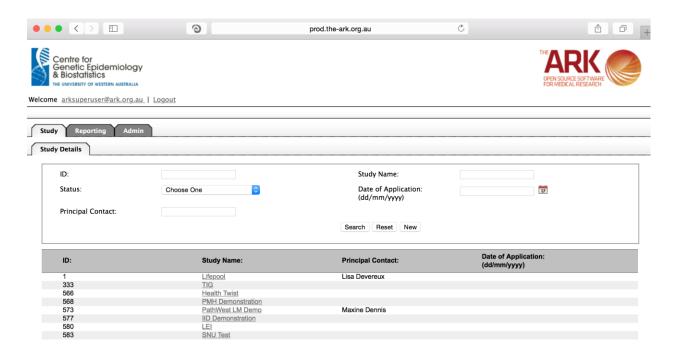
The Ark Documentation.

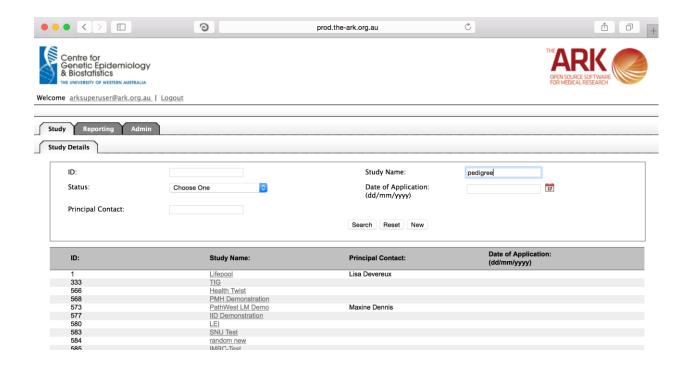
Data Centre

Log in and select your study by clicking the hyperlink corresponding to the name of your study.

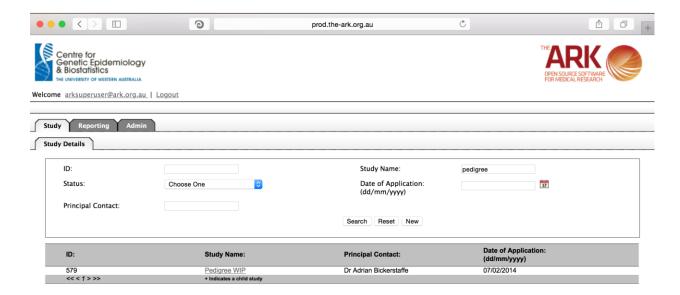
Once you have chosen your study we will be able to adjust study details or select modules and functions available in that study (for which you have privileges).



Alternatively if you have a very long list of studies under your management you may use the search pallette/panel at the top of the page. You can search/filter by id, name, principal contact, status of the study or date of application. Once you have added your filtering criteria, click the search button;



As you can see the filter has helped us find our chosen study and we can click the hyperlinked study name as discussed before;



Once you have chosen your study we will be able to adjust study details or select modules and functions available in that study (for which you have privileges). The module we wish to select is Genomics – Data Centre.

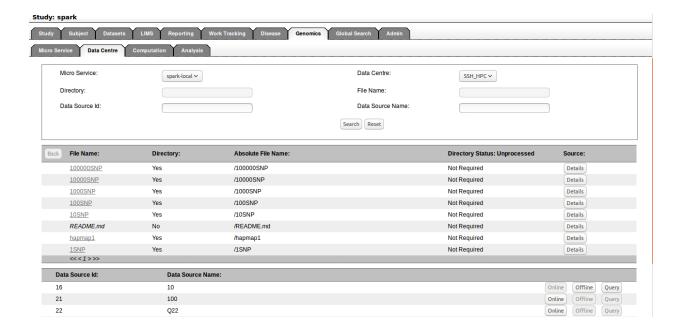
Genomics module is created with the intention of allowing collaborating researchers to administer the existing PLINK formatted GWAS datasets and execute GWAS analysis inside High-performance computing sources via a single web user interface.

This simple page is for searching available files and directories stored in the data centers attached to a micro service (These files/directories hold the original datasets to select researchers to start an analysis) and extracting the GWAS information by specific set of filters. As you can see the same standard layout of a list and filterable fields starts us off. To start a file/directory search, first need to select a micro service and then data centre attached to the service by clicking the drop down box.

Search for a data source reside in data centre.

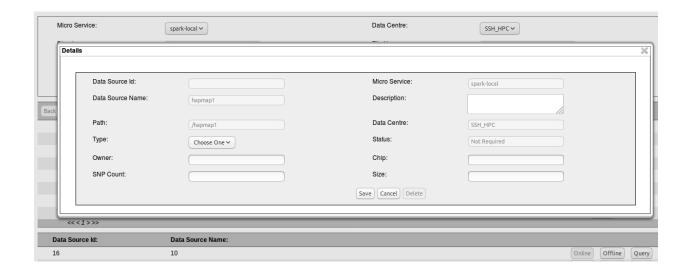
After that clicking the search button, top panel list the directories and files in the root directory of the data centre. By clicking the directory name users can traverse into the directory content and view the sub directories and files. This is an iterative process and users can traverse until the last leaf directory in the file structure. When navigate to specific directory users can use back button to traverse back to its parent directory.

The bottom panel lists the data sources available for selected Micro Service and Data Centre. In addition to above filtering criteria Data Source Id and Data Source Name filters applicable for the data source list. Data Source is the virtual representation of an selected directory in the data centre.



Creating a data source

Click on the detail button of the selected directory will open the details pop-up window. There users can define the existing Type, Owner, Chip, SNP count, and size of the dataset. Click on the Save button will create a new data source which list in the data source panel.



Make data source available for querying

Click on the Online button in data source panel will enable data source to be querying. When click the Online button will start a background process to enable dataset ready for querying. To check the online process status users need to refresh the screen by click the Search button. Once the Online process is completed, online button will be disabled and enable the Offline and Query buttons.

Remove data source from querying

Click on the Offline button in selected data source exist in data source panel which is already in online state. When click the Offline button will start a background process to remove data source from online stage and disable the querying. To check the Offline process status users need to refresh the screen by click the Search button. Once the Offline process is completed, online button will be enabled and disable the Offline and Query buttons.

Query data source

Click on the Query button in selected data source exist in data source panel which is already in online state. When click the Offline button will open a query Builder pop-up window. There user can select a Id type, and Individual Id. Else selecting a Id type only

will query for all the individuals belongs to the Id type. Other options are Family Id, Gender, Affected Status, and Output type. All the query parameters are optional and without specifying a query parameters will generate a copy of existing dataset. Then click the Query button to start the querying background process and close the pop-up window. There is a new data source is listed in the data source panel after click the Query button.

