

# Atlas of Living Australia Beta Test Plan

## Data Quality Filters

Thank you for taking the time to test the functionality developed in response to data quality (DQ) issues.

The text below indicates the problem that is being addressed, the solution developed by the Data Quality project team and notes on testing the functionality.

### How to test:

Please follow the links provided to the test environment and, referencing the functionality description, consider the questions in the testing section for that functionality.

e.g. on this URL [https://biocache-dq-test.ala.org.au/occurrences/search?q=%3A\\*](https://biocache-dq-test.ala.org.au/occurrences/search?q=%3A*) can you see what filters have been applied, switch them on and off and see what's happened?

Please record your testing results in the form here:

[https://docs.google.com/forms/d/e/1FAIpQLSd2UWZJU9jSQQTzg3V8I0JsopM0H2gO\\_bINYvudt74gUcYlw/viewform](https://docs.google.com/forms/d/e/1FAIpQLSd2UWZJU9jSQQTzg3V8I0JsopM0H2gO_bINYvudt74gUcYlw/viewform)

Or email them to [dataquality@ala.org.au](mailto:dataquality@ala.org.au)

As you explore the interface and functionality please make a note in the **Testing comments** section of the form of any functionality that doesn't work as expected or is unclear as to what is happening. If you have any other comments please add them as well.

## Data Quality Filters

### Problem(s):

- I don't know whether to use particular records or data sets for my purpose
- Users of the ALA are using data accessed via the ALA without filtering out records that are not fit for purpose.
- Users that do not know to filter data are using low quality records along with high quality records
- I don't know whether to use particular records or data sets for my purpose because quality indicators are not visible or I don't understand what the information means

### Solution:

Pre filter data on quality attributes to initially return a "higher quality" set of records

- That data is filtered must be very clear to users
- Filters can be turned off/on
- Filter details, including the description and query applied are visible
- The interface indicates any interaction between user filters and DQ filters

### Testing notes:

There are multiple parts to this solution:

1. A framework to administer the filters – enter a description and select the fields and thresholds, turn filters on and off.
2. The interface to apply, display and interact with the filters
3. The details of each individual filter – what fields and thresholds should be applied to only return “higher quality” records

1 - is an administration function and will be tested by the ALA

2 and 3 are now available for testing, individual filters may be added, removed or modified and made available for testing as they are refined and developed

### Functionality (Description and URL):

[https://biocache-dq-test.ala.org.au/occurrences/search?q=%3A\\*](https://biocache-dq-test.ala.org.au/occurrences/search?q=%3A*)

#### UI and filter application

- “Quality filters applied” section on search results above the record list
- “Quality filters” section in left hand facets pane
- <Disable All Quality filters> button
- On each filter:
  - Checkbox to enable/disable
  - “pop out” button to the DQ filter to the facets interface on the left side pane
  - Mouse over on filter name with description of the filter group
  - Information mouse over (i) with the filter query (fq) technical details
    - Click (j) to display a dialogue with more detailed information on fields
  - Count of records excluded by the filter
- Highlight of user filter and DQ filter where they interact

For each filter:

- Description
- Filter field(s) and threshold for each field

### Testing:

#### UI and filter application

[https://biocache-dq-test.ala.org.au/occurrences/search?q=%3A\\*](https://biocache-dq-test.ala.org.au/occurrences/search?q=%3A*)

- Is it clear what filters have been applied?
- Can you switch them off/on and is it clear what’s happened?
- Can you send the DQ filter to the facets to be modified?
- Can you remove the facet and re apply the DQ filter?
- Can you turn all filters on/off
- If a filter you set interacts with DQ filter can you see an indication that is occurring
- Can you understand what the filters are doing?
  - What fields they are operating on and with what thresholds
  - How many records are impacted

#### Filters

- Quality assertions
  - Description understandable and useful

- Choice of filter field(s) and threshold for each field
  - Name not supplied
  - Name not recognised
  - Invalid scientific name
  - Habitat mismatch
  - Cultivated escapee
  - Supplied coordinates are zero
  - Zero latitude
  - Zero longitude
  - Coordinates centre of country
  - Supplied coordinates centre of state
  - Coordinate conversion from grid reference failed
  - Coordinate conversion from decimal coordinates failed
  - Coordinates inverted
  - Coordinates out of range
- Location
  - Description understandable and useful
  - Choice of filter field(s) and threshold for each field
    - Habitat mismatch
    - Geospatial valid (this is an aggregate of a number of tests)
    - Coordinate uncertainty is not greater than 10km (records with no uncertainty are included in the records returned)
- Identification
  - Description understandable and useful
  - Choice of filter field(s) and threshold for each field
    - No taxonomic issue
- User assertions
  - Description understandable and useful
  - Choice of filter field(s) and threshold for each field
    - exclude Unconfirmed OR unresolved user assertions
- Event date
  - Description understandable and useful
  - Choice of filter field(s) and threshold for each field
    - Records after 1900
- Duplicates
  - Description understandable and useful
  - Choice of filter field(s) and threshold for each field
    - Exclude duplicate records
- Outliers
  - Description understandable and useful
  - Choice of filter field(s) and threshold for each field
    - Exclude records that are an outlier against 3 or more layers

## Github issue(s):

### UI and filter application

- UI for DQ filters - <https://github.com/AtlasOfLivingAustralia/DataQuality/issues/73>
- Read and apply DQ filter - <https://github.com/AtlasOfLivingAustralia/DataQuality/issues/58>
- See what data filters have been applied - <https://github.com/AtlasOfLivingAustralia/DataQuality/issues/37>
- Toggle DQ filters on/off - <https://github.com/AtlasOfLivingAustralia/DataQuality/issues/54>
- Manage interaction with user filters and DQ filters - <https://github.com/AtlasOfLivingAustralia/DataQuality/issues/60>
- Indication that a user filter query is interacting with a DQ filter - <https://github.com/AtlasOfLivingAustralia/DataQuality/issues/76>

### Filters

- Quality assertions - <https://github.com/AtlasOfLivingAustralia/DataQuality/issues/30>
- Location - <https://github.com/AtlasOfLivingAustralia/DataQuality/issues/31>
- Identification - <https://github.com/AtlasOfLivingAustralia/DataQuality/issues/32>
- User assertions - <https://github.com/AtlasOfLivingAustralia/DataQuality/issues/36>
- Event date - <https://github.com/AtlasOfLivingAustralia/DataQuality/issues/33>
- Duplicates - <https://github.com/AtlasOfLivingAustralia/DataQuality/issues/34>
- Outliers - <https://github.com/AtlasOfLivingAustralia/DataQuality/issues/35>