

CCM Tools

Introductory hands-on worksheet

Introduction & Accessing CCM

This worksheet gives you search examples you can work through to explore the main features of the CCM interface. Do not feel you have to follow the sheets exactly from start to finish; the scenarios and examples are simply there to help you explore the facilities. If you have questions please ask.

- » To access the CCM Tools interface go to: <https://ccm.copac.jisc.ac.uk>
- » Select the **Search** tab below the CCM logo in the top menu bar
- » Login using your Shibboleth username and password

Working with small amounts of material using the Standard Number Search

Scenario

You are in the University of Manchester library and you are reviewing a section of your South Asian related materials at the shelves. During this process you identify some older items that are in poor physical condition. This is not a heavily used part of the collection and when you check your catalogue you find that none of the items has been in circulation for several years. So you need to decide whether these materials need to be conserved or whether they can be withdrawn.

Using the Standard Number Search screen

Because you are dealing with a small number of items you can use the **Standard Number Search** form. First you must carry out a search for the ISBNs of the relevant materials:

- » Make sure you are in the **Standard Number Search** section of the main search screen
- » Enter the following ISBNs as a comma separated list into the **Standard numbers** field:
0140207694, 0858831775, 0333255801, 0195803930, 0333240634, 9780203885369
- » Check that the **Number type** option is set to 'ISBN'
- » Under **REFINE SEARCH** set the **Deduplicate** option to 'records with a shared ISBN'
- » Select **Search**

You will be presented with a Copac style result set of brief record details.

Search result display and visualisation

In the result display screen you can view a graphical representation of the results, or refine the results. There are also the normal Copac style options to browse multiple result pages, use the export options, or view the full record details.

- » Above the list of results, select the **blue graph thumbnail** on the left-hand side

This gives you a graph of the search results, showing the number of records held by each library in the result set. From this you can see that quite a few libraries have 4 of the documents but only two libraries have all 6.

- » Now select the **pink graph thumbnail** to the right

This shows you the number of contributors holding each of the documents in your list. You will only see this graph for small result sets.

You can see from this graph that 4 of the documents are widely held, so this might be useful supporting information when you are considering your own local criteria for withdrawal. However two documents are much less widely available.

- » Put your mouse pointer over the bar for ISBN **9780203885369**
- » A box will appear telling you that this is held by 3 libraries, of which Manchester is the only one outside London

As one of the holding libraries is the British Library you might decide to do a follow-up check of Copac or the BL catalogue to see if the book is held at Boston Spa, as this could be useful information to take into account during your decision making.

- » Now put your mouse pointer over the bar for ISBN **0858831775**

- » As before a box appears which shows that this appears to be a single volume of a periodical or multivolume work

Cataloguing for multi-volume items can be variable, making deduplication more problematic, so a follow-up search for the title would be needed to decide exactly how rare this document might be, but again it looks like you might have the only copy outside of London.

So, depending on the nature of the materials and likely cataloguing/deduplication issues, some follow-up may be required - but the results should help to focus your time on the less clear-cut cases and provide evidence to support decisions about the conservation needs of these materials.

Creating your own library groups

You can further refine your results by choosing only to search the holdings of a particular set of libraries – for instance libraries in a particular region, of a particular type or size, or specialising in a particular subject area. Returning to the **Standard Number Search form** you will see that you can manually select as many libraries as you wish from the list in the **Library box**, or a region or regions from the list in the **Region box**. However you can also create and save your own custom lists, which will then be available for you to search against whenever you log in to CCM Tools in the future.

- » Click the link to **Create/Edit Groups** under the **My Library Group** heading at the right-hand side of the screen
- » Type the name **Art Libraries** into the **Library group name** field on the following page, and then select the following libraries from the list below:
Courtauld Institute, National Art Library, National Portrait Gallery, Paul Mellon Centre, Royal Academy of Arts, Tate, V&A Museum
- » Click the green **Save** button
- » When you return to the main search form you will see that your **Art Libraries** group has appeared under the **My Library Groups** heading
- » Select it using the check box and then perform a standard number search as you did before using the following ISBNs:

0952702703, 0905080157, 0820423181, 0951142305, 0946511012, 0862417597

You will see that all of the same results display and visualisation options are available as for your previous search. Records are only returned for those items held by at least one of the libraries in your Art Libraries group (in this case all of them are held by at least one library, and in fact none of the items is particularly rare). However the results still provide details of libraries not in your list which hold copies of the items, and these are included in the visualisation graphs. This means that you would need to undertake a little more analysis of your own in order to understand the holdings situation for these items just within the libraries included in your custom group.

Working with large data sets using the Batch Search

To follow this batch search example you will need to use the test data file on your desktop:

» **ISBNs for withdrawal 1116.txt**

Scenario

You need to make space in part of your store, so you are undertaking a major stock withdrawal programme. You've used your LMS to generate a list of potential items for withdrawal using your local criteria. This has been edited to produce a comma separated file of ISBNs.

Using the Batch Search form

You need to carry out a search for the set of ISBNs you've identified as candidates for withdrawal:

- » Switch to the **Batch Search** section within the main search screen
- » Select **Choose File** and from the desktop select the file named **ISBNs for withdrawal 1116.txt**
- » Check that the **Number type** option is set to ISBN
- » Under **REFINE SEARCH** set the **Deduplicate** option to 'records with a shared ISBN'.
- » Select **Search**

When you run a batch search you are taken to the **Batch search in progress** screen which summarises the details of the search you've just requested. This search is fairly quick and you will be taken automatically to the display of results. However, a search for a large number of records can take a while, so whilst you are waiting you can do other searches if you wish.

At any point you can use the **Batch Search History** button towards the top right of the screen to review your batch searches. All previous batch searches will be listed in your batch search history, so you can review previous results sets.

Search results are retained for 1 month, but the searches themselves are currently retained indefinitely and can be re-run from here. Because the database is updated frequently the search results may be slightly different when you re-run a search.

Search results

In the results display screen you can view a graphical representation of the results, or refine the results. There are also the normal Copac style options to browse multiple result pages, use the export options, or view the full record details.

- » Select the drop-down menu beneath the **REFINE RESULTS** option on the left of the screen

This allows you filter your results by the number of holding libraries. As you are primarily interested in withdrawing material, you want to know which items are held by large numbers of institutions and can thus safely be withdrawn.

- » Select 'Records with 6 or more holding libraries' then click **Refine**

The result display is updated to show only those records with 6 or more holding libraries. The change in result set size is clear from the number of records and text shown in the breadcrumb trail towards the top of the display.

Search result visualisation

- » Above the list of results, select the **blue graph thumbnail** on the left-hand side

This gives you an overview of the collections in which each of the documents in your list are held. You decide you want to be more cautious about your withdrawals so change your view of the results again:

-
- » Under the **REFINE VISUALISATION** option on the left of the screen, use the drop-down menu to view only records held by 9 or more libraries

You feel this is a better cut-off for the time being and you can now see which materials are commonly held by other local libraries. However, to get a better understanding of the geographic distribution you might wish to check the map view.

- » Select the **map thumbnail** at the top of the screen
- » A map will load, showing you the location and distribution of collections which hold items included in your search
- » Click on one of the **red map points** – you will see the name of the holding library and the number of items from the search which it holds
- » Click on one of the **library names** in the list to the right – you will see the library's location on the map and the number of items from the search which it holds
- » Note that you can also use the zoom buttons to focus in on a region or get a wider national overview

You can now visualise more easily the geographical distribution of holdings for the items returned by your search, and as a result might decide to experiment with further search options, such as rerunning your search against a particular region or custom library group in order to obtain a more focused results set for analysis.

Exporting results

Once you have refined your results, you can export the data in a range of formats depending on what you want to do with it.

- » Look for the **Export** option at the bottom of the screen
- » Open the dropdown menu: this shows you the different formats available for export
- » Select the 'Item holdings data .csv' option from this list, then select the **Export** button

This provides you with a data file that you can open as a spreadsheet. You can now adapt this data as required to create your own graphs etc. using your spreadsheet facilities. The data export includes the ISBN from the records and could be used, for instance, to import the data into your LMS to generate lists with key bibliographical data such as Bib id, author, title, volume, publication date, class mark, barcode etc., to aid withdrawal of items from your collection.

Using the Keyword Search form

The keyword search allows you to be more exploratory by, for example, looking for collections by a particular author.

Scenario

The Music department at the University of Nottingham is developing a new interest in early music, so you are currently reviewing your collections in this area with a view to expansion - and you are in particular looking at materials for the lute.

- » Select the **Keyword Search** tab
- » In the **Subject** field enter **Lute songs**
- » Select **Search**

Search results

- » Above the list of results, select the **blue graph thumbnail** on the left-hand side

You can move your mouse over the bars of the graph to see the library name and number of records. Nottingham has 47 records with this subject but other libraries have rather more. Since this is an unfamiliar area, you might look at the results for comparable institutions that appear to have good collections in this area to see what types of recent materials, which publishers etc, they are collecting.

- » Next select the **green graph thumbnail** in the centre
- » This shows you the number of collections in which each of the documents in your result set is held, which might help you to identify particularly popular materials

If your library is a contributor to Copac your own library will always be listed first showing how many records represent materials only held by your institution.

To get a different view:

- » Go back to the search form and repeat the search
- » However this time, in the **Library** list under **REFINE SEARCH**, limit the search to materials held by **Nottingham University**

This time the graphs will more clearly emphasise the range of any apparently unique items held by Nottingham. Needless to say a subject search depends greatly on the record quality and use of consistent subject headings - which can be problematic, particularly for older materials. Whilst uniqueness is also affected by the ability to deduplicate effectively. However, if you are trying to assess the importance of a particular collection, for instance with a view to raising its profile within or outside of your institution, the results should be suggestive and help to identify areas worthy of further exploration.

Experimenting with the Deduplication options

Within Copac there is a detailed deduplication process that errs on the side of caution to reduce the likelihood of merging records for different documents in error. However, this does mean that in any result set there are likely to be duplicate records that we have been unable to match because of cataloguing variations.

However, when you are looking at data at the collection level (depending on the task in hand) you may want to reduce the number of duplicate records in a result set - for example, to get a clearer overview of the likely rarity of materials in a collection. To support this you have the option to impose additional deduplication on your search result set.

There are 4 levels of search result deduplication available:

- » Records with a shared ISBN
- » Level 1: uses Author, Title, Date, Publisher, Edition, Pagination
- » Level 2: uses Author, Title, Date, Publisher
- » Level 3: uses Author, Title

The ISBN deduplication merges records sharing the same ISBN. Errors and variations in the records can mean this results in some misconsolidations. It is likely to be uncommon, though we have no figures for this, but it does happen.

There are also three levels of multi-field deduplication offering increasingly broad matching of records and hence more likelihood that records from different libraries will match. The likelihood of records matching incorrectly will increase with each level; but cataloguing variations also mean the more stringent matches will retain greater numbers of duplicate records.

Deduplication searches

We will now try out the different deduplication options in the **Keyword Search** tab using the same search - for **Twelfth Night** - combined with each form of deduplication.

The suggested search finds quite a large result set – and whenever you think this is likely to occur with a keyword search it is a good idea to select the **Treat as a batch search** option in the middle of the search screen. Treating the search as a batch search allows you to do other searches whilst you are waiting for your search to be processed, whilst the data visualizations are available for much larger result sets in the batch search. If you get a very large result set in a 'live' keyword search the visualizations won't be available as they take too long to produce.

- » Using the Keyword search form do a **Title** search for **Twelfth Night** – select the option to do this as a batch search.
- » Note that whilst you are waiting for the batch search, on the **Batch search in process** screen, you can select the option to **do a new search**
- » Go back to the **Keyword Search** form and repeat the same search with each of the deduplication options – doing each one as a batch search.
- » Once the last search has run to completion the results will be automatically displayed. Note the number of records found.
- » Now go to the Search History and select each of your Twelfth Night searches in turn to check the number of records found by each type of deduplication (and none).

Depending on your search the various deduplication methods may have little impact on the results or a very significant impact. The search for Twelfth Night illustrates a modest change in results with each of the deduplication methods, but it may not always work so neatly.

You might like to consider the pros and cons of the various deduplication options and situations in which they might be useful. For example, the broad-brush 'Title and author' match would be fine for getting a general sense of how two collections compare, without getting side-tracked by differences in edition or multivolume works, whilst a much a lower level of deduplication (or even none at all) would be more appropriate in assessing the rarity of particular materials.

We would be very interested in hearing about your use of the deduplication options as you trial the CCM tools – please share your thoughts about what does and doesn't work for you, as well as anything new you'd like to see, by emailing us at help.copac@jisc.ac.uk.