

The Ark Documentation.

How to use the subject custom field uploader to UPDATE custom field data.

Instruction Path a (with perfect data and no validation)

Log in and select our study.

Centre for
Genetic Epidemiology
& Biostatistics
THE UNIVERSITY OF WESTERN AUSTRALIA

THE ARK
OPEN SOURCE SOFTWARE
FOR MEDICAL RESEARCH

Welcome [arksuperuser@ark.org.au](#) | Logout

Study Reporting Global Search Admin

Study Details

| | | | |
|---|---|--------------------------------------|--|
| ID: | <input type="text"/> | Study Name: | <input type="text"/> |
| Status: | <input type="button" value="Choose One"/> | Date of Application: (dd/mm/yyyy) | <input type="text"/>  |
| Principal Contact: | <input type="text"/> | | |
| <input type="button" value="Search"/> <input type="button" value="Reset"/> <input type="button" value="New"/> | | | |

| ID: | Study Name: | Principal Contact: | Date of Application: (dd/mm/yyyy) |
|-----|---------------|--------------------|--------------------------------------|
| 587 | Test Study #1 | 장난장난장난장난 | |
| 588 | Test Study #2 | | |
| 589 | Preeclampsia | | |
| 590 | shortertable | | |

<< < 1 > >>

Once we have chosen our study we will see the following Modules and Functions Select Datasets



Welcome [arksuperuser@ark.org.au](#) | [Logout](#)

Study: Test Study #1

Study Subject Datasets LIMS Reporting Work Tracking Disease Global Search Admin

Study Details Study Components Manage Users Subject Custom Fields Subject Custom Field Upload Subject Upload

| | | | | | |
|--|-----------------|--|------------------------------|---|----------|
| Study: | | Test Study #1 | | Summary of study statistics: Total Subjects: 10 | |
| ID: | 587 | Study Name: | Test Study #1 * | Status: | Active * |
| Description: | 장난 | Estimated Year of Completion: | | | |
| Date of Application: | (dd/mm/yyyy) | Co-Investigator: | 안녕 안녕 | | |
| Chief Investigator: | Trav Arkson장난 * | Contact Phone: | 장난장난 | | |
| Contact Person: | 장난장난장난장난 | Logo: | Choose File no file selected | | |
| Application Modules: * | | Max 100px in height, & less than 100Kb in size | | | |
| Available | | Selected | | | |
| <input type="checkbox"/> Disease | | <input type="checkbox"/> Study <input type="checkbox"/> Subject <input type="checkbox"/> Datasets <input type="checkbox"/> Genotypic <input type="checkbox"/> LIMS | | | |
| Auto-generate Subject UID: <input type="checkbox"/> | | Auto-consent subjects: <input type="checkbox"/> | | | |
| Note: Disabled if Subjects exist | | | | | |
| Subject IID Pattern: <input type="button" value="Choose One"/> <input type="button" value="Choose One"/> | | Start Subject counter at: <input type="text"/> | | | |

We will now see the following functions (tabs), and our available data dictionary (listing our fields). To look at the fields in detail click on one of the Field Names (eg; Drug)



Welcome [arksuperuser@ark.org.au](#) | [Logout](#)

Study: Test Study #1

Study Subject Datasets LIMS Reporting Work Tracking Disease Global Search Admin

Data Dictionary Data Dictionary Upload Data Set Definition Field Data Upload

| | | | |
|---|---|----------------|----------------------|
| ID: | <input type="text"/> | Field Name: | <input type="text"/> |
| Field Type: | <input type="button" value="Choose One"/> | Description: | <input type="text"/> |
| Units: | | | |
| Minimum Value: | <input type="text"/> | | |
| | | Maximum Value: | <input type="text"/> |
| <input type="button" value="Search"/> <input type="button" value="Reset"/> <input type="button" value="New"/> | | | |

| ID: | Field Name: | Field Type: | Description: | Question/Label For Field: | Units: | Min: | Max: |
|-----|----------------|-------------|-----------------------|---------------------------------|--------|------|------|
| 326 | Drug | CHARACTER | 악물려 | 복용 약물 | | | |
| 322 | fdfdf | CHARACTER | Systolic | whats systolic | | | |
| 325 | Height | NUMBER | height in centimeters | What is your height? | cm | 150 | 190 |
| 319 | HOW_U_LIKE_ARK | CHARACTER | like the ark? | How do you like the ark so far? | | | |
| 324 | mix나 | CHARACTER | 난난 | mix나 | | | |
| 320 | SysBP | CHARACTER | Systolic | whats systolic | | | |
| 323 | 난 | CHARACTER | one helo | helo koreand | | | |
| 321 | 장난장난장난장난 | CHARACTER | Systolic | whats systolic | | | |

<< < > >>



You now get to look at the Type of values accepted plus and validations such as minimum and maximums.



Welcome [arksuperuser@ark.org.au](#) | Logout

Study: Test Study #1

| Study | Subject | Datasets | LIMS | Reporting | Work Tracking | Disease | Global Search | Admin | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|--|--------------------------------------|-----------------------------------|-----------|---------------|---------|---------------|-------|-----|-----|-------------|------|-------------|-----------|--------------|-----|-------------------------------|--|--|--|---------------------------|-------|--|--|------------------------------|--|--|--|--------|--|--|--|----------------|--|--|--|----------------|--|--|--|--|--|--------------------------------------|--|--|--|----------------|--|--|--|----------------|--|---|--|--|--|
| Data Dictionary | Data Dictionary Upload | Data Set Definition | Field Data Upload | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <table border="1"> <tr> <td>ID:</td> <td>326</td> <td>Field Name:</td> <td>Drug</td> </tr> <tr> <td>Field Type:</td> <td>CHARACTER</td> <td>Description:</td> <td>약물력</td> </tr> <tr> <td colspan="2">Note: Disabled if data exists</td> <td colspan="2"></td> </tr> <tr> <td>Question/Label For Field:</td> <td colspan="3">복용 약물</td> </tr> <tr> <td colspan="2">Note: Maximum 255 characters</td> <td colspan="2"></td> </tr> <tr> <td>Units:</td> <td colspan="3"></td> </tr> <tr> <td>Minimum Value:</td> <td colspan="3"></td> </tr> <tr> <td>Missing Value:</td> <td colspan="3"></td> </tr> <tr> <td colspan="2">Note: Indicates field data values that are missing</td> <td>Encoded Values: e.g.: 0=Yes;1=No;</td> <td></td> </tr> <tr> <td colspan="2"></td> <td>Maximum Value:</td> <td></td> </tr> <tr> <td colspan="2"></td> <td>Default Value:</td> <td></td> </tr> <tr> <td colspan="4" style="text-align: right;"> <input type="button" value="Save"/> <input type="button" value="Cancel"/> <input type="button" value="Delete"/> </td> </tr> </table> | | | | | | | | | ID: | 326 | Field Name: | Drug | Field Type: | CHARACTER | Description: | 약물력 | Note: Disabled if data exists | | | | Question/Label For Field: | 복용 약물 | | | Note: Maximum 255 characters | | | | Units: | | | | Minimum Value: | | | | Missing Value: | | | | Note: Indicates field data values that are missing | | Encoded Values: e.g.: 0=Yes;1=No; | | | | Maximum Value: | | | | Default Value: | | <input type="button" value="Save"/> <input type="button" value="Cancel"/> <input type="button" value="Delete"/> | | | |
| ID: | 326 | Field Name: | Drug | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Field Type: | CHARACTER | Description: | 약물력 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Note: Disabled if data exists | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Question/Label For Field: | 복용 약물 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Note: Maximum 255 characters | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Units: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Minimum Value: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Missing Value: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Note: Indicates field data values that are missing | | Encoded Values: e.g.: 0=Yes;1=No; | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | Maximum Value: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | Default Value: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <input type="button" value="Save"/> <input type="button" value="Cancel"/> <input type="button" value="Delete"/> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

To look at, or to create and Dataset definitions (you could also view a dataset definition as a questionnaire or grouping of relevant (longitudinal) questions)



Welcome [arksuperuser@ark.org.au](#) | Logout

Study: Test Study #1

| Study | Subject | Datasets | LIMS | Reporting | Work Tracking | Disease | Global Search | Admin | | | | | | | | | | | | |
|---|--|-------------------------------------|-----------------------------------|-----------|---------------|---------|---------------|-------|-----|----------------|--------------|------------|---|---------------|--|---|-----------|--|--|--|
| Data Dictionary | Data Dictionary Upload | Data Set Definition | Field Data Upload | | | | | | | | | | | | | | | | | |
| <p>You may view or edit one of the existing data set definitions or you may create a new data set.</p> <p>New</p> <table border="1"> <thead> <tr> <th>ID:</th> <th>Data Set Name:</th> <th>Description:</th> <th>Published:</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Small dataset</td> <td></td> <td>✓</td> </tr> <tr> <td colspan="4"><< <1> >></td> </tr> </tbody> </table> | | | | | | | | | ID: | Data Set Name: | Description: | Published: | 1 | Small dataset | | ✓ | << <1> >> | | | |
| ID: | Data Set Name: | Description: | Published: | | | | | | | | | | | | | | | | | |
| 1 | Small dataset | | ✓ | | | | | | | | | | | | | | | | | |
| << <1> >> | | | | | | | | | | | | | | | | | | | | |

To look at your Dataset Definition, click on the Data Set Name.

Now we are looking at the dataset, where we can add or remove fields IF there is not existing datasets using this dataset definition. It also helps us know what data to put in our dataset upload.

The screenshot shows the ARK software interface. At the top left is the logo for the Centre for Genetic Epidemiology & Biostatistics, University of Western Australia. At the top right is the ARK logo: "THE ARK" in red with a stylized orange globe graphic, and "OPEN SOURCE SOFTWARE FOR MEDICAL RESEARCH" below it. The top navigation bar includes links for Study, Subject, Datasets, LIMS, Reporting, Work Tracking, Disease, Global Search, Admin, Data Dictionary, Data Dictionary Upload, Data Set Definition (which is currently selected), and Field Data Upload. The main content area is titled "Study: Test Study #1". It shows a form for defining a new dataset:

| Data Set Name: | Small dataset * | Description: | <input type="text"/> | | |
|---|--|---|----------------------|-------------------------------------|-------------------------------------|
| Custom Fields: | Published: <input checked="" type="checkbox"/> | | | | |
| Available | | Selected | | | |
| <input type="checkbox"/> HOW_U_LIKE_ARK SysBP 장난장난장난장난 fdfdf 난 | | <input type="checkbox"/> Height Drug | | | |
| File of Custom Fields: Note: One field per line in file. Will replace selected list <input type="button" value="Choose File"/> no file selected <input type="button" value="Upload Field List"/> | | | | | |
| <input type="button" value="Save"/> <input type="button" value="Cancel"/> <input type="button" value="Delete"/> | | | | | |
| ID: | Field Name: | Question/Label For Field: | Sequence: | Required: | Allow multiple selections: |
| 349 | Height | What is your height? | 1 | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> |
| 350 | Drug | 복용 약물 | 2 | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> |
| << < 1 > >> | | | | | |

We now understand what data we can upload for this dataset. Click on Field Data Upload.

The following page will appear. You will see the following sections from top to bottom;

1. The files upload section. This is where you can begin the steps to upload subjects.
2. The Labels and Template download section. This will be described further below.
3. The Result list section. This section lists all previous uploads (if any...there are none yet in this example, but will be after we finish an upload).

 THE UNIVERSITY OF WESTERN AUSTRALIA

Welcome [arksuperuser@ark.org.au](#) | Logout

Study: Test Study #1

S

Study Subject Datasets LIMS Reporting Work Tracking Disease Global Search Admin

Data Dictionary Data Dictionary Upload Data Set Definition Field Data Upload

Step 1/5: Select data file to upload

Select the file containing data, the file type and the specified delimiter, click Next to continue.

Filename: no file selected *
Data Set: *
Comment:
Delimiter Type: *

Note: When uploading an Excel file (.xls), delimiter type is ignored. However, we recommend converting (.xls to a CSV file) and examining the data prior to uploading.
The variability of xl version means we cannot be certain of the quality of data the XL file will represent

Note: Click "Download Template" for an example file. Header columns are required.

ID: Filename: File Format: UserID: Start Time: Finish Time: Status:
[**<< < > >>**](#)

←1

←2

←3

The next step is to upload your data. You have two choices for your starting point;

1. You can select Download Template if you are unsure about the format. The template lists which fields are mandatory, what type/format of data is expected and valid values for fields that are limited in that way (see below).

| A | B | C | D | E |
|--|--|---|--|---|
| SUBJECTUID DESCRIPTION: The unique identifier assigned for this subject. This must match the subjectUID in the database NOTE: Re | RECORD_DATE_TIME The date this record was taken DD/MM/YYYY format | YOUR_FIRST_CUSTOM_FIELD_NAME_HERE Value for first custom field | YOUR_SECOND_CUSTOM_FIELD_NAME_HERE Value for second field | |

As you can see, this template is more of a guide than a generated file. You have to fill out the field names in the headers eg;

| 1 | SUBJECTUID,RECORD_DATE_TIME,Drug,Height |
|---|---|
| 2 | 12345,1/01/01,페니실린,153 |

2. Alternatively, you can access an example from the usefulTools/csvTemplatesForUploaders directory of our code repository. If you have not been given a local version of our repository you can access it from

<https://code.google.com/p/ark-informatics/source/browse/#svn%2FARK%2Ftrunk%2FusefulTools%2FcsvDataForUploaders>

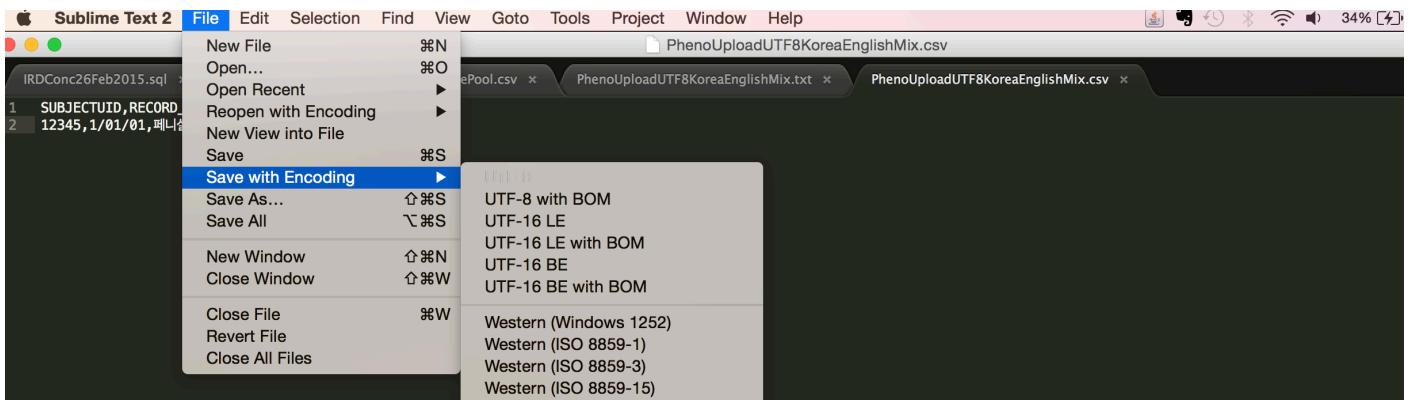
For the example below I utilized (updated) the PhenoUploadUTF8KoreaEnglishMixI.csv template file. You could also use the file you originally used to insert the data, but I strongly advise that you ONLY use the data you intend to update in order to avoid losing the current state of your data.

Important Note For Internationalisation (i18n):

This guide is created assuming you have taken all the steps necessary to ensure your Ark instance is completely setup for UTF-8 (contact ark-gohad@uwa.edu.au if you need help with this).

Always save your data as UTF-8. Many text editors such as Sublime Edit will do this to your CSV files. But the biggest thing is to make sure previous applications such as Excel DO NOT destroy your previous work. If possible always try to simply start off your document with UTF-8 as your first encoding as this is an international standard to ensure that you can have any international character set.

If you must convert character sets, here is an example of how I converted a UTF-16 csv containing Korean and English to UTF-8 for use in The Ark.



Once you have taken the steps listed above to ensure you have the fields you wish to upload, save the file, then you simply;

1. Select the filename of that file in the “The files upload section”.
2. Leave the delimiter type as COMMA, if appropriate.
3. Add a comment describing this dataset you are uploading. This can be important for longitudinal data, as it helps to clearly define what your are looking at later if you have many datasets/questionnaires over time and want to quickly recognize one.
4. Click the Next button once.



Welcome [arksuperuser@ark.org.au](#) | Logout

Study: Test Study #1

Study Subject Datasets LIMS Reporting Work Tracking Disease Global Search Admin

Data Dictionary Data Dictionary Upload Data Set Definition Field Data Upload

Step 1 /5: Select data file to upload

Select the file containing data, the file type and the specified delimiter, click Next to continue.

Filename: no file selected *
Data Set: *
Comment:
Delimiter Type: *

Note: When uploading an Excel file (.xls), delimiter type is ignored. However, we recommend converting (.xls to a CSV file) and examining the data prior to uploading.
The variability of .xl version means we cannot be certain of the quality of data the XL file will represent.

Note: Click "Download Template" for an example file. Header columns are required.

| | | | | | | | |
|--------|-----------|--------------|----------|-------------|--------------|--|---------|
| ID: | Filename: | File Format: | User ID: | Start Time: | Finish Time: | <input type="button" value="Download Template"/> | Status: |
| <<< >> | | | | | | | |

This will take you to the step 2. This page will let you know if the file structure is correct for the upload type. If an error occurs you can click cancel and correct the mistake.

The screenshot shows the THE ARK software interface. At the top left is the Centre for Genetic Epidemiology & Biostatistics logo from The University of Western Australia. At the top right is the THE ARK logo with the tagline "OPEN SOURCE SOFTWARE FOR MEDICAL RESEARCH". Below the header, a navigation bar includes links for Study, Subject, Datasets (which is highlighted in grey), LIMS, Reporting, Work Tracking, Disease, Global Search, and Admin. Under the Datasets menu, sub-links are shown: Data Dictionary, Data Dictionary Upload, Data Set Definition (which is highlighted in grey), and Field Data Upload. The main content area is titled "Step 2/5: File Format Validation" and contains the message: "The file format has been validated. If there are no errors, click Next to continue." A section labeled "Validation Messages:" is present but empty. Below this is a "Data Preview: (First 20 rows of file)" table:

| ROW | SUBJECTUID | RECORD_DATE_TIME | Drug | Height |
|-----|------------|------------------|------|--------|
| 1 | 12345 | 1/01/01 | 페니실린 | 133 |

At the bottom of the preview table are three buttons: Next, Cancel, and Finish.

But on this occasion, we can see no validation messages at step 2, so I can assume everything is ok and click next at the bottom of the page (the next step will validate all of your **data** Vs what is expected, so can take some time on very large files. Please be patient and don't click back, as it is busy doing work and will let you know when it is complete).

SIDENOTE : Ensuring Excel doesn't add extra commas and lines;

If you passed validation on Step one of the previous page disregard this information. If you are having an error message, this can often help you out.

Step 1: Select all (or at least those that may have once had data) columns after our last column of actual data

| | A | B | C | D | E | F |
|----|------------|---------------|----------------|---------|---|---|
| 1 | SUBJECTUID | CANCER_BREAST | CANCER_DX_DATE | 1/01/01 | | |
| 2 | 998877 | Unknown | | | | |
| 3 | | | | | | |
| 4 | | | | | | |
| 5 | | | | | | |
| 6 | | | | | | |
| 7 | | | | | | |
| 8 | | | | | | |
| 9 | | | | | | |
| 10 | | | | | | |
| 11 | | | | | | |

Step 2: Delete selected columns

| | A | B | C | D | E | F |
|----|------------|---------------|----------------|---------|---|---|
| 1 | SUBJECTUID | CANCER_BREAST | CANCER_DX_DATE | 1/01/01 | | |
| 2 | 998877 | Unknown | | | | |
| 3 | | | | | | |
| 4 | | | | | | |
| 5 | | | | | | |
| 6 | | | | | | |
| 7 | | | | | | |
| 8 | | | | | | |
| 9 | | | | | | |
| 10 | | | | | | |
| 11 | | | | | | |

Step 3 & 4; Repeat the process by selecting and deleting all rows after our last row of data.

A screenshot of Microsoft Excel showing a context menu open over a row of data. The menu options are:

- Cut
- Copy
- Paste
- Paste Special...
- Insert
- Delete

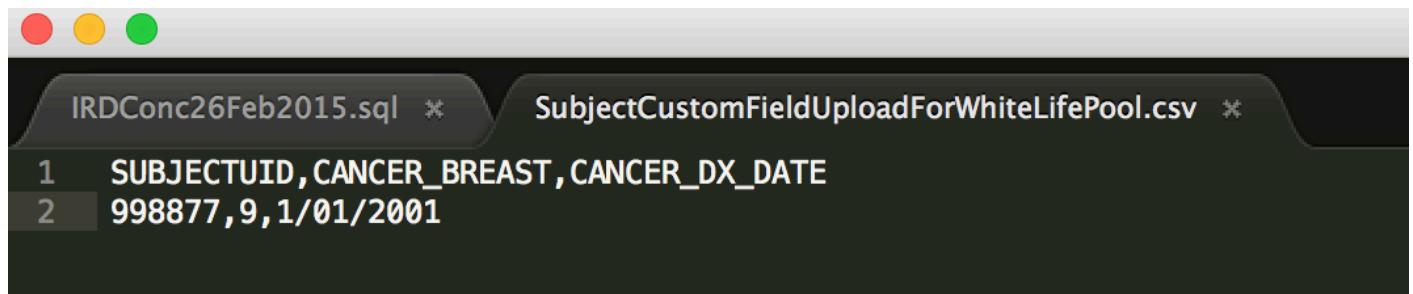
The menu also includes keyboard shortcuts: ⌘X, ⌘C, ⌘V, ⌘V, ⌘V, and ⌘V respectively. The background shows a spreadsheet with columns A through G and rows 1 through 11. Row 1 contains "SUBJECTUID", "CANCER_BREAST", and "CANCER_DX_DATE". Row 2 contains "998877", "Unknown", and "1/01/01". Rows 3 through 10 are empty. Row 11 contains the context menu options.

Step 5: Now save as CSV file.

A screenshot of the 'Save As' dialog in OS X. The dialog shows a list of files in the 'Favorites' sidebar and a main pane displaying a table of files. The table has columns for Name, Date Modified, Size, and Kind. The 'Format' dropdown is set to 'Comma Separated Values (.csv)'. Below the table, there is a 'Description' section with the text: 'Exports the data on the active sheet to a text file that uses commas to separate values in cells.' There are also links to 'Learn more about file formats' and 'Compatibility Report...'. At the bottom, there are buttons for 'Options...', 'Compatibility Report...', 'Cancel', and 'Save'.

| Name | Date Modified | Size | Kind |
|-------------------------|---------------------|-----------|--------------|
| SubjectCu...ifePool.csv | Today 9:05 am | 62 bytes | comm...val |
| SubjectCu...ifePool.xls | Today 9:05 am | 18 KB | Micros...ksl |
| SubjectUpl...White.csv | Yesterday 8:06 am | 636 bytes | comm...val |
| ABC(pad t...DATES.csv | 22 Sep 2014 2:34 pm | 7 KB | comm...val |
| ABC(pad t...000096.csv | 22 Sep 2014 2:34 pm | 7 KB | comm...val |
| ABC(pad that) 17000.csv | 22 Sep 2014 2:34 pm | 2.4 MB | comm...val |
| ACoupleOfIDS.csv | 22 Sep 2014 2:34 pm | 13 bytes | comm...val |
| ACoupleOf...TUDY.csv | 22 Sep 2014 2:34 pm | 23 bytes | comm...val |
| arkfunctionsInCsv.csv | 22 Sep 2014 2:34 pm | 7 KB | comm...val |
| AutoGenLI...Upload.csv | 22 Sep 2014 2:34 pm | 229 bytes | comm...val |
| BioCOLLE...1fields.csv | 22 Sep 2014 2:34 pm | 2 KB | comm...val |
| BioSPECI...31fields.csv | 22 Sep 2014 2:34 pm | 2 KB | comm...val |
| BIOUpload.xls | 22 Sep 2014 2:34 pm | 6 KB | Micros...ksl |
| CustomFie...stomer.csv | 22 Sep 2014 2:34 pm | 49 bytes | comm...val |
| CustomFieldData.csv | 22 Sep 2014 2:34 pm | 75 bytes | comm...val |
| CustomFie...llValid.csv | 22 Sep 2014 2:34 pm | 403 KB | comm...val |

Step 6: Check your data in a plain text editor (eg; Notepad, sublime edit, vi, vim,TextEdit). It should now have no extra commas and no extra lines.



```
IRDConc26Feb2015.sql *
SubjectCustomFieldUploadForWhiteLifePool.csv *

1 SUBJECTUID,CANCER_BREAST,CANCER_DX_DATE
2 998877,9,1/01/2001
```

So now our upload data is as expected, we can try again.

We click cancel at the bottom of The Ark page and start the upload again.

We are now at step 3. This step shows if anything went wrong in step 2. The validation messages say that we have a value out of the expected range (ie; this dataset has a rule that height must be greater than 150, yet our listed height is 133). So I have to fix the data and try again.

Centre for
Genetic Epidemiology
& Biostatistics
THE UNIVERSITY OF WESTERN AUSTRALIA

THE ARK
OPEN SOURCE SOFTWARE
FOR MEDICAL RESEARCH

Welcome [arksuperuser@ark.org.au](#) | Logout

Study: Test Study #1

Study Subject Datasets LIMS Reporting Work Tracking Disease Global Search Admin

Data Dictionary Data Dictionary Upload Data Set Definition Field Data Upload

Step 3/5: Data Validation
The data in the file is now validated, correct any errors and try again, otherwise, click Next to continue.

Validation Messages:

ERROR: Subject 12345 has a value: 133 which is less than the minimum allowed value of 150.0

[Download Validation Messages](#)

Data Preview: (First 20 rows of file)

| ROW | SUBJECTUID | RECORD_DATE_TIME | Drug | Height |
|-----|------------|------------------|------|--------|
| 1 | 12345 | 1/01/01 | 페니실린 | 133 |

KEY: Indicates data in error Indicates data with a warning Indicates data to insert Indicates subject does not exist [Download File](#)

the-ark 1.2 Copyright © the-ark.org.au. All rights reserved.

[WICKET AJAX DEBUG](#)

Click cancel and go and fix your data, and then repeat steps one and two again.

If we have corrected all of our data error we will now see an empty box for the validation messages.

Welcome [arksuperuser@ark.org.au](#) | Logout

Study: Test Study #1

Study Subject Datasets LIMS Reporting Work Tracking Disease Global Search Admin

Data Dictionary Data Dictionary Upload Data Set Definition Field Data Upload

Step 3/5: Data Validation

The data in the file is now validated, correct any errors and try again, otherwise, click Next to continue.

Validation Messages:

Data Preview: (First 20 rows of file)

| ROW | SUBJECTUID | RECORD_DATE_TIME | Drug | Height |
|-----|------------|------------------|------|--------|
| 1 | 12345 | 1/01/01 | 페니실린 | 153 |

KEY: Indicates data in error Indicates data with a warning Indicates data to insert Indicates subject does not exist

[Next](#) [Cancel](#) [Finish](#)

If things look good, you can therefore click next if you are happy with the data you see.

(SIDE NOTE; It is worth noting the dataset uploader DOES NOT UPDATE data. Dataset data is LONGITUDINAL, and we assume that each time you upload data it is intended to be a new longitudinal set)

You have now confirmed you are accepting the data, click next one more time and a BATCH process will begin.



Welcome [arksuperuser@ark.org.au](#) | [Logout](#)

Study: Test Study #1



Step 4/5: Confirm Upload

Data will now be written to the database, click Next to continue, otherwise click Cancel.

[Next](#) [Cancel](#) [Finish](#)

You now have completed the final step and can click finish whenever you want...



Welcome [arksuperuser@ark.org.au](#) | [Logout](#)

Study: Test Study #1



Step 5/5: Data Upload Finished

The data has been successfully submitted, When the actual upload is finished it will have a status of "Successfully Completed".

[Next](#) [Cancel](#) [Finish](#)

[WICKET AJAX DEBUG](#)

the-ark 1.2 Copyright © the-ark.org.au. All rights reserved.

BUT you must know that the data has been submitted BUT it has not been completed until the result list page says that your upload is in a “Successfully Completed” status.

This is what a successful upload will look like on the result list section. You can also see what you two previous unsuccessful attempts looked like (the first time we failed validation of the file (headers or file type/format), the second time we failed validation of the data.



Welcome [arksuperuser@ark.org.au](#) | Logout

Study: Test Study #1

Study Subject Datasets LIMS Reporting Work Tracking Disease Global Search Admin

Data Dictionary Data Dictionary Upload Data Set Definition Field Data Upload

o Data upload of file: PhenoUploadUTF8KoreaEnglishMix.csv was uploaded successfully

Step 1/5: Select data file to upload
Select the file containing data, the file type and the specified delimiter, click Next to continue.

Filename: no file selected *

Data Set: *

Comment:

Delimiter Type: *

Note: When uploading an Excel file (.xls), delimiter type is ignored. However, we recommend converting (.xls to a CSV file) and examining the data prior to uploading.
The variability of xl version means we cannot be certain of the quality of data the XL file will represent.

Note: Click "Download Template" for an example file. Header columns are required.

| ID: | Filename: | File Format: | UserID: | Start Time: | Finish Time: | Download Template | Status: |
|-----|------------------------------------|--------------|-------------------------|--------------------------|--------------------------|--|---|
| 8 | PhenoUploadUTF8KoreaEnglishMix.csv | CSV | arksuperuser@ark.org.au | 2015-03-26 14:49:20.0 | 2015-03-26 14:49:20.0 | <input type="button" value="Download File"/> | <input type="button" value="Download Report"/> Successfully completed |
| 7 | PhenoUploadUTF8KoreaEnglishMix.csv | CSV | arksuperuser@ark.org.au | 2015-03-26 14:47:29.0 | | <input type="button" value="Download File"/> | Error while validating data |
| 6 | PhenoUploadUTF8KoreaEnglishMix.txt | TXT | arksuperuser@ark.org.au | 2015-03-26 14:44:39.0 | | <input type="button" value="Download File"/> | Error validation file |

<< < 1 > >>

As you can see, in the results list for our successful upload, we have the options of Download File (this will down the original file you used) or Download Report which will give you a summary of what was uploaded, and any assumptions (or less serious errors) that were made upon successful rows. See an example below;

The screenshot shows theARK software interface. At the top left is the Centre for Genetic Epidemiology & Biostatistics logo from The University of Western Australia. At the top right is THE ARK logo with the tagline "OPEN SOURCE SOFTWARE FOR MEDICAL RESEARCH". The main window title is "uploadReport538.txt". The content area displays a log message:

```
**:  
the  
Thi*****  
** Upload Report  
Sti Thu Mar 26 14:49:19 EST 2015  
Us*****  
Fi Study: Test Study #1  
Fi UserID: arksuperuser@ark.org.au  
Do Filename: PhenoUploadUTF8KoreaEnglishMix.csv  
To File Format: CSV  
In Delimiter Type: COMMA  
0 Total file size: 0.00 MB  
2 Processed 1 rows.  
2 Inserted 2 rows of data.  
Updated 0 rows of data.
```

Below the log, there's a dropdown menu set to "COMMA" with an asterisk. A note says "Examining the data prior to uploading." with a "Finish" button. A table lists uploaded files:

| Name | Type | Uploader | Date | Action | Status |
|--------------------------------------|------|-------------------------|-----------------------|-------------------------------|--|
| 7 PhenoUploadUTF8KoreaEnglishMix.csv | CSV | arksuperuser@ark.org.au | 14:49:20.0 2015-03-26 | Download File | Download Report Successfully completed |
| 6 PhenoUploadUTF8KoreaEnglishMix.txt | TXT | arksuperuser@ark.org.au | 14:47:29.0 2015-03-26 | Download File | Error while validating data |
| | | | 14:44:39.0 | Download File | Error validation file |

At the bottom, there are navigation links "<<< 1 >>>" and a copyright notice "the-ark 1.2 Copyright © the-ark.org.au. All rights reserved." To the right of the copyright notice is a "WICKET AJAX DEBUG" link.

As always, we recommend you test what you just did to ensure the results were as you expected.

To do this, we click on the Subject Module Tab near the top of the page.

We can use a search term to find out subjects we wish to test

Click on his SubjectUID to test put this subject in Context.

Your next step is to click “Clinical Data” tab to see if the data was uploaded as expected. Hopefully all the data you expected is there. As we can see, our new data set we just added is there (I labelled it Trying English and Korean utf8 so I could easily recognize it later).

Welcome [arksuperuser@ark.org.au](#) | [Logout](#)

Study: Test Study #1 Subject UID: 12345

Study Subject Datasets LIMS Reporting Work Tracking Disease Global Search Admin

Demographic Data Study-specific Demographic Data Contact Attachments Consent Correspondence Clinical Data Subject Biospecimen Biospecimen Search Pedigree

| ID: | Data Set: | Comment: | Status: | Record Date: | Reviewed Date: |
|-----|---------------|--------------------------------|--------------------|--------------|----------------|
| 1 | Small dataset | | In Progress | 03/12/2014 | |
| 2 | Small dataset | Trying English and Korean utf8 | Uploaded From File | 26/03/2015 | |

New Choose One Get Phenotypic data

the-ark 1.2 Copyright © the-ark.org.au. All rights reserved. WICKET AJAX DEBUG

Click on the Dataset that we wish too look at.

We can see that our data has been updated correctly. It has a mixture of many international characters that are available as a part of the Universal UTF-8 charset (which enables many languages such as English, Korean and many more).

The screenshot shows the ARK software interface. At the top, there are logos for the Centre for Genetic Epidemiology & Biostatistics (The University of Western Australia) and THE ARK OPEN SOURCE SOFTWARE FOR MEDICAL RESEARCH. The main header displays "Study: Test Study #1" and "Subject UID: 12345". A navigation bar on the left includes links for "Study", "Demographic", "New", and "Choices". The central area is a modal dialog titled "Subject Dataset Details". Inside the dialog, there are several input fields: "ID" (value: 2), "Record Date" (value: 26/03/2015), "Reviewed By" (dropdown menu: "Choose One"), "Comment" (text area: "Trying English and Korean utf8"), "Data Set" (dropdown menu: "Published datasets only" selected), "Status" (dropdown menu: "Uploaded From File" selected), and "Reviewed Date" (calendar field). Below these fields, there is a note: "What is your height? 153 cm" and "복용 약물 페니실린". At the bottom of the dialog, there is a warning message: "<<< 1 >>> Warning: You must click the "Save" button for changes to be kept. Moving away (including to the next page of fields) will lose any unsaved data." and three buttons: "Save", "Cancel", and "Delete". At the very bottom of the screen, a footer bar contains the text "the-ark 1.2 Copyright © the-ark.org.au. All rights reserved." and a link "WICKET-AJAX DEBUG".

We also intend to put out a general guide to i18n and how to make sure your character encryption is created and always kept in UTF-8 to ensure it can work anywhere in the future.