



The Ark

Quick-start Guide

Contents:

STUDY MODULE	3
Creating A Study	3
Managing Study Users.....	5
Uploading Subject Demographic Data	6
SUBJECT MODULE	8
Adding a New Subject.....	8
Searching for a Subject	10
Editing a Subject's Details	12
Adding, searching and editing Phone Numbers.....	13
Adding, Searching and Editing Addresses.....	15
Attachments	16
Subjects, Collections and Biospecimens.....	16
The Subject Biospecimen Summary Page.....	17
LIMS MODULE (Laboratory Information Management System)	18
Adding New Collection(s) or Biospecimen(s)	18
Locating and Managing Biospecimens.....	18
Allocating a Biospecimen to a Freezer Location.....	20
Biospecimen Transactions	21
Searching for a Biospecimen(s)	22
Batch Allocation.....	23
Unallocating Biospecimens.....	23
DATASET MODULE	24
The Data Dictionary	25
<i>The Data Dictionary Uploader</i>	28
Defining Datasets	30
Bulk Dataset Data Upload.....	32
Manually entering the data subject-by-subject.....	33

STUDY MODULE

In The Ark, Studies are the “lowest common denominator” of data entities. Every other data type in The Ark (i.e. Subjects, Biospecimens, Phenotypic) belong either directly or indirectly to a Study.

Creating A Study

The screenshot shows a web-based application window titled "Centre for Genetic Epidemiology & Biostatistics THE UNIVERSITY OF WESTERN AUSTRALIA". The URL in the address bar is "arksuperuser@ark.org.au" and the IP address is "192.168.33.10". The top right corner features the "THE ARK" logo with the tagline "OPEN SOURCE SOFTWARE FOR MEDICAL RESEARCH". The main menu includes "Study", "Global Search", and "Admin". The "Study" tab is selected. Below it, a sub-menu "Study Details" is open, showing a message: "There are no studies available." There are four input fields: "ID:" (empty), "Status:" (dropdown menu showing "Choose One"), "Study Name:" (empty), and "Principal Contact:" (empty). To the right of these fields are two more: "Date of Application:" (empty) and "Principal Contact:" (empty). At the bottom of the form are three buttons: "Search", "Reset", and "New". A footer note at the bottom left reads "the-ark 1.2 Copyright © the-ark.org.au. All rights reserved." and a link "WICKET-AJAX DEBUG" is visible at the bottom right.

To create a new Study, from the Study Selection tab, click on the New button underneath the search form fields.

On the New Study form, you should fill in all appropriate fields, however note that Study Name, Status, and Chief Investigator are required fields and need to be entered.

You should also specify which of the Application Modules are available for your study to use. If a module is not moved to the Selected side, then it will not be accessible for this study.

Auto-generation of Subject, Biospecimen or Biocollection UIDs is disabled if any Subject, Biospecimens or Biocollections respectively exist within the study, so this configuration should occur before any Subjects are enrolled in the study.

For the configuration of UID auto-generation, you need to specify 3 things;

- Prefix: This is a one to three-character prefix that will be applied to all UIDs,
- Token: The token is a single character that separates the prefix from the numerical part of the UID,
- Counter: The number of the subject/biospecimen/biocollection that the UID is associated with. You need to specify a length for the counter, and the entity number will be padded by zeros to make it that length. i.e. If you specify a length of 5, the second UID generated will have [Prefix][Token]00002, and the 150th will be [Prefix][Token]00150. UIDs can still be generated when the number of UIDs exceeds the length, and will be created without any padded zeros.

Managing Study Users

To add a user to a study, go to the Manage Users sub-tab within the Study tab, and click on the New button.

The screenshot shows a web-based application window titled 'Demo'. At the top left is the logo for 'Centre for Genetic Epidemiology & Biostatistics, THE UNIVERSITY OF WESTERN AUSTRALIA'. At the top right is the 'THE ARK' logo with the tagline 'OPEN SOURCE SOFTWARE FOR MEDICAL RESEARCH'. The main navigation bar includes links for 'Study', 'Subject', 'Datasets', 'LIMS', 'Work Tracking', 'Global Search', and 'Admin'. Under 'Admin', the 'Manage Users' sub-tab is selected. The page title is 'Study: Demo'. The 'Manage Users' form contains fields for 'Login User Name', 'Email', 'First Name', 'Last Name', and 'Confirm Password'. Below these fields is a password strength checker stating: 'A combination of: - at least 1 digit, - a lower case letter, - an upper case letter, - a special character (e.g. @#\$%^&*!) and - length between 6-20 characters.' A table below lists 'Module Name' (Study, Subject) and 'Role' (Study Read-Only user, Choose One). At the bottom are 'Save', 'Cancel', and 'Remove' buttons. The footer of the page includes the text 'the-ark 1.2 Copyright © the-ark.org.au. All rights reserved.' and a link 'WICKET AJAX DEBUG'.

To create the user, you need to complete all the text fields in the top half of the form.

- Login User Name: A valid email address for that user to use as a login.
- Email: A valid email address to be used to contact the user.
- First Name: The new users first name
- Last Name: The new users last name
- Password: When entering Password and Confirm Password, you should be aware of the rules for password creation. You need to comply with all of the following rules;
 - At least 1 digit,
 - At least 1 lower case letter,
 - At least 1 upper case letter,
 - At least 1 special character (e.g. @#\$%^&*!)
 - A password length between 6 and 20 characters.

If the user already exists in The Ark, but is associated with a different study, you will be prompted to add their existing user account to your study. Their password will remain the same, however the other details (Email, First Name, and Last Name) will be entered as per the new study.

In this view, you can also specify the access control levels for each of the available study modules. Each module can be set to have Read Only, Data Manager, or Administrator access which allow for different levels of read/write access, depending on what has been configured in the administrator section of The Ark.

Uploading Subject Demographic Data

There are often occasions where you need to enter data for many subjects at once, and you can achieve that through the Study Data Upload tab under Study.

The screenshot shows a web-based application window titled "Demo". At the top left is the logo for "Centre for Genetic Epidemiology & Biostatistics, THE UNIVERSITY OF WESTERN AUSTRALIA". At the top right is the "THE ARK OPEN SOURCE SOFTWARE FOR MEDICAL RESEARCH" logo. The main navigation bar includes links for Study, Subject, Datasets, LIMS, Work Tracking, Global Search, Admin, Study Details, Study Components, Manage Users, Custom Fields Category, Custom Fields, Subject Custom Field Upload, and Study Data Upload. The "Study Data Upload" tab is currently selected. A sub-menu for "Step 1/5: Select data file to upload" is open, containing instructions: "Select the file containing data, upload type and the specified delimiter, click Next to continue." It includes fields for "Filename:" (with a "Choose File" button and a note "no file selected"), "Upload Type:" (set to "Subject Demographic Data" with a required asterisk), and "Delimiter Type:" (set to "COMMA"). Below these are notes about Excel file compatibility and header requirements, along with "Next", "Cancel", and "Finish" buttons. At the bottom of the main area, there are download links for various template files: "Download Subject Template", "Download (Subject/Family)Custom Field Data Template", "Download Consent Field Template", "Download Pedigree Template", and "Download Subject Attachment Template". The status bar at the bottom of the screen displays "the-ark 1.2 Copyright © the-ark.org.au. All rights reserved." and "WICKET-AJAX DEBUG".

To upload data, it needs to be in the correct format which is shown in the templates available for download. Headers are required and must be identical to those in the template files. The template files also show whether a field is required or not.

When the file is ready to upload, select the file using the Choose File dialog. Then ensure that the correct upload type is selected, in this case “Subject Demographic Data”. Then you need to choose the delimiter being used within the file, this will usually be a COMMA but may be of type - COMMA, TAB, PIPE (|), COLON (:) or AT SYMBOL (@). Then you may click the Next button to commence the file validation process.

If the file format is validated correctly by the system then you will then be shown a sample of the file that you are uploading, as The Ark understands it, with the column headers and data for the first 10 subjects. If the file format validation fails the checks you should be notified at this point of why the failure occurred, and the only option will be to Cancel the upload. If the first 10 lines rendered on the screen do not appear correct, even if the format check completes successfully, then click Cancel. If the first 10 lines have not displayed correctly, you will most likely need to amend your upload file or change the parameters defined on the previous screen for the upload. If the file format validation process has passed successfully then click Next.

If the data in the file is validated correctly by the system then you will again be shown the first 10 subject from the file and you will be shown whether the entries are updated, inserted or if nothing happens. If the data validation step has completed successfully then click Next, otherwise click Cancel.

After clicking Next, you will be prompted to complete the process by clicking Next again. At this point you may still cancel out of the upload. If the upload is successful, you will see a message to that effect. Clicking Finish will then take you back to the main Study Data Upload screen.

You can also use this section to upload Custom Field Data, Consent Data, Pedigree information and Subject attachments which all have a similar approach, but are outside the scope of this document.

SUBJECT MODULE

Within The Ark, all data is associated with a subject. This includes all phenotypic data and biospecimen samples. It is not possible to load any phenotypic or biospecimen data without a valid subject already having been defined to the system.

Adding a New Subject

After logging into The Ark and placing a study in context, click on the Subject tab to bring up the Subject Search screen.

Subject UID:	Full Name:	Previous Last Names:	Date of Birth: (dd/mm/yyyy)	Vital Status:	Gender:	Subject Status:	Consent Status:	Other IDs:
DEM-00000001	Mason Neumann		17/11/1983	Alive	Male	Subject	Consented	
DEM-00000002	Dedan Crooks		02/02/1979	Alive	Male	Subject	Pending	
DEM-00000003	Elizabeth Dallas		20/09/1935	Alive	Female	Subject	Pending	
DEM-00000004	Hannah Fugit		10/08/1965	Alive	Male	Subject	Pending	
DEM-00000005	Piper Garratt		02/12/1941	Alive	Female	Subject	Consented	
DEM-00000006	Tyson Clarkson		06/01/1980	Unknown	Male	Subject	Pending	
DEM-00000007	Holly Irwin		13/01/1948	Alive	Unknown	Subject	Consented	

On the Subject Demographic Data Search, click on “New”.

Enter the following details (as a minimum) in the form:

- Subject unique identifier (Subject UID): Can be manually entered, or if the study is configured appropriately it will be auto-generated,
- Subject Status: Indicates whether a subject is actively enrolled in the study,
- Consent Status: Indicates whether consent has been given by the subject for the study. This field gets set to “Pending” by default.

All remaining fields are optional, and may be left blank.

Please be aware of the following fields’ meanings:

- Consent to active contact – Does the participant consent to contact for future studies?
- Consent to passive data gathering – Does the participant consent his/her medical info to be collected?
- Consent to use data – Does the participant consent for disclosure of the findings to other family members?

- Consent status – Did the participant complete the consent form at the time of examination or sample collection?
- Heard About Study From – Please indicate referral location (Optional).

Once you complete the form, click on “Save”. A message will appear in green at the top the form.

The screenshot shows a web-based application window titled "Demo". At the top left is the logo for "Centre for Genetic Epidemiology & Biostatistics, THE UNIVERSITY OF WESTERN AUSTRALIA". At the top right is the "THE ARK" logo with the tagline "OPEN SOURCE SOFTWARE FOR MEDICAL RESEARCH". The main header includes "Welcome arksuperuser@ark.org.au | Logout" and tabs for "Study", "Subject", "Datasets", "LIMS", "Work Tracking", "Global Search", and "Admin". Below this is a navigation bar with tabs for "Demographic Data", "Study-specific Demographic Data", "Contact", "Attachments", "Consent", "Correspondence", "Clinical Data", "Subject Biospecimen", "Biospecimen Search", "Pedigree", and "Affection". A message in green at the top of the main content area says: "o The Subject with Subject UID: DEM-00000001 has been created successfully and linked to the study in context: Demo." The main form contains fields for personal information like Title, First Name, Last Name, Date of Birth, Gender, Marital Status, Email, and "Heard About Study From". It also includes sections for "Consent Date", "Consent Type", "Consent to Passive Data Gathering", "Consent to Use Data", and "Consent Status". Buttons for "Other ID Source" and "Add New" are visible. At the bottom of the form is a "Consent History" section with a checkbox. The footer of the page includes the text "the-ark 1.2 Copyright © the-ark.org.au. All rights reserved." and a link "WICKET AJAX DEBUG".

Searching for a Subject

Subject Search screen has an area at the top where you may enter your search criteria. The results of your search will be displayed as a list in the bottom part of the screen. When you first enter the Search screen you will be shown the first page of subjects sorted by Subject Unique ID. You may modify the search criteria as follows:

The screenshot shows the 'The Ark' software interface for searching subjects. At the top, there's a header with the Centre for Genetic Epidemiology & Biostatistics logo, a 'Demo' title, and The Ark logo. Below the header is a navigation bar with tabs: Study, Subject, Datasets, LIMS, Work Tracking, Global Search, and Admin. The 'Subject' tab is selected. Underneath the tabs is a sub-navigation bar with tabs: Demographic Data, Study-specific Demographic Data, Contact, Attachments, Consent, Correspondence, Clinical Data, Subject Biospecimen, Biospecimen Search, Pedigree, and Affection. The 'Demographic Data' tab is selected. The main search form contains fields for Study (set to 'Demo'), Subject UID (text input), Middle Name (text input), Date of Birth (date picker), Gender (dropdown), Other IDs (text input), First Name (text input, containing 'Mason'), Last Name (text input), Vital Status (dropdown), Subject Status (dropdown), and Consent Status (dropdown). Below the search form is a table showing a single subject record:

Subject UID:	Full Name:	Previous Last Names:	Date of Birth: (dd/mm/yyyy)	Vital Status:	Gender:	Subject Status:	Consent Status:	Other IDs:
DEM-00000001	Mason Neumann		17/11/1983	Alive	Male	Subject	Consented	

At the bottom of the search form, there are buttons for Search, Reset, and New. The footer of the page includes copyright information: 'the-ark 1.2 Copyright © the-ark.org.au. All rights reserved.' and 'WICKET-AJAX DEBUG'.

- You may enter all or part of a name in the First Name, Middle Name, Last Name or Previous Last Name fields and the system will return all participants that contain the text that you entered in their names.
- You may enter a date of birth by typing into the Date of Birth field or using the date picker.

Once you have found the correct subject and have clicked on the Subject UID link, that subject will be placed "in context". When a subject is in context, their data is shown when navigating to different pages of The Ark. The subject that is currently in context is shown above the main tabs.

In order to navigate through the returned search results you may click on:

- "<<" to go back to the first page of search results
- "<" to go back to the previous page of search results
- 1 or 2 or 3... to go to the corresponding page of search results
- ">" to go back to the next page of search results
- ">>" to go back to the last page of search results

To reset the search fields, click on the *Reset* button.

To download the displayed fields for all of the subjects returned by the search you may click on one of the small icons at the bottom of the screen. The data can then be downloaded as either a csv file, an Excel file or as a pdf file.

If you are unsure of the study the subject belongs to, you can also use the Global Search functionality. This search will only search within the studies the user belongs to, and it may appear to return duplicate results if a matching subject is in a study and a sub-study.

S Centre for
Cancer Epidemiology
& Biostatistics
THE UNIVERSITY OF WESTERN AUSTRALIA

Welcome arksuperuser@ark.org.au | Logout

Study Subject Datasets LIMS Work Tracking Global Search Admin

Demo

THE ARK
OPEN SOURCE SOFTWARE
FOR MEDICAL RESEARCH

Demographic Data Biospecimen Search

Study: Choose One
Subject UID:
Middle Name:
Date of Birth: (dd/mm/yyyy)
Gender: Choose One

Other IDs:
First Name:
Last Name:
Vital Status:
Subject Status:
Consent Status:
Other IDs:

Search Reset New

Subject UID:	Study:	Full Name:	Previous Last Names:	Date of Birth: (dd/mm/yyyy)	Vital Status:	Gender:	Subject Status:	Consent Status:	Other IDs:
DEM-00000001	Demo 2	Martin Tuma		19/03/1965	Unknown	Male	Subject	Pending	
DEM-0000000001	Demo	Mason Neumann		17/11/1983	Alive	Male	Subject	Consented	
DEM-0000000002	Demo	Declan Crooks		02/02/1979	Alive	Male	Subject	Pending	
DEM-0000000003	Demo	Elizabeth Dallas		28/06/1935	Alive	Female	Subject	Pending	
DEM-0000000004	Demo	Hannah Smith		10/03/1995	Alive	Male	Subject	Pending	
DEM-0000000005	Demo	Piper Garnett		02/12/1941	Alive	Female	Subject	Consented	
DEM-0000000006	Demo	Tyson Clarkson		06/10/1980	Unknown	Male	Subject	Pending	
DEM-0000000007	Demo	Holly Irwin		13/01/1948	Alive	Unknown	Subject	Consented	

Total: 8

<< f >>

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

WICKET AJAX DEBUG

Editing a Subject's Details

Go to the Subject Detail screen by clicking on the corresponding participants Subject UID. On the Subject Detail screen edit subject details and click on “Save”. A message in green will appear at the top of the screen as seen in the example below.

The screenshot shows the ARK software interface for editing a subject's details. The title bar says "Demo". The main header includes the Centre for Genetic Epidemiology & Biostatistics logo and the ARK logo. The top navigation bar has tabs for Study, Subject, Datasets, LIMS, Work Tracking, Global Search, and Admin. The "Subject" tab is selected. Below the tabs, a message says "The Subject with Subject UID: DEM-00000001 has been updated successfully and linked to the study in context Demo". The main form contains demographic data fields:

Subject UID:	DEM-00000001 *	Other ID:	Other ID Source	Other ID
Title:	Unknown	Middle Name:	Frederick	
First Name:	Mason	Preferred Name:		
Last Name:	Neumann	Vital Status:	Alive	
(previous) Note: Updated when Last Name changed		Cause of Death:		
Date of Birth: (dd/mm/yyyy)	17/11/1983 Current age: 32 years, 3 months	Subject Status:	Subject	*
Date Last Known Alive:		Preferred Contact Method:	Choose One	
Date of Death: (dd/mm/yyyy)		Other Email:		
Gender:	Male	Email Status: (other):	Choose One	
Marital Status:	Choose One	Comments:		
Email: (preferred)				
Email Status: (preferred)	Choose One			
Heard About Study From:				
Consent Date:		Consent Status:	Consented	*
Consent Type:	Choose One	Consent Downloaded:	Choose One	
Consent to Passive Data Gathering:	Choose One	Consent To Active Contact:	Choose One	
Consent to Use Data:	Choose One			
□ Consent History				

At the bottom right of the form, there is a small red box containing the text "WICKET AJAX DEBUG".

Adding, searching and editing Phone Numbers

A subject may have multiple phone numbers. When you click on the Contact tab you will be presented with both the Phone screen and the Address screen, the top section of the page refers to Phones.

The screenshot shows the ARK software interface for managing subject data. The top navigation bar includes the Centre for Genetic Epidemiology & Biostatistics logo, the study name "Demo", and the subject UID "DEM-00000001". The main content area is titled "Demo" and displays two phone entries:

ID:	Area Code:	Phone Number:	Phone Type:	Phone Status:	Valid From:	Valid To:	Preferred Phone Number:
1	02	42061196	Home	Current			<input checked="" type="checkbox"/>
2		0410256985	Mobile	Current			<input type="checkbox"/>

Below the phones, there is an address entry:

Street Address:	City:	State:	Post Code:	Country:	Address Type:	Date Received: (dd/mm/yyyy)	Preferred Mailing Address:
12 Norton Street	ST IVES CHASE	New South Wales	2075	Australia	Residential		<input checked="" type="checkbox"/>

At the bottom of the screen, there is a footer note: "the-ark 1.2 Copyright © the-ark.org.au. All rights reserved." and a "WICKET AJAX DEBUG" link.

To add a new phone number, click on the “New” and complete the fields on the Phone Details screen

The screenshot shows the ARK software interface for adding a new phone number. The top navigation bar includes the Centre for Genetic Epidemiology & Biostatistics logo, the study name "Demo", and the subject UID "DEM-00000001". The main content area is titled "Demo" and displays a form for entering a new phone number:

ID: <input type="text"/>	Area Code: <input type="text"/>	Phone Type: <input type="button" value="Choose One"/>
Phone Number: <input type="text"/> *	Preferred Phone Number: <input type="checkbox"/>	Note: The first entry of a phone number for a subject is automatically set to be preferred.
Source: <input type="text"/>	Phone Status: <input type="button" value="Choose One"/>	Date Received: <input type="text"/>
Silent Number: <input type="button" value="Choose One"/>	Valid From: <input type="text"/> <input type="button" value=""/>	Valid To: <input type="text"/> <input type="button" value=""/>
Comment: <input type="text"/>	Save Cancel Delete	

At the bottom of the screen, there is a footer note: "the-ark 1.2 Copyright © the-ark.org.au. All rights reserved." and a "WICKET AJAX DEBUG" link.

- The Phone ID is a system-generated field and cannot be modified

- The Phone Type may be Mobile, Work or Home
- Area code is not a mandatory field
- The Phone Number is a mandatory field
- Phone Status may be Current, Current Alternative, Current Under Investigation, Incorrect or Disconnected or Valid Past
- Source is used to document where the number came from
- The date the number was received by research staff

Adding, Searching and Editing Addresses

A subject may have multiple addresses. When you click on the Contact tab you will be presented with both the Phone screen and the Address screen, the bottom section refers to Addresses.

To add a new address, click on “New”.

The screenshot shows a web-based application window titled "Demo". At the top left is the "Centre for Genetic Epidemiology & Biostatistics THE UNIVERSITY OF WESTERN AUSTRALIA" logo. At the top right is the "THE ARK OPEN SOURCE SOFTWARE FOR MEDICAL RESEARCH" logo. The main header bar includes "Study: Demo" and "Subject UID: DEM-00000001". Below the header is a navigation menu with tabs: Study, Subject, Datasets, LIMS, Work Tracking, Global Search, Admin, Demographic Data, Study-specific Demographic Data, Contact, Attachments, Consent, Correspondence, Clinical Data, Subject Biospecimen, Biospecimen Search, Pedigree, and Affection. The "Contact" tab is currently selected. The main content area displays a form for entering a new address. The fields include: Building Name, Unit Number or C/O (input field); Street Address (input field with asterisk); City (input field with asterisk); State (dropdown menu labeled "Choose One"); Post Code (input field with asterisk); Valid From (input field with calendar icon); Country (dropdown menu labeled "Australia"); Source (checkbox); Preferred Mailing Address (checkbox with note: "Note: The first entry of an address for a subject is automatically set to be preferred."); Address Type (dropdown menu labeled "Choose One"); Address Status (dropdown menu labeled "Choose One"); Date Received (input field with calendar icon); Valid To (input field with calendar icon); and Comments (text area). At the bottom of the form are "Save", "Cancel", and "Delete" buttons. A small note at the bottom center says "the-ark 1.2 Copyright © the-ark.org.au. All rights reserved." and a red box in the bottom right corner says "WICKET AJAX DEBUG".

The Address Detail screen shows the following fields:

- Building Name
- Street Address
- City (Suburb)
- Post Code (Zip code)
- Country
- State - Note that if the current installation of The Ark does not have all of the states loaded for a given country then this field will have been populated using free text entry. Only if the states are loaded into the system for the chosen country will the state have been populated from a drop-down list relevant to the country chosen.
- Address Type - Home or Work
- Address Status - Current, Current - Alternative, Current - Under Investigation, Incorrect Address, Valid Past Address
- Preferred Mailing Address - Yes or No. Note that only one address may be the current mailing address for any given subject.
- Date Received - the date upon which the address was received by the study coordinators
- Describes where the address came from
- Comments

To edit an existing address, go to the Address Detail screen by clicking on the street address of the corresponding record in the list. Once you complete editing, click on “Save”.

Attachments

It is possible to upload files and store them as attachment for any subject.

Note that there is an upper limit of 200 MB per image. It is possible to associate an attachment with a predefined study component but this association is not necessary.

Once you have found the required attachment record using the search screen you may then download the record by clicking on the Download File button associated with that record. At this point you also have the option of deleting the file by clicking on the Delete button, assuming you have the appropriate permissions.

Subjects, Collections and Biospecimens

All biospecimens are associated with a subject that must be defined within Subject tab. Each subject can then have one or more biospecimen Collections. A Collection can be recorded in the system to record the event or time at which one or more biospecimens were taken. For example, if a patient walks into a hospital and a phlebotomist takes a number of blood samples then a Collection record would firstly be created recording the date upon which the blood samples were taken. Likewise, the collection event could be when a surgeon took some tissue samples or where a clinician took tissue or saliva samples.

Once the Collection is defined in The Ark, then one or more physical biospecimens can then be associated with the Collection. Each biospecimen will receive a separate record in the system. All biospecimens must be associated with a Collection.

To illustrate this, you may have a subject called Joe Biggs. On the 1st of January 2014 Joe Biggs might have visited a clinic to give skin biopsy. A single collection would then be created with the following basic fields, however Collection UID is the only required field:

- Collection UID – This can be manually entered, or automatically generated by The Ark, depending on the study configuration
- Collection Name – This is a user-friendly identifier for the collection. From the example above, we could set the Collection Name to be “Skin Biopsy at Clinic”
- Collection Date – The date upon which the biospecimens were taken
- Comments – Any additional comments.

After the Collection has been defined one or more biospecimens can then be associated with the Collection. In Joe Biggs' case a tissue sample and a blood sample may have been collected. Each physical biospecimen should then be defined to the system and associated with the Collection already created. Each Biospecimen will have the following basic fields which are required, however there are additional optional fields:

- Biospecimen UID – The biospecimen Unique Identifier, which is either manually or automatically created, depending on the Study configuration,
- Sample Type – Each biospecimen requires a type, e.g. Blood / Buffy Coat, Saliva etc.,
- Collection – The UID of the Collection this biospecimen has been associated with,
- Initial Quantity – The initial quantity of the biospecimen taken, as well as the unit in the dropdown box next to Initial Quantity,
- Treatment – How the sample was initially treated, e.g. Frozen,

The Subject Biospecimen Summary Page

This page shows all collections and biospecimens that are associated with the subject in context. It also shows a tree with the biospecimen hierarchy, as a biospecimen may be a child of another biospecimen if it is the result of some processing e.g. Extracted DNA from a Blood Sample.

The screenshot shows the ARK software interface for the 'Demo' study. The top navigation bar includes links for Study, Subject, Datasets, LMS, Work Tracking, Global Search, Admin, Demographic Data, Study-specific Demographic Data, Contact, Attachments, Consent, Correspondence, Clinical Data, Subject Biospecimen, Biospecimen Search, Pedigree, and Affection. The main content area displays demographic data for a subject with Subject UID DEM-00000001, First Name Mason, Date of Birth 17/11/1983, and Gender Male. Below this is a table of biospecimens:

ID:	Biospecimen UID:	Study:	Collection UID:	Collection Name:	Collection Date:	Comments:
1	DES-000001	Demo	DEC-00001		(dd/mm/yyyy)	14/03/2016

At the bottom of the screen, a footer bar includes the text 'the-ark 1.2 Copyright © the-ark.org.au. All rights reserved.' and a red button labeled 'WIKI KET AJAX DERUG'.

To search for another subject, you can either click on the Cancel button on the Demographic Data tab, or go back to the Subject Module by clicking on the Subject tab.

LIMS MODULE (Laboratory Information Management System)

The LIMS allows laboratory personnel to capture and track patient, biospecimen and inventory data. It integrates clinical and laboratory research information into a central repository, making it easier to add, search, retrieve and track data at any time. It provides biospecimen management with user-customisable inventory layout and management. It tracks biospecimen and sub-specimen to patients, sample banks, studies and storage locations.

The main LIMS tab is not used to manage Biospecimens on a per subject level. It is used to configure the LIMS inventory management (explained later), to configure Biospecimen Custom Fields, Barcode labels and to perform bulk uploads of biospecimen data.

Adding New Collection(s) or Biospecimen(s)

You may create new Collections and Biospecimens by clicking on the New button below each section on the Subject Biospecimen tab or you may edit an existing Collection or Biospecimen by clicking on the corresponding Collection Name or Biospecimen UID in the lists.

In addition to Biospecimen UID, Study, Sample Type and Collection Name, the Biospecimen list also includes the following information:

- Quantity of sample available
- Location - if the sample has been allocated into inventory there will be a view link which will show you the location information in a modal pop-up window
- An indication as to whether a barcode label has ever been printed for the sample

Locating and Managing Biospecimens

The screenshot shows a software application window titled "Biospecimen Detail". The interface is divided into several sections:

- Subject Details:** Shows Subject UID: DEM-00000001, Last Name: Neumann, First Name: Mason, Date of Birth: 17/11/1983.
- Specimen Identification:** ID: 1, Sample Type: Blood / EDTA-purple top (10mL), Sample Date: 14/03/2016, Processed Date: (dd/mm/yyyy), Comments: (empty), Grade: Choose One, Anticoagulant Type: Choose One, Protocol: Choose One, Quality: Choose One.
- Barcode and Status:** Biospecimen UID: DES-000001, Parent UID: DEC-00001, Collection: DEC-00001, Sample Time: 12:00 AM, Processed Time: 12:00 AM, Barcode: (checkbox checked), Is barcoded: (checkbox checked), Assumes label is attached when barcode is printed, Stored In: Choose One, Status: Choose One, Purity (260/280): (empty).
- Quantity and Treatment:** Quantity (avail.): 10.0 mL, Treatment: Unprocessed, Amount (qty * conc): 0.
- Transaction Log:** Shows a table with columns: ID, Transaction Date, Quantity, Status, Reason, Recorded by, Access Request. One entry is shown: ID 1, Transaction Date 14/03/2016, Quantity 10.0 mL, Status Initial Quantity, Reason (empty), Recorded by arksuperuser@ark.org.au, Access Request Delete.
- Location Details:** Shows "Biospecimen not allocated". Buttons: Allocate, Save, Cancel, Delete, History.
- Bottom Buttons:** Clone, Process, Aliquot, No. of labels: 1, Print Barcode(s), Print Straw Barcode(s).

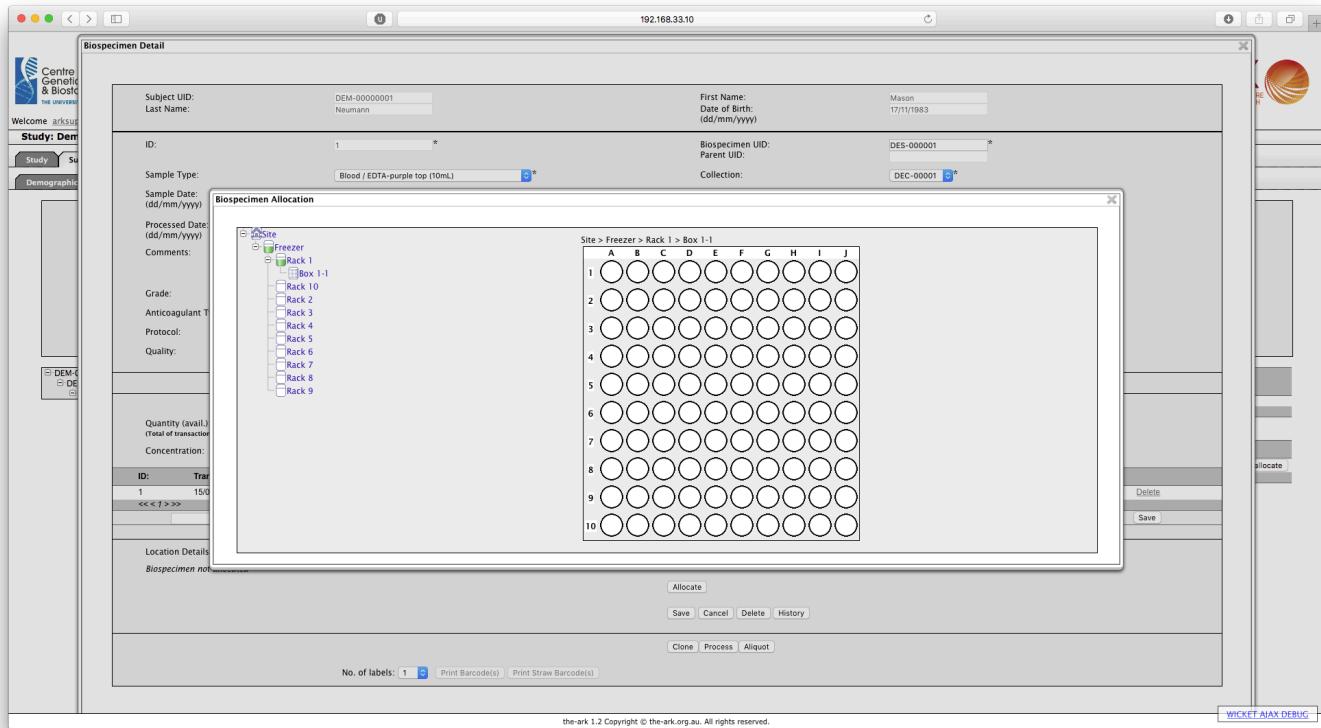
All biospecimens must be associated with a collection, which in turn must be associated with a Subject. There are a number of operations that can be performed on a biospecimen in addition to the original creation of the sample:

- Allocating a sample - putting the sample into inventory
- Deleting a sample - removing the sample from the system
- Modifying the biospecimen by editing the core and custom field information
- Cloning a sample - making a copy of the sample. This is typically done when several physical samples, such as bloods, have been taken from a subject at the same time.
- Processing a biospecimen - taking all of a sample and turning into something else. For example, taking a blood sample and processing all of it into DNA
- Creating an aliquot from a biospecimen - taking a subset of the material and creating a new sample from this subset
- Printing a barcode label for the sample.

Allocating a Biospecimen to a Freezer Location

To allocate a biospecimen to a freezer location click on Allocate from the Biospecimen Detail screen. This will take you to the Biospecimen Allocation pop-up window.

1. You can click on a site to expand the list to display the freezers associated with that site.
2. You can click on a freezer to expand the list to display the racks associated with that freezer.
3. You can click on a Rack to expand the list to display the boxes associated with that rack.



Navigate down the hierarchy to the box to which you wish to allocate your biospecimen. Click on the box and the box details will be shown.

Hovering over an empty cell will show the coordinates of that cell. Clicking on an empty cell in the box will allocate the current biospecimen to that cell.

Hovering the cursor over a full cell will show summary details of the biospecimen allocated to that cell.

Biospecimen Transactions

Every time the volume of material in a biospecimen is updated a new transaction will be created. These transactions provide a history of the sample and can be seen on the main Biospecimen Detail screen. Processes such as sample creation, processing a sample and creating sub aliquots will all add transactions to a sample.

Whenever a transaction is associated with an increase in volume then a green upwards arrow will be displayed in the transaction summary. A decrease is shown as a red downwards arrow.

Cloning Biospecimens

Cloning a biospecimen creates a carbon copy of your biospecimen but does not clone the transactions. To clone your biospecimen, click Clone from the Biospecimen Detail screen.

When your biospecimen is cloned, The Ark creates an identical biospecimen. The barcode will depend on your study setup.

Processing a Sample

Processing a sample will cause the quantity on the original sample to be reduced to zero and cause a new child biospecimen to be created. The type and details of this child biospecimen may be entered as required.

Creating an Aliquot

An aliquot is a type of transaction where a volume of the biospecimen sample is taken away from the parent biospecimen and allocated to the aliquot.

The View biospecimen screen enables you to create aliquots.

To create an aliquot:

- Click Aliquot.
The Add biospecimen screen appears.
- Select a Type of Biospecimen: Sub-aliquot or Processing
- Follow instructions for Creating a new Biospecimen and click Save.

The biospecimen is saved. This automatically creates transactions in the parent biospecimen where a volume of the biospecimen sample is taken away and allocated to the aliquot.

Searching for a Biospecimen(s)

On the Biospecimen Search tab, you may search for the specific biospecimen(s) across multiple subjects within the same study.

The screenshot shows a web-based application window titled "Demo". At the top, there's a navigation bar with tabs: Study, Subject, Datasets, LIMS, Work Tracking, Global Search, Admin, Demographic Data, Study-specific Demographic Data, Contact, Attachments, Consent, Correspondence, Clinical Data, Subject Biospecimen, Biospecimen Search (which is the active tab), Pedigree, and Affection. On the left, there's a sidebar with a logo for "Centre for Genetic Epidemiology & Biostatistics, THE UNIVERSITY OF WESTERN AUSTRALIA" and a "Logout" link. The main content area has several input fields: "Study:" dropdown set to "Demo", "ID:" text input, "Sample Type:" dropdown set to "Choose One", and "Subject UID:" text input. To the right, there are fields for "Collection UID:", "Biospecimen UID:", and "Sample Date:" (dd/mm/yyyy). Below these are "Search" and "Reset" buttons. The main table area displays a list of biospecimens with columns: ID, Biospecimen UID, Study, Subject UID, Collection UID, Sample Type, Quantity (avail.), Location, and Is barcoded?. Each row contains a "View" link and a red "X" icon. The table lists 9 rows of data. At the bottom, there are navigation links (<< < > >>) and a "the-ark 1.2 Copyright © the-ark.org.au. All rights reserved." footer. A "WICKET AJAX DEBUG" message is visible in the bottom right corner.

ID:	Biospecimen UID:	Study:	Subject UID:	Collection UID:	Sample Type:	Quantity (avail.):	Location:	Is barcoded?:
1	DES-0000001	Demo	DEM-00000001	DEC-00001	Blood / EDTA-purple top (10mL)	10.0 mL		
2	DES-0000002	Demo	DEM-00000001	DEC-00001	Blood / EDTA-purple top (10mL)	10.0 mL		
3	DES-0000003	Demo	DEM-00000001	DEC-00001	Saliva / for Metagenomics (NOT Orogenic)	1.0 units		
4	DES-0000004	Demo	DEM-00000002	DEC-00002	Blood / EDTA-purple top (10mL)	10.0 mL		
5	DES-0000005	Demo	DEM-00000003	DEC-00003	Blood / EDTA-purple top (10mL)	10.0 mL		
6	DES-0000006	Demo	DEM-00000004	DEC-00004	Blood / EDTA-purple top (10mL)	10.0 mL		
7	DES-0000007	Demo	DEM-00000005	DEC-00005	Blood / EDTA-purple top (10mL)	10.0 mL		
8	DES-0000008	Demo	DEM-00000006	DEC-00006	Blood / EDTA-purple top (10mL)	10.0 mL		
9	DES-0000009	Demo	DEM-00000007	DEC-00007	Blood / EDTA-purple top (10mL)	10.0 mL		

In the Global Search tab, biospecimens can be searched across all studies that the user has access to.

Batch Allocation

The Batch Allocation function is used to rapidly allocate already barcoded samples into inventory. To allocate samples to a box:

1. Go to LIMS module.
2. Within the Inventory tab navigate to the specific box where the samples are to be allocated
3. Click on the Batch Allocate button
4. Start scanning the barcoded samples one-by one, placing the scanned samples in the next available location.

Note that the samples will be allocated starting at the top left of the box and working along each row to the right and then down the rows from the top to the bottom. Already filled locations will be automatically skipped as the samples are allocated.

The screenshot shows the ARK LIMS software interface. At the top, there's a navigation bar with tabs for Study, Subject, Datasets, LIMS, Work Tracking, Global Search, Admin, Inventory, Biospecimen Upload, Custom Field Category, Custom Field, Custom Field Upload, and Barcode Labels. The 'Inventory' tab is selected. On the left, a tree view shows a Site node with a Freezer node, which has a Box 1 node, which in turn has a Box 1-1 node. Under Box 1-1, there are 10 rows labeled 1 through 10, each containing 10 columns labeled A through J. The grid is mostly empty circles, indicating available locations for allocation. A legend at the bottom defines symbols: an empty circle for 'An empty cell', a grey circle for 'A used cell', a black circle for 'An inaccessible used cell', and a barcode icon for 'A barcoded cell'. There's also a 'Download as XLS' link. The title bar says 'Demo' and the address bar shows '192.168.33.10'.

Unallocating Biospecimens

Biospecimens may be removed or unallocated from the biospecimen inventory. To unallocate a biospecimen, locate the biospecimen and go to the detail screen. If the biospecimen is currently allocated, then you will be able to click on the Unallocate button.

DATASET MODULE

A Dataset within The Ark is defined as a set of data about a group of subjects. Dataset data can be questionnaire results, clinical exam results, or laboratory results. It can be any type of data where the result can be represented in textual form, as a date or as a number.

The Dataset module in The Ark has no limitations in the number of variables that may be defined for a study or the number of times a value may be recorded for a specific variable for a specific subject over time.

The core of the Dataset module is the Data Dictionary that where each data variable is defined for the study. A Dataset may include data on any number of variables predefined in the Data Dictionary for the study in context.

The process for creating a new Dataset is to:

1. Determine the set of variables about which data was collected
2. Determine the questions or textual prompts that will be used with each variable. For example, the question may be: "Does the subject smoke?". The response may be an encoded character variable SMOKE with a value of 0=No and 1=Yes.
3. Define the variable to the Data Dictionary for the study, if it has not already been defined.
4. Create a new Dataset Definition, for example "Questionnaire 1"
5. Associate the appropriate variables with the new Dataset Definition
6. Publish the Dataset Definition.
7. Either bulk load or manually enter the Dataset data for each subject.

Note: Once a Dataset has been published, it can't be modified. You should ensure that the Dataset is correct before publishing.

The Data Dictionary

When you select the Dataset tab you will initially be shown the Data Dictionary Search screen with a set of search results below.

The screenshot shows a web-based application window titled "Demo". At the top left is the logo for "Centre for Genetic Epidemiology & Biostatistics, THE UNIVERSITY OF WESTERN AUSTRALIA". At the top right is the "THE ARK" logo with the tagline "OPEN SOURCE SOFTWARE FOR MEDICAL RESEARCH". The top navigation bar includes links for "Study", "Subject", "Datasets", "LIMS", "Work Tracking", "Global Search", and "Admin". Below this is a secondary navigation bar with links for "Dataset Categories", "Dataset Category Upload", "Data Dictionary", "Data Dictionary Upload", "Data Set Definition", and "Field Data Upload". The main content area contains several input fields and dropdown menus. On the left, there are fields for "ID", "Description", "Minimum Value", "Maximum Value", and "Units". On the right, there are fields for "Field Name", "Selected PhenoCategories", "Data Type", and "Question/Label For Field". Below these fields are buttons for "Search", "Reset", and "New". At the bottom of the page, there is a footer with the text "the-ark 1.2 Copyright © the-ark.org.au. All rights reserved." and a red "WICKET AJAX DEBUG" button.

Maintaining the study Data Dictionary can be achieved by adding and editing individual variables using the Data Dictionary data entry screens or may be achieved by bulk loading new or updated variables using the Data Dictionary uploader. To create a new variable, click on the New button on the search screen. To view or edit existing variable definitions, click on the Field Name of one of the variables in the results list.

Each variable definition may utilise the following attributes:

- Field ID - an internal system-generated identifier for the variable
- Field Name - the common name of the variable. This must be unique within a study.
- Field Type - this may be CHARACTER, NUMBER or DATE. Note that the type may not be changed if there is already subject data associated with this variable for the study in context.
- Description - a free text description of the variable. This should be sufficiently explanatory to enable researchers to work out what the variable actually represents, and possibly how the data was collected.
- Question/Label For Field - this is the question text that will be used by the system whenever the variable is to be included in a questionnaire.
- Units - this is a drop-down list where the units of measure for a variable may be recorded, such as metres, km, litres, etc. Note that this list may be extended by sending a request to The Ark development team.
- Encoded Values - Mapping the textual value of a character variable to a numeric value. For example, a sex of Male may be recorded as a "1" within the system and a sex of Female may be recorded as a "2". The way in which the encoding is captured within The Ark data dictionary is to define value pairs separated by semicolons, for example "0=Female; 1=Male;". Each encoded value pair should be on a new line, e.g.

0=Female; 1=Male; 2=Unknown;

There should be a semicolon at the end of the line.

- Minimum Value - the minimum value that should be stored for this variable
- Maximum Value - the maximum value that should be stored for this variable
- Missing Value - This allows a value to be associated with collected data that makes it clear that the data value was not present at the time of collection, as opposed to no data for a given questionnaire, exam, etc. being collected for that subject.

The Data Dictionary Uploader

The Data Dictionary Uploader can be used to load a large number of variables into the system from a file.

The screenshot shows a web-based application window titled 'Demo'. At the top left is the logo for 'Centre for Genetic Epidemiology & Bioinformatics, THE UNIVERSITY OF WESTERN AUSTRALIA'. At the top right is the 'THE ARK' logo with the tagline 'OPEN SOURCE SOFTWARE FOR MEDICAL RESEARCH'. The main header says 'Study: Demo'. Below the header is a navigation menu with tabs: Study, Subject, Datasets, LIMS, Work Tracking, Global Search, Admin, Dataset Categories, Dataset Category Upload, Data Dictionary, Data Dictionary Upload, Data Set Definition, and Field Data Upload. The 'Data Dictionary Upload' tab is currently selected. The main content area is titled 'Step 1/5: Select data file to upload'. It contains instructions: 'Select the file containing data, the file type and the specified delimiter, click Next to continue.' There is a 'Filename:' input field with 'Choose File' and 'no file selected' text, and a 'Delimiter Type:' dropdown set to 'COMMA'. A note below states: 'Note: delimiter type is ignored when uploading Excel (.xls) files. Variability across Excel versions means that The Ark cannot always extract data from Excel spreadsheets accurately. Converting Excel files to a UTF-8 comma separated value (CSV) format is recommended.' Below this are 'Next', 'Cancel', and 'Finish' buttons. At the bottom of the form is a row with fields: 'Upload ID:', 'Filename:', 'File Format:', 'UserID:', 'Start Time:', 'Finish Time:', 'Download Template', and 'Status:'. The status field has a dropdown with '<< >>' options. At the very bottom of the page is a footer with the text 'the-ark 1.2 Copyright © the-ark.org.au. All rights reserved.' and a link 'WICKET-AJAX DEBUG'.

To upload new variables into the Data Dictionary:

1. Firstly, you need to choose the file that contains the field definitions that you want to upload. The Data Dictionary Uploader supports files of type XLS, CSV or TXT but strongly advises you to use .CSV format. Note that you may click on the Download Template button to download a template file that you can use as a basis for your field definition file.
2. Then you need to choose the delimiter being used within the file, this will usually be a COMMA but may be of type - COMMA, TAB, PIPE (|), COLON (:) or AT SYMBOL (@)
3. Then you may click the Next button to commence the file validation process.
4. If the file format is validated correctly by the system then you will then be shown a sample of the file that you are uploading, as The Ark Data Dictionary Uploader understands it, with the column headers and data for the first 10 variables. If the file format validation fails the checks you should be notified at this point of why the failure occurred, and the only option will be to Cancel the upload. If the first 10 lines rendered on the screen do not appear correct, even if the format check completes successfully, then click Cancel. If the first 10 lines have not displayed correctly, you will most likely need to amend your upload file or change the parameters defined on the previous screen for the upload. If the file format validation process has passed successfully then click Next.
5. If the data in the file is validated correctly by the system then you will again be shown the first 10 variables from the file and you will be asked if you want to update any data for variables that are already defined for the study in context. If you do not click on the checkbox then only

new variables will be loaded. If the data validation step has completed successfully then click Next, otherwise click Cancel.

6. After clicking Next, you will be prompted to complete the process by clicking Next again. At this point you may still cancel out of the upload
7. If the upload is successful, you will see a message to that effect. Clicking Finish will then take you back to the main Data Dictionary Upload screen.

Viewing the status of past uploads

The results of all past variable definition uploads are displayed on the bottom part of the Data Dictionary Upload screen. There will be one row displayed per past upload. Each row of information contains the following:

- The filename of the file that was uploaded
- The user id of the user that uploaded the file
- The start and finish times and dates of each upload
- A Download File button that can be clicked to download a copy of the file that was uploaded into the system.
- A Download Report button that can be clicked to download a copy of the report that was generated during the upload.

Defining Datasets

All Data Dictionaries uploaded into The Ark needs to be associated with a Dataset Definition. A Dataset is then a set of information about a group of subjects that logically belongs together, such as a set of questionnaire results or clinical exam results. Once a Dataset Definition has been defined to The Ark for the study in context you may add or update the Data Dictionaries on an ongoing basis until it is been PUBLISHED.

Note that the Dataset Definition tab is designed to allow the creation and editing of existing Dataset Definitions. It is not for editing Dataset data associated with a specific subject or subjects.

When you click on the Dataset Definition tab in the Dataset module you will be shown an alphabetically sorted list of the Datasets already defined for the study. You may then select a Dataset Definition to view or edit, you may create a new Dataset Definition, or you may view the next page of Dataset Definitions.

The screenshot shows a web-based application window titled 'Demo'. At the top left is the logo for 'Centre for Genetic Epidemiology & Biostatistics, THE UNIVERSITY OF WESTERN AUSTRALIA'. At the top right is the 'THE ARK OPEN SOURCE SOFTWARE FOR MEDICAL RESEARCH' logo. The main header says 'Study: Demo'. Below the header is a navigation menu with tabs: Study, Subject, Datasets, LIMS, Work Tracking, Global Search, Admin. The 'Datasets' tab is highlighted. Underneath the menu is a sub-menu with tabs: Dataset Categories, Dataset Category Upload, Data Dictionary, Data Dictionary Upload, Data Set Definition, Field Data Upload. The 'Data Set Definition' tab is highlighted. A large text area below the sub-menu contains the instruction: 'View or edit existing dataset definitions, or define a new dataset.' A small 'New' button is located in this area. Below this area are two input fields: 'ID:' and 'Data Set Name:', followed by a 'Description:' field and a 'Published:' field. At the bottom of the page are links: 'Go to # on this page', 'the-ark 1.2 Copyright © the-ark.org.au. All rights reserved.', and 'WICKET-AJAX DEBUG'.

To view an existing Dataset Definition, click on the name of the Dataset Definition. From the detail view screen, you may then edit the Dataset Definition, assuming that you have the authorisation to do so and that the Dataset Definition has not yet been published. To create a new Dataset Definition, click on the New button.

- The Dataset Name is a mandatory field
- You may add a Description of the Dataset
- The Custom Fields selector contains a list of those fields from the study data dictionary that are associated with the Dataset Definition. You may use the left and right arrow buttons on this control to add or remove fields from the Dataset Definition. Available fields are on the left and chosen fields are on the right.

The up and down arrows can be used to change the sequence in which the fields in a Dataset Definition are displayed. Note that this will change the order in which fields are displayed on the Electronic Data Capture (EDC) screens.

You cannot remove fields from a Dataset Definition if there is already data associated with those fields in the Dataset.

After you have chosen the fields for a Dataset you may then determine which of the fields on the Dataset are mandatory and which are not. By default no fields are mandatory.

Clicking on the Field Name of the appropriate field on the bottom half of the screen will take you to a new screen where you can make a field mandatory.

You may make a field mandatory by clicking on the Required check box.

If the field has Encoded Values, then there is an option to allow multiple selections (i.e. Checkboxes) for the data entry. This option is accessible on the same screen as the Required check box.

Bulk Dataset Data Upload

The appropriate Dataset must be defined, and published into which data will be loaded. Note that this facility is just for loading data and not for performing updates.

The process to load the data is as follows:

1. Choose the file that contains the Dataset data that you want to upload. The Dataset Data Uploader supports files of type XLS or CSV. We strongly advise you to prefer CSV format.
2. Choose the name of the Dataset into which the data will be loaded.
3. Determine whether you want the data to be uploaded even if not all of the data meets the field constraints defined in the data dictionary.

Once the upload is complete you will be able to access the corresponding upload report.

The results of past Dataset data uploads are displayed on the bottom part of the upload screen.

The screenshot shows a web-based application window titled "Demo". At the top left is the logo for "Centre for Genetic Epidemiology & Biostatistics, THE UNIVERSITY OF WESTERN AUSTRALIA". At the top right is the "THE ARK" logo with the tagline "OPEN SOURCE SOFTWARE FOR MEDICAL RESEARCH". The main header says "Study: Demo". Below the header is a navigation menu with tabs: Study, Subject, Datasets, LIMS, Work Tracking, Global Search, Admin, Dataset Categories, Dataset Category Upload, Data Dictionary Upload, Data Set Definition, and Field Data Upload. The "Data Set Definition" tab is currently selected. A sub-menu for "Field Data Upload" is open, showing "Step 1/5: Select data file to upload". It instructs the user to "Select the file containing data, the file type and the specified delimiter, click Next to continue." There are fields for "Filename:" (with a "Choose File" button), "Data Set:" (with a "Choose One" button), "Comment:", "Update (if exists):" (checkbox), and "Delimiter Type:" (radio buttons for COMMA and TAB, with TAB selected). Below these fields are notes about Excel compatibility and a "Download Template" link. At the bottom of the sub-menu are "Next", "Cancel", and "Finish" buttons. The footer of the page includes "the-ark 1.2 Copyright © the-ark.org.au. All rights reserved." and "WICKET-AJAX DEBUG".

Manually entering the data subject-by-subject

To enter data for one subject, follow the process below:

1. Click on “Subject Module” tab.
2. Select the subject that the data belong.
3. Click on “Clinical data” tab.
4. On the clinical data screen, click on New.

The screenshot shows a web-based application window titled "Demo". At the top left is the logo for the "Centre for Genetic Epidemiology & Biostatistics, THE UNIVERSITY OF WESTERN AUSTRALIA". At the top right is the "THE ARK" logo with the tagline "OPEN SOURCE SOFTWARE FOR MEDICAL RESEARCH". The URL in the address bar is "192.168.33.10". The main header includes "Welcome arksuser@ark.org.au | Logout", "Study: Demo", and "Subject UID: DEM-00000001". Below the header is a navigation menu with tabs: Study, Subject, Datasets, LIMS, Work Tracking, Global Search, Admin. The "Subject" tab is selected. A secondary navigation menu below shows tabs: Demographic Data, Study-specific Demographic Data, Contact, Attachments, Consent, Correspondence, Clinical Data, Subject Biospecimen, Biospecimen Search, Pedigree, Affection. The "Clinical Data" tab is selected. The main content area has fields for "ID:", "Data Set:", "Comment:", "Status:", "Record Date:", and "Reviewed Date:". Below these fields are buttons for "New", "Select dataset", and "Quick export". At the bottom of the page, there is a copyright notice: "the-ark 1.2 Copyright © the-ark.org.au. All rights reserved." and a link to "WIKI & API DOCUMENTATION".

5. In the "Subject Dataset Details" screen, select the Data Set (ie, questionnaire, clinical examination), enter the record date (this could be chosen as the current date by clicking on the calendar icon) and select the status of the dataset.
6. Click on the “Save” button.

Demo

The screenshot shows the ARK software interface with a central window titled "Subject Dataset Details". The window contains fields for "ID", "Record Date" (set to 17/04/2014), "Reviewed By" (set to "Choose One"), "Comment", "Data Set" (set to "EPS_E"), "Status" (set to "Uploaded From File"), and "Reviewed Date". Below these fields are additional input fields for "DOB" (11/01/1950), "XDAT" (19/05/2012), and several dropdown menus for participant details like "Blood or saliva sample given?", "Does the participant have glasses?", "Does the participant wear contact lenses?", "Does the participant have their glasses with them today?", "Does the participant have their contact lenses with them today?", and visual acuity measurements (IPD, RS, LS, RC, LC) with values ranging from -0.75 to 65. A warning message at the bottom of the window states: "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." At the bottom right of the window are "Save", "Cancel", and "History" buttons.

As soon as you select save the first page of the corresponding record's fields will be displayed.

The screenshot shows the ARK software interface with the "Subject Dataset Details" window open. The fields have been updated from the previous screenshot. The "Record Date" is now set to 30/04/2014, and the "DOB" is now set to 11/01/1950. The "XDAT" field remains at 19/05/2012. The dropdown menus for participant details show different responses compared to the previous screenshot. The visual acuity measurements remain the same. The copyright notice at the bottom of the window reads: "the-ark 1.1c Copyright © the-ark.org.au. All rights reserved."

You may then update the fields for the Dataset record. Note that you must “Save” the data you enter or update before you move to the next page of the Dataset record or before you “Cancel” out of the Dataset record if you want your changes or default values to be saved.

On the “Clinical data” screen, you may choose a dataset and click on “Get Phenotypic data” to view the entered values for that dataset.