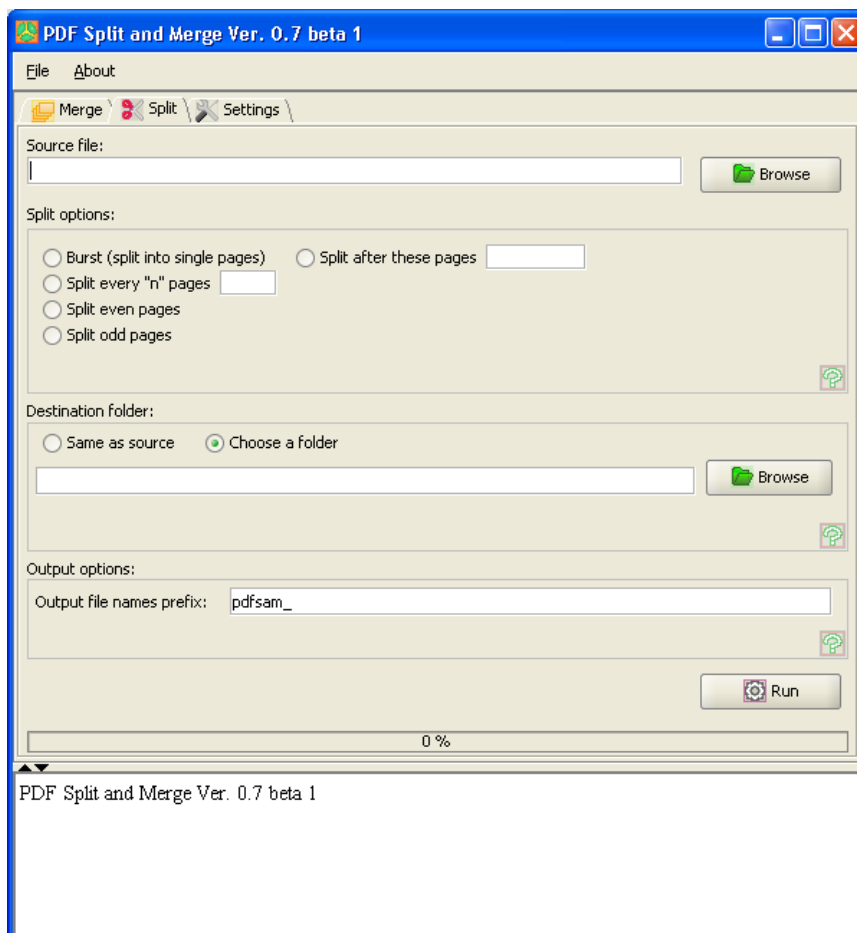
If no error occurs you'll see the progress bar growing till 100% and a black message on the log panel saying "Total processed pages: xxxx-", otherwise you'll find a red error message.

In the example above we had an error message saying we cannot merge pages between 2 and 10 because the document has only 5 pages.

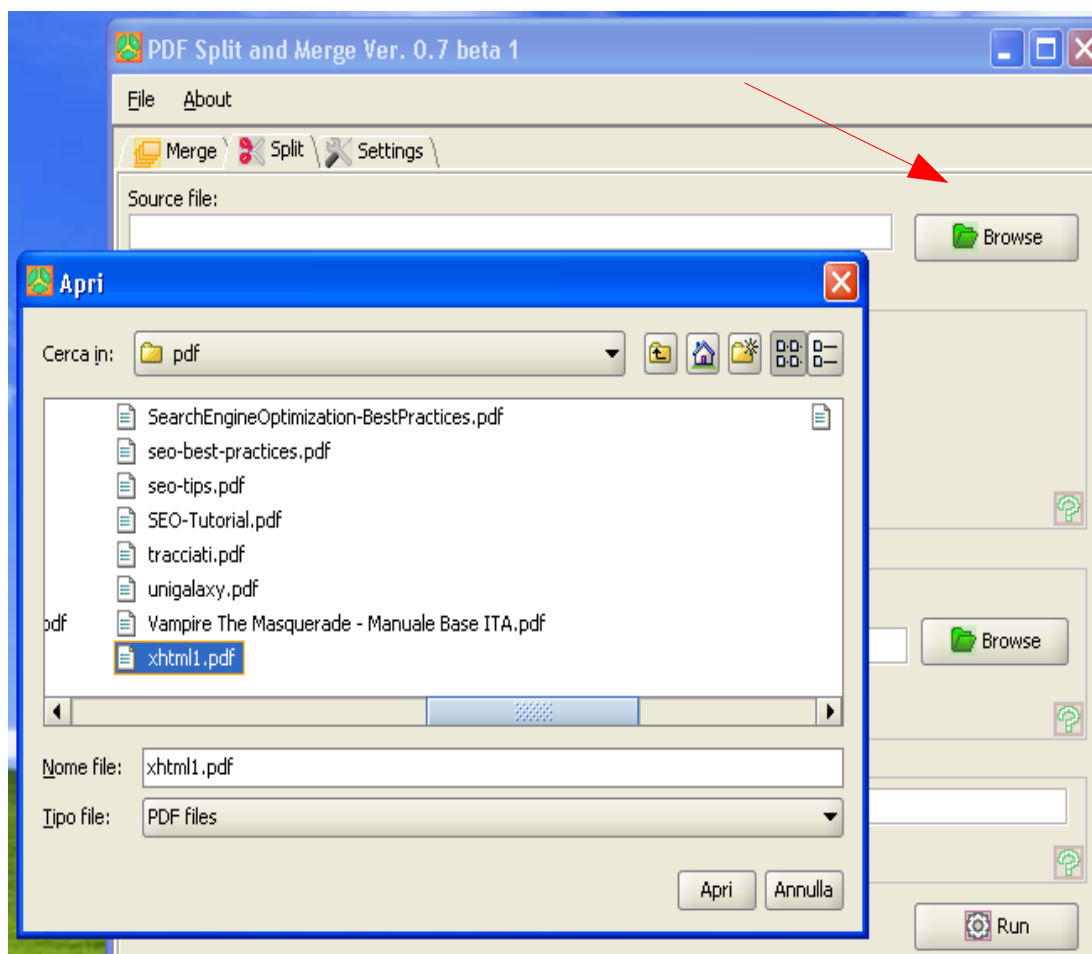


SPLIT

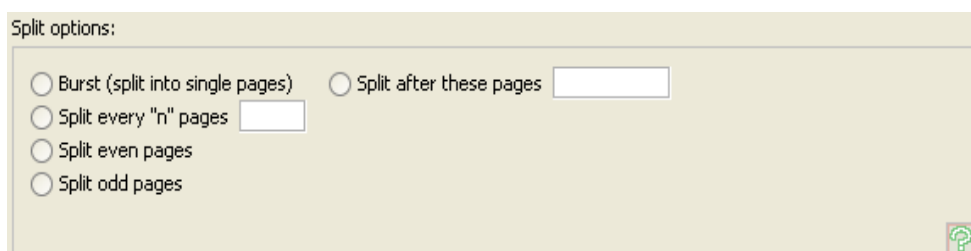
Here is the split main window:



Click the "Browse" button to select the pdf file to split.

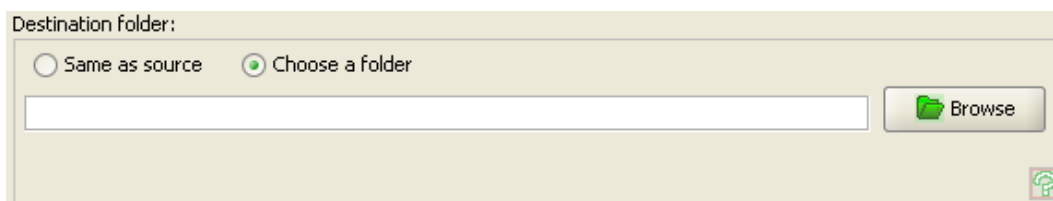


Select the split option you need:




- *Burst* splits the pdf file into single pages, one pdf file-one page.
- *Split every "n" pages* splits the pdf file into files made of "n" pages. (Ex. If you selected a file made of 100 pages and you split it every 25 pages, you'll have 4 pdf files made of 25 pages).
- *Split even pages* splits the pdf file every even page.
- *Split odd pages* splits the pdf file every odd page.
- *Split after these pages* splits the pdf file at the submitted page numbers. Numbers must be separated by "-" (Ex. 12-45-75). You may specify just one page number to split the document in two subdocuments. Note that page numbers will be ordered and zero, negative or too big numbers are pruned.


Select the destination folder for resulting pdf files:



Destination folder:


☐ Same as source ☒ Choose a folder

 Browse




Destination folder can be chosen clicking the "Browse" button and finding the folder on your filesystem or selecting "Same as source", with this option you'll have all output pdf files in the same directory as the source file.

Select output prefix if needed.



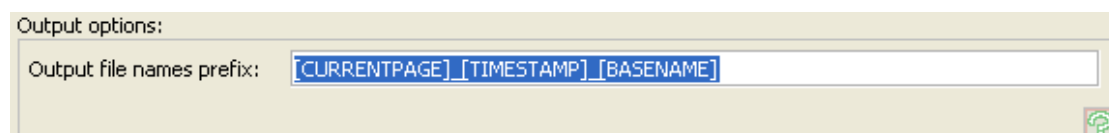
Output options:

Output file names prefix:




Output files will be named with the same name of the input pdf file and with a progressive leading number. You can specify a prefix for the output name and the final result will be something like "0_prefix_filename.pdf".

Pdfsam 0.7 supports complex prefix. If the prefix contains one among [CURRENTPAGE] and [TIMESTAMP] complex prefix are enabled and pdfsam perform variable substitution. Available variables are [CURRENTPAGE], [TIMESTAMP], [BASENAME]. The variable [BASENAME] doesn't enable complex prefix because it doesn't ensure unique output filenames but it can be use together with [CURRENTPAGE] or [TIMESTAMP] or both.



Output options:

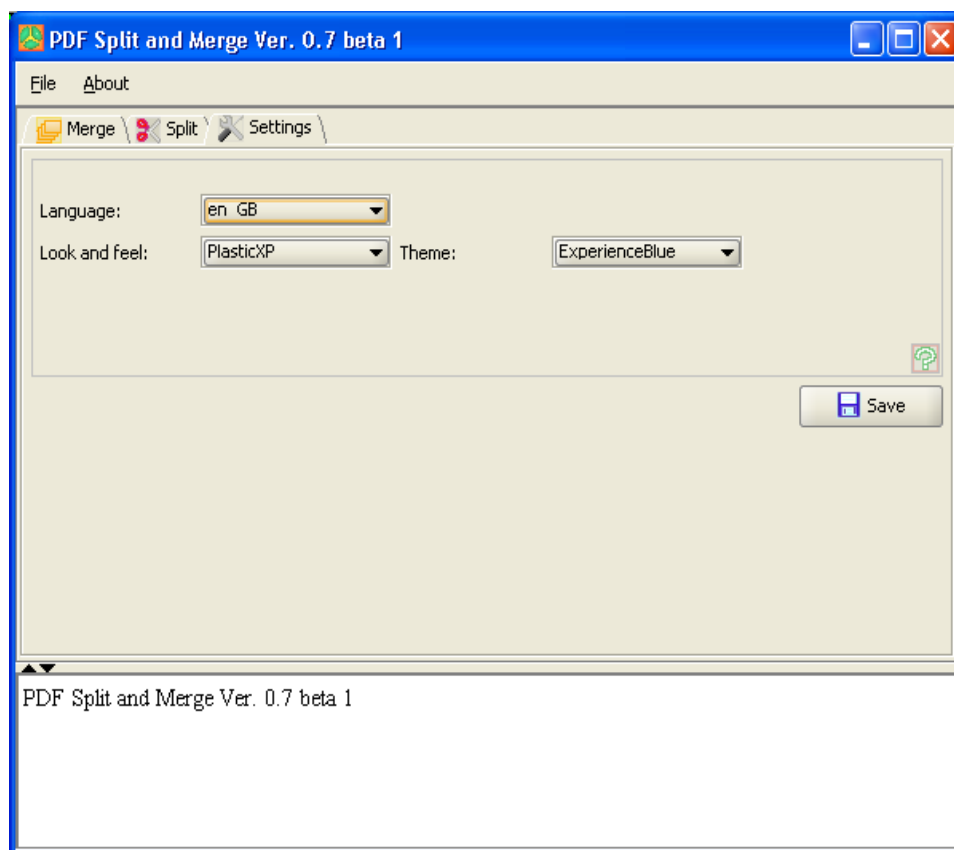
Output file names prefix:



Click "Run" and you find the operation result in the log section.

SETTINGS

This is the settings main window:



Here you can change pdfsam language and look and feel. Click “Save” after your changes but remember that every changed option will be visible after a pdfsam restart.

CONSOLE

pdfsam-console is a command line application. It uses jcmdline to parse input parameters and it executes the proper command. It's the core application and it provides pdf manipulation. The main GUI loads plugins, once clicked the plugin "Run" button (merge or split), the plugin performs input data validation and, if everything is correct, it creates an arguments array and then it sends it to the console.

To run the console type the following command: `java -jar /PATH_TO_PDFSAM/lib/pdfsam-console-CURRENT_VERSION.jar`

Here is a list of the console arguments:

Usage: Pdfsam options command

where:

command = command to execute {[concat], [split]} (required)

and options are:

Merge

-?

prints usage to stdout; exits (optional)

-f <file>

pdf files to concat. a list of existing pdf files (EX. -f /tmp/file1.pdf -f /tmp/file2.pdf)

-h

prints usage to stdout; exits (optional)

-help

displays verbose help information (optional)

-l <file>

csv file containing pdf files list to concat in comma separated value format (optional)

-log <file>

text file to log output messages (optional)

-o <file>

pdf output file. if it doesn't exist it's created, if it exists it must be writeable (required).

-u <string>

page selection script. You can set a subset of pages to merge. Accepted values: "all" or "num1-num2" separated by ":" (EX. -f /tmp/file1.pdf -f /tmp/file2.pdf -u all:all:), (EX. -f /tmp/file1.pdf -f /tmp/file2.pdf -u all:12-14:) to merge file1.pdf and pages 12,13,14 of file2.pdf. If -u is not set default behaviour is to merge document completely (optional).

-overwrite

if passed, overwrite existing output file

-copyfields

use this if input documents contain forms. This option will lead to a high memory usage.

-version

displays command's version (optional).

Split

-?

prints usage to stdout; exits (optional)

-f <file>

input pdf file to split (required)

- h
prints usage to stdout; exits (optional)
- help
displays verbose help information (optional)
- n <num>
page number to split at if -s is SPLIT or NSPLIT (optional)
- log <file>
text file to log output messages (optional)
- o <dir>
output directory (required).
- overwrite
overwrite existing output file (optional)
- p <s>
prefix for the output files name (optional). If it contains "[CURRENTPAGE]" or "[TIMESTAMP]" it performs variable substitution. (Ex. [BASENAME]_prefix_[CURRENTPAGE] generates FileName_prefix_005.pdf). Available prefix variables: [CURRENTPAGE], [TIMESTAMP], [BASENAME]
- s <string>
split type {[BURST], [ODD], [EVEN], [SPLIT], [NSPLIT]} (required)
- version
displays command's version (optional).

Examples:

java -jar pdfsam-console-0.4.3.jar -f c:\t.pdf -o c:\ -s ODD split

ava -jar pdfsam-console-0.4.3.jar -f c:\file1.pdf -f c:\file2.pdf -o c:\out_file.pdf -u all:1-3: concat

TWEAK&TIPS

When you see this



move the mouse over it to have some tooltip help.

If you have some trouble loading plugins try this:

- Open the file config.xml with good editor (Notepad+..) .
- Fill the tag *"plugins_absolute_dir"* with the absolute path to your plugins directory .
- Save .

pdfsam tries to find plugins directory but, if there's some trouble, you can specify the absolute path as described.

You can extract portion of a single file:

- In merge section select only one file.
- Set a the page subset you need to extract (Ex. 12-23).
- Select an output file.
- Merge.

THANKS TO OPEN SOURCE PROJECTS INVOLVED

This simple tool is free software and this has been possible thanks to many other open source projects.

I'd like to thank:

Ubuntu linux distro because most of the time i work on pdfsam at home, on my personal pc.

Eclipse, the development platform i used.

OpenOffice, the suite i used to write this document.

Notepad++, the editor i used many times to make changes in text files (php, txt etc) while working on win32 platform.

SourceForge because it hosts pdfsam givin' me a lot of usefull services.

WordPress for the website.

phpBB for the forum.

iText, jGoodies, jcmdline, jaxen, dom4j libraries to make pdfsam a working tool.

Elisa Bortolotti for her icons.

Free Software Foundation for the GPL license.

All the pdfsam community for contributions, tests, support and suggestions.

LICENSE

The GNU General Public License (GPL)

Version 2, June 1991

Copyright (C) 1989, 1991 Free Software Foundation, Inc.
59 Temple Place, Suite 330, Boston, MA 02111-1307 USA

Everyone is permitted to copy and distribute verbatim copies
of this license document, but changing it is not allowed.

TERMS AND CONDITIONS FOR COPYING, DISTRIBUTION AND MODIFICATION

0. This License applies to any program or other work which contains a notice placed by the copyright holder saying it may be distributed under the terms of this General Public License. The "Program", below, refers to any such program or work, and a "work based on the Program" means either the Program or any derivative work under copyright law: that is to say, a work containing the Program or a portion of it, either verbatim or with modifications and/or translated into another language. (Hereinafter, translation is included without limitation in the term "modification".) Each licensee is addressed as "you".

Activities other than copying, distribution and modification are not covered by this License; they are outside its scope. The act of running the Program is not restricted, and the output from the Program is covered only if its contents constitute a work based on the Program (independent of having been made by running the Program). Whether that is true depends on what the Program does.

1. You may copy and distribute verbatim copies of the Program's source code as you receive it, in any medium, provided that you conspicuously and appropriately publish on each copy an appropriate copyright notice and disclaimer of warranty; keep intact all the notices that refer to this License and to the absence of any warranty; and give any other recipients of the Program a copy of this License along with the Program.

You may charge a fee for the physical act of transferring a copy, and you may at your option offer warranty protection in exchange for a fee.

2. You may modify your copy or copies of the Program or any portion of it, thus forming a work based on the Program, and copy and distribute such modifications or work under the terms of Section 1 above, provided that you also meet all of these conditions:

a) You must cause the modified files to carry prominent notices stating that you changed the files and the date of any change.

b) You must cause any work that you distribute or publish, that in whole or in part contains or is derived from the Program or any part thereof, to be licensed as a whole at no charge to all third parties under the terms of this License.

c) If the modified program normally reads commands interactively when run, you must cause it,

when started running for such interactive use in the most ordinary way, to print or display an announcement including an appropriate copyright notice and a notice that there is no warranty (or else, saying that you provide a warranty) and that users may redistribute the program under these conditions, and telling the user how to view a copy of this License. (Exception: if the Program itself is interactive but does not normally print such an announcement, your work based on the Program is not required to print an announcement.)

These requirements apply to the modified work as a whole. If identifiable sections of that work are not derived from the Program, and can be reasonably considered independent and separate works in themselves, then this License, and its terms, do not apply to those sections when you distribute them as separate works. But when you distribute the same sections as part of a whole which is a work based on the Program, the distribution of the whole must be on the terms of this License, whose permissions for other licensees extend to the entire whole, and thus to each and every part regardless of who wrote it.

Thus, it is not the intent of this section to claim rights or contest your rights to work written entirely by you; rather, the intent is to exercise the right to control the distribution of derivative or collective works based on the Program.

In addition, mere aggregation of another work not based on the Program with the Program (or with a work based on the Program) on a volume of a storage or distribution medium does not bring the other work under the scope of this License.

3. You may copy and distribute the Program (or a work based on it, under Section 2) in object code or executable form under the terms of Sections 1 and 2 above provided that you also do one of the following:

a) Accompany it with the complete corresponding machine-readable source code, which must be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

b) Accompany it with a written offer, valid for at least three years, to give any third party, for a charge no more than your cost of physically performing source distribution, a complete machine-readable copy of the corresponding source code, to be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

c) Accompany it with the information you received as to the offer to distribute corresponding source code. (This alternative is allowed only for noncommercial distribution and only if you received the program in object code or executable form with such an offer, in accord with Subsection b above.)

The source code for a work means the preferred form of the work for making modifications to it. For an executable work, complete source code means all the source code for all modules it contains, plus any associated interface definition files, plus the scripts used to control compilation and installation of the executable. However, as a special exception, the source code distributed need not include anything that is normally distributed (in either source or binary form) with the major components (compiler, kernel, and so on) of the operating system on which the executable runs, unless that component itself accompanies the executable.

If distribution of executable or object code is made by offering access to copy from a designated place,

then offering equivalent access to copy the source code from the same place counts as distribution of the source code, even though third parties are not compelled to copy the source along with the object code.

4. You may not copy, modify, sublicense, or distribute the Program except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense or distribute the Program is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

5. You are not required to accept this License, since you have not signed it. However, nothing else grants you permission to modify or distribute the Program or its derivative works. These actions are prohibited by law if you do not accept this License. Therefore, by modifying or distributing the Program (or any work based on the Program), you indicate your acceptance of this License to do so, and all its terms and conditions for copying, distributing or modifying the Program or works based on it.

6. Each time you redistribute the Program (or any work based on the Program), the recipient automatically receives a license from the original licensor to copy, distribute or modify the Program subject to these terms and conditions. You may not impose any further restrictions on the recipients' exercise of the rights granted herein. You are not responsible for enforcing compliance by third parties to this License.

7. If, as a consequence of a court judgment or allegation of patent infringement or for any other reason (not limited to patent issues), conditions are imposed on you (whether by court order, agreement or otherwise) that contradict the conditions of this License, they do not excuse you from the conditions of this License. If you cannot distribute so as to satisfy simultaneously your obligations under this License and any other pertinent obligations, then as a consequence you may not distribute the Program at all. For example, if a patent license would not permit royalty-free redistribution of the Program by all those who receive copies directly or indirectly through you, then the only way you could satisfy both it and this License would be to refrain entirely from distribution of the Program.

If any portion of this section is held invalid or unenforceable under any particular circumstance, the balance of the section is intended to apply and the section as a whole is intended to apply in other circumstances.

It is not the purpose of this section to induce you to infringe any patents or other property right claims or to contest validity of any such claims; this section has the sole purpose of protecting the integrity of the free software distribution system, which is implemented by public license practices. Many people have made generous contributions to the wide range of software distributed through that system in reliance on consistent application of that system; it is up to the author/donor to decide if he or she is willing to distribute software through any other system and a licensee cannot impose that choice.

This section is intended to make thoroughly clear what is believed to be a consequence of the rest of this License.

8. If the distribution and/or use of the Program is restricted in certain countries either by patents or by copyrighted interfaces, the original copyright holder who places the Program under this License may

add an explicit geographical distribution limitation excluding those countries, so that distribution is permitted only in or among countries not thus excluded. In such case, this License incorporates the limitation as if written in the body of this License.

9. The Free Software Foundation may publish revised and/or new versions of the General Public License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns.

Each version is given a distinguishing version number. If the Program specifies a version number of this License which applies to it and "any later version", you have the option of following the terms and conditions either of that version or of any later version published by the Free Software Foundation. If the Program does not specify a version number of this License, you may choose any version ever published by the Free Software Foundation.

10. If you wish to incorporate parts of the Program into other free programs whose distribution conditions are different, write to the author to ask for permission. For software which is copyrighted by the Free Software Foundation, write to the Free Software Foundation; we sometimes make exceptions for this. Our decision will be guided by the two goals of preserving the free status of all derivatives of our free software and of promoting the sharing and reuse of software generally.

NO WARRANTY

11. BECAUSE THE PROGRAM IS LICENSED FREE OF CHARGE, THERE IS NO WARRANTY FOR THE PROGRAM, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE THE PROGRAM "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU. SHOULD THE PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION.

12. IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO MAY MODIFY AND/OR REDISTRIBUTE THE PROGRAM AS PERMITTED ABOVE, BE LIABLE TO YOU FOR DAMAGES, INCLUDING ANY GENERAL, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE THE PROGRAM (INCLUDING BUT NOT LIMITED TO LOSS OF DATA OR DATA BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY YOU OR THIRD PARTIES OR A FAILURE OF THE PROGRAM TO OPERATE WITH ANY OTHER PROGRAMS), EVEN IF SUCH HOLDER OR OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

END OF TERMS AND CONDITIONS