

Auto Multiple Choice - Bug # 22: messy sheet papers

Status:	Closed	Priority:	Normal
Author:	red sea	Category:	
Created:	03/21/2012	Assignee:	
Updated:	04/01/2012	Due date:	
Description:	I'm on the order of the papers before scanned, sorted and saved in PDF file(arranged). but become messy after i import them into AMC (data capture).		

History

03/21/2012 07:02 pm - red sea

!http://www.freeimagehosting.net/t/89tlw.jpg!

03/21/2012 07:09 pm - red sea

sorry..

i mean the sheet id 1 should be given for the first sheet paper in pdf file, and the sheet i 2 should be given for 2nd sheet paper in pdf file

03/21/2012 07:10 pm - red sea

and o for the correct teacher sheet.

03/21/2012 09:40 pm - Alexis Bienvenüe

In the case answer sheets were photocopied, the sheet ID created by AMC for each page (the one after the colon) cannot be reliably connected with the page number from the PDF scan. This is because of parallelized processing. Moreover, at the data capture stage, AMC does not know which is the teacher sheet, so it cannot tag it with sheet ID 0.

AMC is not designed to take care of the order the scans were given. The right way to identify the scans in AMC is to use the name field.

03/22/2012 12:12 am - red sea

ok, Thank you for your generosity and your patience.

03/23/2012 04:26 pm - Alexis Bienvenüe

Since revision r956, when scan files are split (for example when you selected a multi-page PDF scan file), and when you don't use parallelization (set "number of processes" to 1 in Edit/Preferences/Main/System), the sheet IDs should be in the same order than the pages in the original PDF. However, this will slow down data capture on multi-processors (or multi-cores) computers...

03/24/2012 07:05 pm - red sea

thanks very much.

i don't understand the last thing?

i have i7 , what i can do to get the fast data capure ("number of processes").

best regards

03/25/2012 11:51 am - Alexis Bienvenüe

> what i can do to get the fast data capure ("number of processes").

It is enabled by default: default value for "number of processes" is 0, and that means that AMC will use all the cores in parallel during automatic data capture.

03/26/2012 11:37 pm - Alexis Bienvenüe

- Status changed from New to Resolved

From revision r963, an option has been added to pre-allocate sheets IDs (available in the window where you select scan files). With this option set, when you run automatic association from some scans, sheets IDs are allocated before data capture to all pages in the scans *in the right order*, starting from the maximum of all sheets IDs already used plus one.

* If all scans are from the same photocopied page and are properly recognized, all goes well.

* If some pages are not recognized, the corresponding pre-allocated sheet ID won't be used.

* If pages are from photocopies of different subjects, the sheet IDs will often jump when you look at the sheets IDs for _one particular page_.

This option is compatible with parallelization, so this solution is probably better than the one I suggested in note number 6 above.

04/01/2012 09:34 pm - Alexis Bienvenüe

- *Status changed from Resolved to Closed*