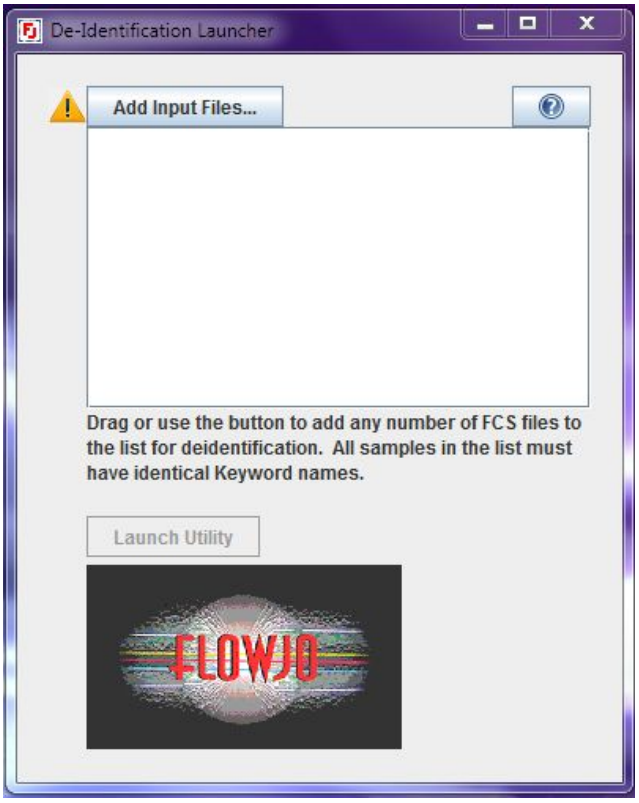
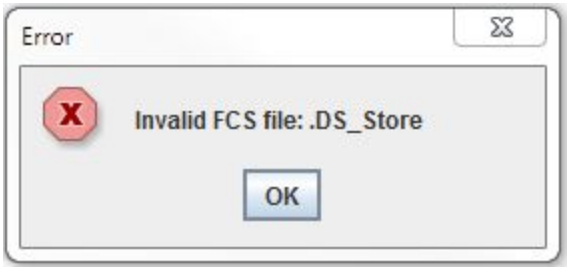
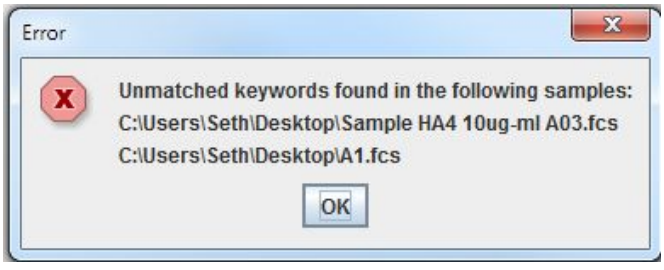
De-Identification Utility: A Quick Tutorial

Double-click the deidentify.jar file to run it on any computer with Java 1.6 or higher.



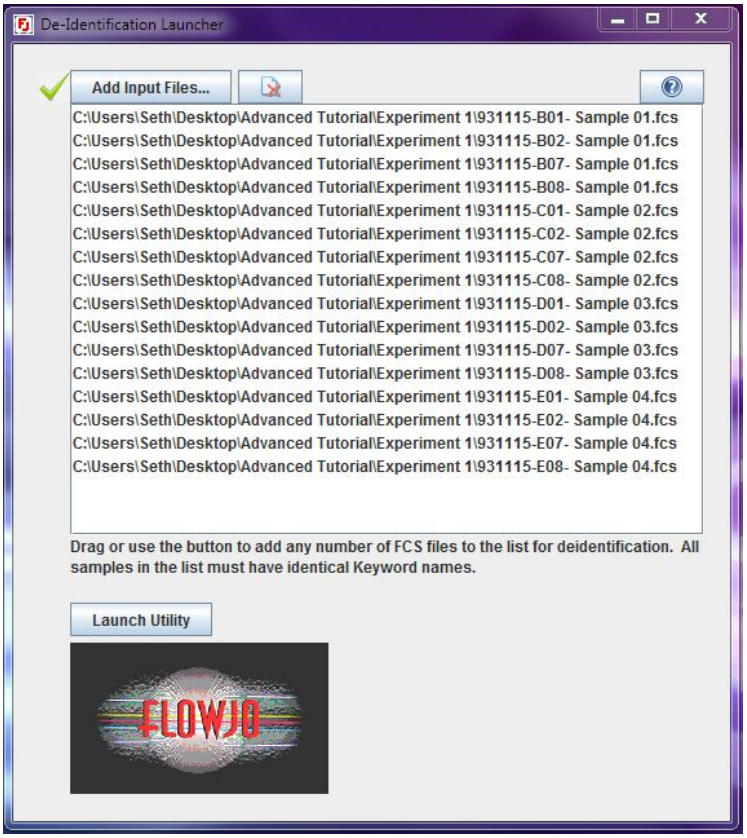
This brings up the launcher, where you will load files to be batch processed. Files must be valid FCS files and must have the EXACT same set of keywords. If either of these requirements are not met, an error dialog will notify you that you are unable to proceed.



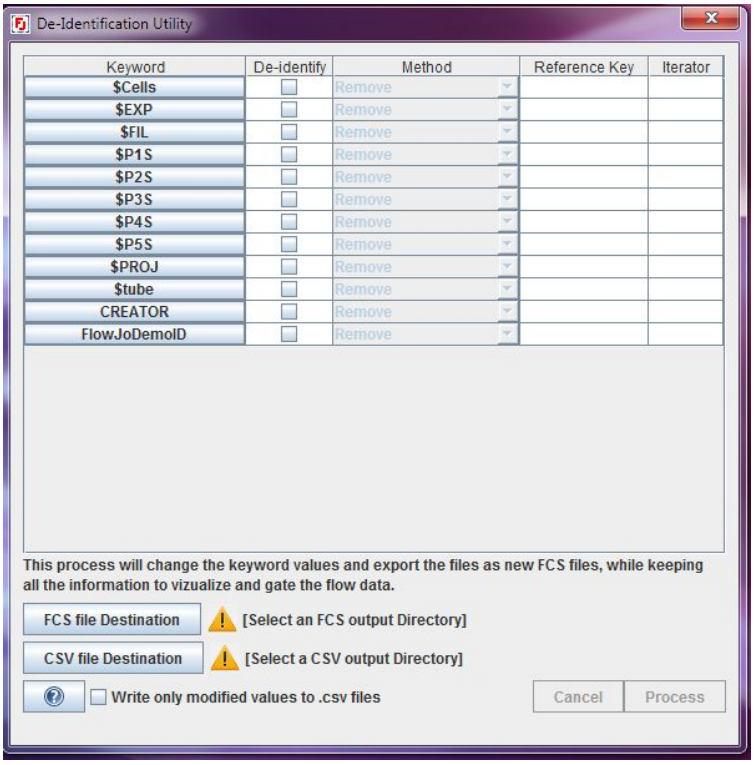


Notice the warning icon by the "Add Input Files..." button and that the "Launch Utility" button is disabled. Begin by adding files from the top dialog button, or dragging in files

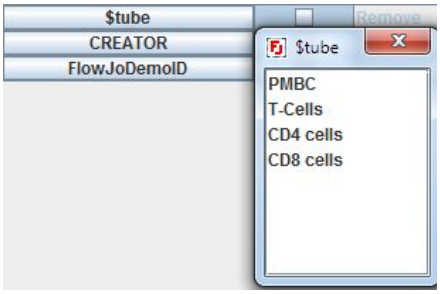
or folders to white list-area. Dragging a folder will add all the files and folders within it, recursively. Note that in extreme cases this can be time consuming and may hang the program depending on the complexity of the folder's tree structure (don't try to drag in your root folder). Once there are files in the list, a delete button will appear next to the add button. Select one or more files and either press this button or the delete key to remove files from the list.



Moving along...



Now that we have successfully launched our set of files to deidentify, we are ready to begin modifying keywords with the utility dialog. Beginning with the first column, you should see each keyword in the files represented as a button. Clicking this button brings up a popup with the list of the union of the keyword values.

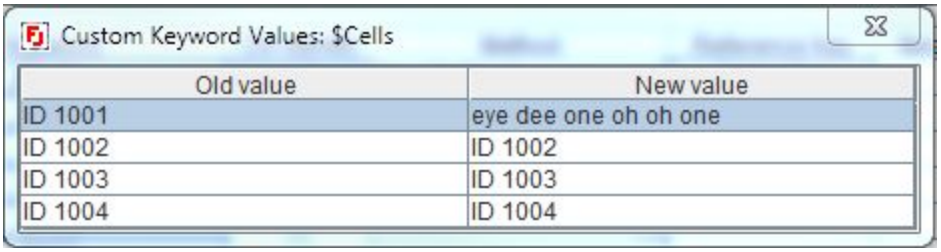


You may notice that not all keywords are available for deidentification. This is because many keywords are required by the data standard, so only non-essential keywords are shown in the list. One caveat is that although $PnS and $PnN are available for deidentification, modification may cause issues with older workspaces. The user will be warned when attempting to modify one of these values.

Now that we know which keywords contain the values we wish to deidentify, select the checkbox for the associated row within the table. This will enable the "method" column's dropdown box in order to qualify the type of deidentification. There are three choices:

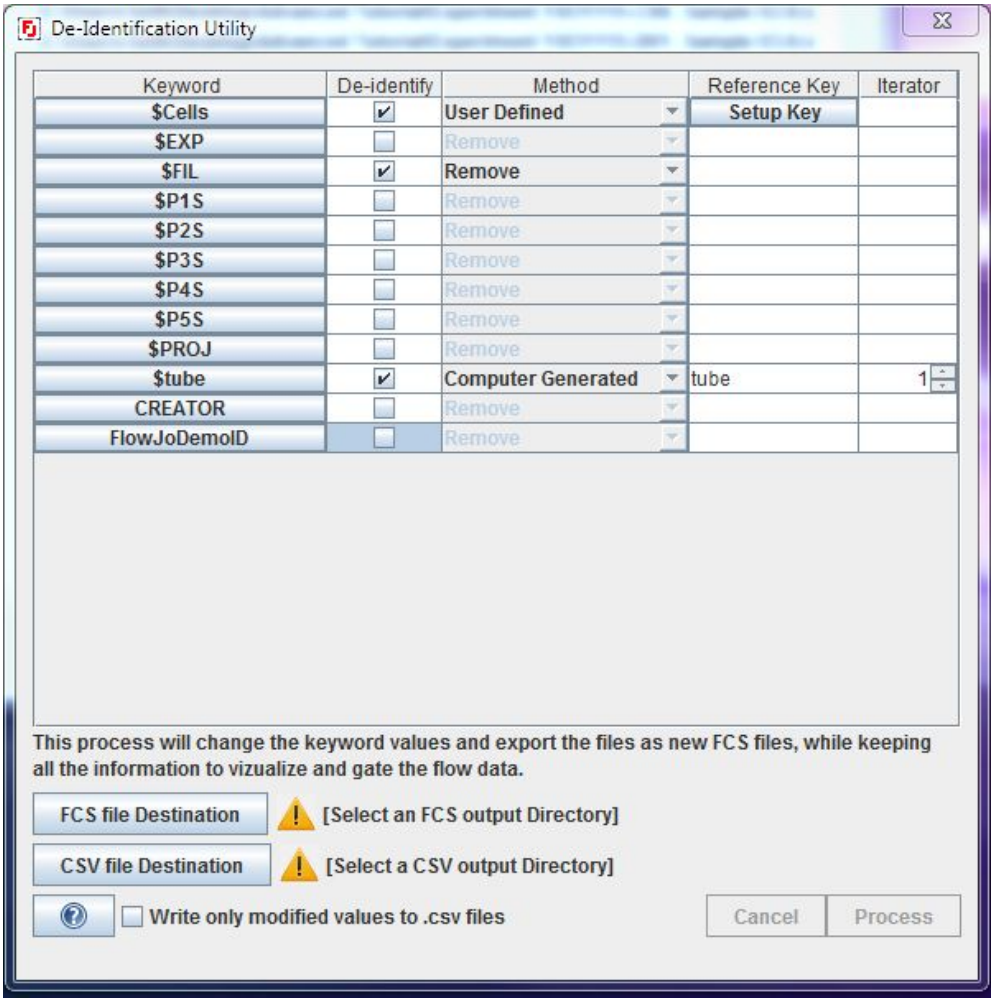
1: The simplest is REMOVE. This method requires no more adjustment, and completely removes the keyword and value from the output files.

2: The second is USER DEFINED. This method functions much like a search and replace routine. Select this option and a button appears to set up a custom key. Click this button to open a reference table. The table contains two columns: an old keyword value (from the input file) and the new keyword value you wish to write to the output file. For the selected keyword, each instance of the old value will be replaced with the new one. Double click on a cell in the new value column to change the text. Note that due to restrictions in the FCS data format, you will encounter errors when including non-ascii characters, or when using the file-specific delimiter(e.g. " / ") as the first or last character in the keyword value. Close the dialog to save your reference key.



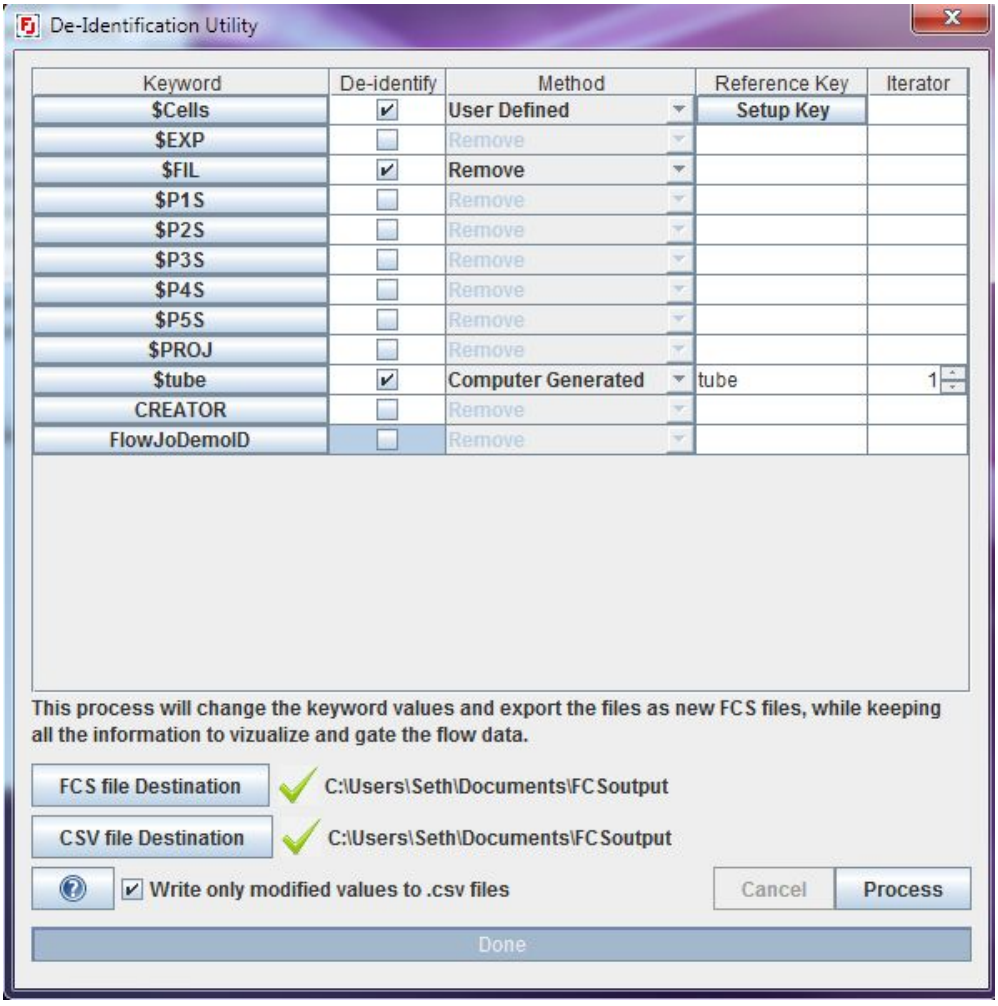
3: The third is COMPUTER GENERATED. This method uses a numerical iterator to automatically rename your keyword values. Select this option and double click on the now-enabled text field in column four to enter a reference prefix. Note that the same ascii and delimiter restrictions exist as with the USER DEFINED values. In the last column, you can select the starting value of the iterator, from 0-9999. Since there are four distinct keyword values for $tube, the output file will have four generated keyword values (e.g. tube0001, tube0002, tube0003, and tube0004) for use as replacement.

Now the table manipulation is complete.

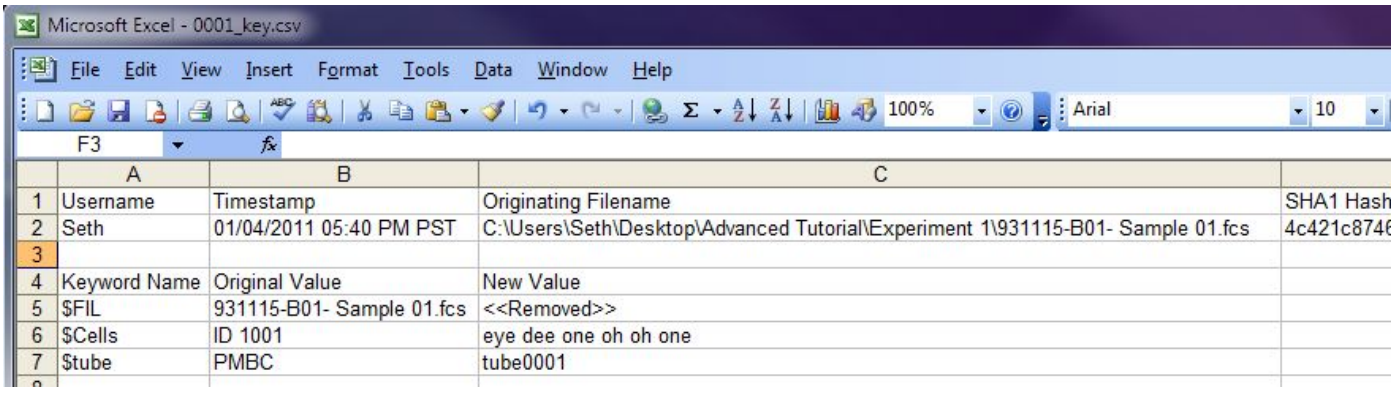


The next step is to deal with the warning icons and disabled "Process Button." We must select output directories for our files. Click the FCS file Destination box to choose a location. Do the same for the CSV file destination. Each output FCS file will produce a corresponding CSV file which serves as a sort of key. This tabular text file contains a digital signature, and depending on the option denoted by the "Write only modified values to .csv files" checkbox, a complete or terse list of before and after keywords and values.

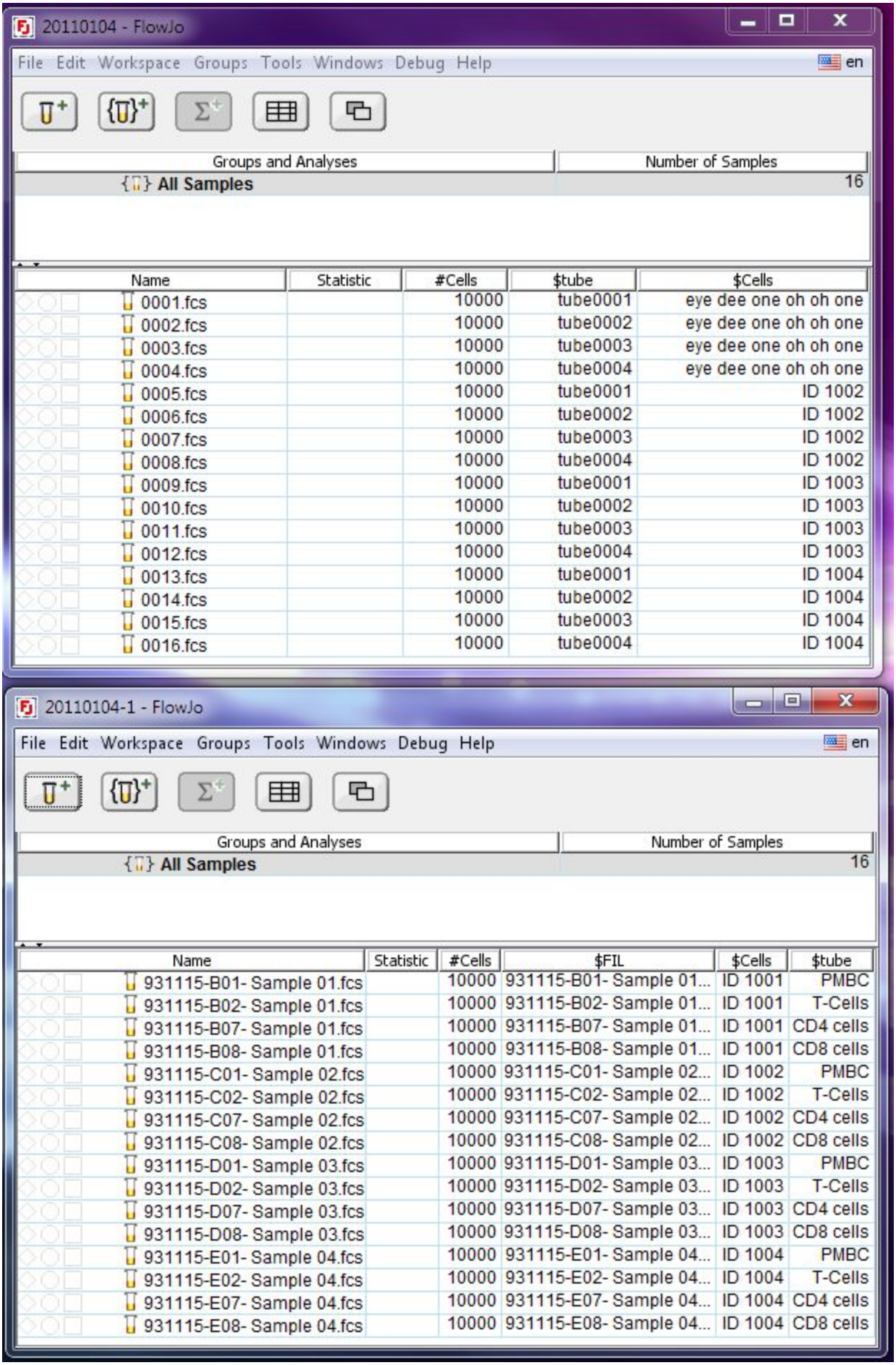
Once the directories are selected, the process button will become enabled and we are ready to produce deidentified files! The files output files have the following naming scheme:
0001.fcs, 0002.fcs, etc
0001_key.csv, 0002_key.csv, etc.
If you have any existing FCS files with the same name you will be prompted to resolve the conflict (skip/overwrite). CSV files will overwrite by default.



Sample CSV output:



Sample FCS output:



20110104 - FlowJo

File Edit Workspace Groups Tools Windows Debug Help en

Groups and Analyses	Number of Samples
{} All Samples	16

Name	Statistic	#Cells	$tube	$Cells
0001.fcs		10000	tube0001	eye dee one oh oh one
0002.fcs		10000	tube0002	eye dee one oh oh one
0003.fcs		10000	tube0003	eye dee one oh oh one
0004.fcs		10000	tube0004	eye dee one oh oh one
0005.fcs		10000	tube0001	ID 1002
0006.fcs		10000	tube0002	ID 1002
0007.fcs		10000	tube0003	ID 1002
0008.fcs		10000	tube0004	ID 1002
0009.fcs		10000	tube0001	ID 1003
0010.fcs		10000	tube0002	ID 1003
0011.fcs		10000	tube0003	ID 1003
0012.fcs		10000	tube0004	ID 1003
0013.fcs		10000	tube0001	ID 1004
0014.fcs		10000	tube0002	ID 1004
0015.fcs		10000	tube0003	ID 1004
0016.fcs		10000	tube0004	ID 1004

20110104-1 - FlowJo

File Edit Workspace Groups Tools Windows Debug Help en

Groups and Analyses	Number of Samples
{} All Samples	16

Name	Statistic	#Cells	$FIL	$Cells	$tube
931115-B01- Sample 01.fcs		10000	931115-B01- Sample 01...	ID 1001	PMBC
931115-B02- Sample 01.fcs		10000	931115-B02- Sample 01...	ID 1001	T-Cells
931115-B07- Sample 01.fcs		10000	931115-B07- Sample 01...	ID 1001	CD4 cells
931115-B08- Sample 01.fcs		10000	931115-B08- Sample 01...	ID 1001	CD8 cells
931115-C01- Sample 02.fcs		10000	931115-C01- Sample 02...	ID 1002	PMBC
931115-C02- Sample 02.fcs		10000	931115-C02- Sample 02...	ID 1002	T-Cells
931115-C07- Sample 02.fcs		10000	931115-C07- Sample 02...	ID 1002	CD4 cells
931115-C08- Sample 02.fcs		10000	931115-C08- Sample 02...	ID 1002	CD8 cells
931115-D01- Sample 03.fcs		10000	931115-D01- Sample 03...	ID 1003	PMBC
931115-D02- Sample 03.fcs		10000	931115-D02- Sample 03...	ID 1003	T-Cells
931115-D07- Sample 03.fcs		10000	931115-D07- Sample 03...	ID 1003	CD4 cells
931115-D08- Sample 03.fcs		10000	931115-D08- Sample 03...	ID 1003	CD8 cells
931115-E01- Sample 04.fcs		10000	931115-E01- Sample 04...	ID 1004	PMBC
931115-E02- Sample 04.fcs		10000	931115-E02- Sample 04...	ID 1004	T-Cells
931115-E07- Sample 04.fcs		10000	931115-E07- Sample 04...	ID 1004	CD4 cells
931115-E08- Sample 04.fcs		10000	931115-E08- Sample 04...	ID 1004	CD8 cells