

Big Data Import Documentation

V1.4

By Eva Bascompte

Configuration

In the External Modules configuration there are 2 options:

Configure Module: Big Data Import

Project Settings	Value
List of emails, separated by commas to send the notification to once the import has finished:	<input type="text" value="myemail@vumc.org"/>
Sender email (by default noreply@vumc.org):	<input type="text" value="noreply@vumc.org"/>

- List of emails: an option to set up a list of emails. These emails will be notified when an import has finished or fails.
- Sender: the email of the sender of these notifications. If none is added by default it will be noreply@vumc.org.

Import files

This tool helps import one or more big CSV files without the need to split them.

To begin Click on the Big Data Import link under the external module section on the left-hand toolbar.

Note: this page does not refresh itself. To see the status of the import you will need to refresh the browser window.

Big Data Import

This tool helps import one or more big CSV files without the need to split them.

Note: this page does not refresh itself. To see the status of the import you will need to refresh the browser window.

Set CSV delimiter character:

, (comma) - default ▼

Select to check for existing records:



Allow blank values to overwrite existing values:



Set format for date and datetime values:

DD/MM/YYYY or YYYY-MM-DD ▼

Select a CSV file to import:

Choose File

No file chosen

Submit

Note: Once you hit start you can't upload more files until the process has finished.

🔍 Checked files:

None

🕒 Pending files:

None

Import Options:

Set CSV delimiter character: you can select any of the same delimiters REDCap exports with.

- , (comma)
- Tab
- ; (semi-colon)
- | (pipe)
- ^ (caret)

Select to check for existing records: When selected the module will check if the file contains existing records in the project. You can only check one file at a time.

Note: if not selected and records in the file are already on the project the data on the project will be overwritten.

Note: See Checked Files option for more information.

Allow blank values to overwrite existing values: this allows blank values to overwrite existing values.

Set format for date and datetime values: you can select two different formats for dates.





- DD/MM/YYYY or YYYY-MM-DD
- MM/DD/YYYY or YYYY-MM-DD

Select a CVS file to import: Only CVS files are accepted. Multiple import files may be queued, but must be added to the import queue one at a time. Once a file has been submitted the green header will show letting the user know a file has been uploaded.

Your file has been uploaded.
If you have set an email, a message will be sent once the import is ready, if not, refresh this page.

Pending Files: A numbered list of files will display with the option to delete them.



 Pending files:


1.  TestClinicalTrialsgo_DATA_2019-06-13_0839.csv 
2.  TestClinicalTrialsgo_DATA_2019-06-13_0839.csv 

*Note: files actively importing can't be deleted, you will have to use the **Cancel Current Import** button.*



Checked Files: Files with check for existing records option selected will appear here.

 Checked files:

1.  Harmonist4Regions_DATA_2019-06-18_1347.csv 

Once the file has been checked, if there are no duplicates it will go to *Pending*. If there are duplicates a new button will appear to continue with the import. In case the user wants to cancel the import the file can be deleted by clicking on 

 Checked files:

1.  Harmonist4Regions_DATA_2019-06-18_1347.csv 

Continue Import



Note: only one file at a time can be checked.

Note: in case of duplicates, the file won't be imported until the user does some action.

Start Process: Once a file is uploaded a Start Process button will appear. On clicking it, the import/check process will begin.

Start Process

Checked files:



1.  BigDataImportTest2_DATA_2019-08-06_1657.csv 

Pending files:

None

1. Failed Import

When an import fails a message will be sent to the emails designated in the configuration. The logging will also show the failed import. Below is an example of the page after a failed import:

Date/Time	Message	Records	Details
2019-06-12 14:10:55	Import process finished with errors 		Import # 2
2019-06-12 14:10:54	Import #2 did NOT complete successfully. Errors in batch 8 of 9 		Show Details

Clicking the “*Show Details*” will display why the import failed.


Details Import #1

```
{
  "errors": "The following fields were not found in the project as real data fields: app_services_pat_outcome__9, other_update, id_problem, app_services_pat_o",
  "warnings": [],
  "ids": [],
  "item_count": 0
}
```

[Close](#)

2. Successful Import



When an import is successful a message will be sent to the emails designated in the configuration. Below is an example of the page after a successful import:

Date/Time	Message	Records	Details
2019-08-26 08:38:49	Import process finished 		Import #4
2019-08-26 08:38:49	Import #4 completed successfully for batch 9 of 9	9	Import #4 Batch 9 of 9

Note: Emails will not be sent if there are no email addresses in the configuration.

3. Checking for existing records



When an import with the existing records option selected finds duplicates, a message will be sent to the emails designated in the configuration. Below is an example of the page after finding existing records:

Date/Time	Message	Records	Details
2019-07-11 08:16:26	Checking process finished with issues 		Import #2
2019-07-11 08:16:26	There are existing records in the project that match the excel file 	2	

Note: Emails will not be sent if there are no email addresses in the configuration.

4. Checking with issues

When an import with the existing records option selected finds blank records, a message will be sent to the emails designated in the configuration. Below is an example of the page after finding existing records:

Date/Time	Message	Records	Details
2019-07-11 09:09:10	Checking process finished with issues 		Import #2
2019-07-11 09:09:10	There are blank records in the file! 	Line: 7	

Note: The import will be cancelled and the file deleted.

Note: Emails will not be sent if there are no email addresses in the configuration.

5. Warnings

When an import has warnings but no errors, the import will continue but display those warnings at the end.

Date/Time	Message	Records	Details
2019-08-07 08:21:56	Import process finished ✓		Import #8
2019-08-07 08:21:56	Import #8 finished with warnings ⓘ		Show Details

Clicking the “*Show Details*” will show warning details.

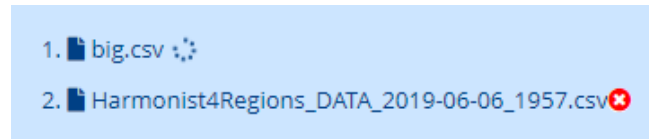
Details Import #8

Array(calc) Calculated fields cannot be imported because they will be calculated automatically during the import process.
All calculated fields will simply be ignored during the import.

Close

Cancel Current Import

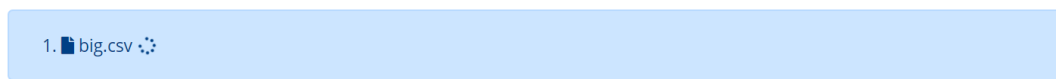
If an import has already started a spinning icon will appear, this icon lets the user know the file is currently being imported.



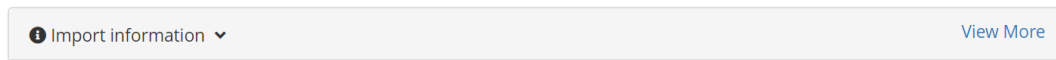
Once a file has begun importing, it cannot be deleted. It can only be canceled.

To cancel the current import, use the “Cancel Current Import” button next to *Recent Log Entries*:

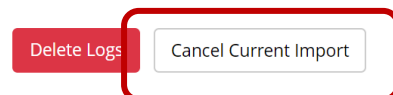
🕒 Pending files:



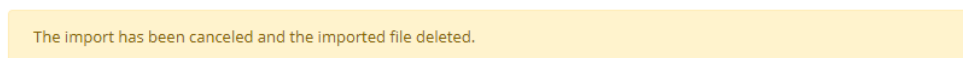
🔍 Checked files:



Recent Log Entries



When the current import is cancelled a new row in the logs will display, the new row displays which import was cancelled. You will also see a new yellow banner stating the import file was deleted.



Big Data Import

Tool to help import one or more big CSV files without the need to split them.

Date/Time	Message	Records	Details
2019-07-11 09:12:34	Import #4 cancelled 🚫		Import #4
2019-07-11 09:12:34	Import #4 completed successfully for batch 195 of 1350	6	Import #4

Note: Any canceled file will be deleted from pending files list.

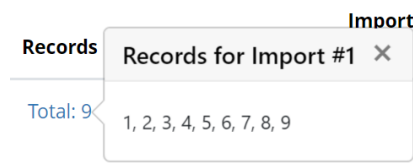
Note: After a file is canceled the next pending file will automatically start uploading. If you want to delete all of them, we recommend manually deleting them using the red X and then canceling the current import.

Import Information

This table will show data from the last 5 imports.

Import information ▾							View More
Date/Time	File	Uploaded By	Records	Status	Checked	Import #	
2019-07-11 09:14:18	big.csv	site_admin	Total: 9	✓	No	5	
2019-07-11 09:12:34	big.csv	site_admin	Total: 9	✗	No	4	
2019-07-11 09:11:04	fail.csv	site_admin	Total: 5	✗	No	3	
2019-07-11 08:16:17	Harmonist4Regions_DATA_2019-06-18_1347.csv	site_admin	Total: 9	✓	Yes	1	

On clicking on the records **total** a new window will display with the records imported.



When clicking on View More, it will redirect to another page where it will display all imports.

Big Data Import test 2

[< Back to Big Data Import](#)

Import information

View and search all current and past imports.

Excel Print

Search:

Status: Select All ▾

Show 50 ▾ entries

Previous 1 Next

Date/Time	File	Uploaded By	Records	Status	Checked	Import #
2019-07-11 09:14:18	big.csv	site_admin	Total: 9	✓	No	5
2019-07-11 09:12:34	big.csv	site_admin	Total: 9	✗	No	4
2019-07-11 09:11:04	fail.csv	site_admin	Total: 5	✗	No	3
2019-07-11 08:16:17	Harmonist4Regions_DATA_2019-06-18_1347.csv	site_admin	Total: 9	✓	Yes	1

Showing 1 to 4 of 4 entries

Previous 1 Next

The user will be able to print or export as an excel the table as well as to filter by status or text content.

Logs

In REDCap's *Logging* a new entry will appear every time a file is **submitted**, **deleted** or and import has been **cancelled** or **continued** import after checking.

Here are examples of how these display:

06/12/2019 8:11pm	site_admin	File submitted via <i>Big Data</i> <i>Import</i> external module	user = site_admin File = 'Harmonist4Regions_DATA_2019-06-06_1957.csv'
06/12/2019 8:04pm	site_admin	<i>Big Data Import</i> process cancelled	user = site_admin File = 'big.csv'
06/11/2019 4:08pm	site_admin	<i>Big Data Import</i> process cancelled	user = site_admin File = 'big.csv'
07/11/2019 2:58pm	site_admin	Continue file import after checking for existing records via <i>Big Data Import</i> EM	user = site_admin File = 'duplicate.csv' Import = 2

Delete Logs

If all data is erased, the settings and all the logs will be deleted from the import tool.

You can also delete the Logs by clicking on *Delete Logs* button

Delete Logs

When clicking on it an alert will display before deleting them.

