
Contributing to the NBK:

Data supply supporting documentation

Updated: 23 August 2019

| | |
|--|----|
| Introduction | 2 |
| Data Supply questionnaire: Notes | 4 |
| Data Management questionnaire: Notes | 10 |
| Library Information questionnaire: Notes | 14 |
| Circulation Data questionnaire: Notes | 16 |

Introduction

Please read this document alongside the Contributor Guide to the National Bibliographic Knowledgebase.

The documentation expands on the Data supply questionnaires accessible from the NBK Contributor Dashboard (<https://contribute.copac.ac.uk/dashboard/>), providing more details about each question in terms of the information required and why this is needed.

If you have questions or comments regarding any aspect of the information being requested there is space for you to include these within your questionnaire response. When you have submitted your questionnaire we will review the information you have provided and respond to any questions you may have, or contact you to follow up any area where we would like more information.

Once we have all the information we need we will be in touch regarding potential dates for the data supply that would fit in with your other local activity.

Submitting the questionnaires:

You do not have to complete the questionnaires all at one time. For example, if you don't have all the information you need at hand, you can use the 'Save' button to save the information you have entered so far and come back to the page later to complete it.

Note: The NBK team will **not** act on any information you have supplied until you select the 'Form complete' option at the end then 'Save' the form again. If there are things you are unsure about just include questions for us in the relevant sections of the form and select 'Form complete' then 'Save'. We will follow up with you any questions you may have when we review the form.

Preparation & ongoing commitment:

Once we have agreed to include your catalogue on the NBK there are likely to be a number of preparations you need to make. This will vary between different libraries and library systems, but as a starting point the following suggestions are based on feedback from current contributors regarding some of the areas you might need to consider.

Which records are you sending?

Unless we have agreed to load a specific sub-set of your catalogue, we expect a full export of all the records you show to users of your local catalogue. So you need to identify how to exclude records that are not locally visible from the export.

If this is a major problem please let us know.

Which locations are you sending?

If you have locations that are not normally visible in your public catalogue you will need to filter these holdings out of your export. This may mean dropping one location from a record for a work held in several of your branch libraries.

If this is a major problem please let us know.

Location code translation

Where you are sending local holdings fields with codes that represent branch and/or collection location you will need to create a location code translation list for us, so we can convert your local codes to something meaningful in the Discover holdings display.

Looking closely at your data

You may find that as you start to look closely at your data you identify things you would like to improve before you share your records. Deciding to upgrade some of your data is likely to increase the time involved in catalogue supply, and you might instead like to make incremental upgrades and send us improved records over time as updates.

Doing the export

Once you've set up your data export you'll need a bit of time to test it to make sure it is working as you expect. The data export and transfer will vary, but for a large catalogue it may take a few hours of background activity.

Once the data is loaded

It is useful to build in a bit of time to check the data once it is loaded on Discover and Cataloguing.

Updating

The time commitment involved in supplying updates will vary between libraries. Some libraries may automate data supply, in which case there is little time involved, other than perhaps checking that the process has worked OK. In other libraries, once an export process is established it is likely to take a few minutes to run. This may be extended depending on the checks you might wish to do before transferring the update file(s).

In future, for some library systems we expect to be able to harvest updates.

Local changes

If you make any local changes that are likely to affect data supply or live circulation data then we need to know about these. Eg. a new Z39.50 server; a change to the IP address you are transferring data from; a change to a branch library name. These types of changes are unlikely to be a major overhead in terms of time, but you do need to tell us otherwise, e.g. you may find our live circulation data is no longer working for your library because your Z39.50 connection details have changed or we no longer have access through your firewall.

Data Supply questionnaire: Notes

The following notes provide some background information regarding the questions we are asking in our Data supply questionnaire.

1. Data transfer

1.1 What is the IP address (or range of addresses) of the machine you will be using to send the data?

We need to know your IP address to give you access to our system for supplying data files.

Most libraries supply data via sftp. However if you don't have a fixed IP address or you are using a Windows system we can also accept data via WebDav which provides an easy drag and drop mechanism for data supply. If you would like to try this option please let us know in 1.3 Comments & questions.

1.2 What is your SSH Public key?

We prefer to use your SSH public key to authenticate access. If this is a problem, please let us know in 1.3 Comments & questions.

Once the data supply is established, if you wish to change your SSH public key you can upload a new key through the Copac & NBK Contributor Dashboard. Please tell us once you have done so.

1.3 Comments & questions

Use this option for any comments or queries you have about the above questions or any other aspect of the data transfer process.

2. Your records

2.1 Approximate number of records you are sending

There is no maximum or minimum number of records for a data load, but it is helpful to know the size of a catalogue.

2.2 Record format:

We can work with a variety of record formats. If necessary we will convert your data to MARC21 so we have consistent data for deduplicating incoming records against the rest of the database. All libraries tend to have

some non-standard data so we try to standardise incoming records, primarily in terms of data format, to support the deduplication process. We normally only reject records completely where there is no title field.

2.3 Library system name

Knowing the library system can be helpful as we may have worked with data from the same system before.

2.4 Field containing your stable record ID

When we update records we match incoming updates against existing records on the basis of your local record ID. It is essential to have a unique stable record ID so that when you send us an update the correct records are changed or deleted.

If you are *not* able to supply a unique stable record ID please let us know as without this we can't reliably update the records on the NBK using the normal update procedures.

2.5 Please confirm your MARC organisation code

We include your MARC organisation code in the processed records; if you don't have a code this can be requested from the British Library:

<http://www.bl.uk/bibliographic/marcagency.html>

2.6 to 2.11 Holdings information

In the Discover records we include local holdings data in the form of branch location, sub-location, shelfmark and copy notes. We also try to include other local notes that you display on your own system eg. 59X local MARC fields. We treat content with a \$5 subfield as local and we include copy notes from the 852\$z, or your local equivalent, in the holdings display.

Having this local content means that we can supply basic location information and copy notes etc. in the Discover holdings display. For some libraries we are able to get live circulation data via Z39.50, but having the holdings data on Discover means you can use Discover as a back-up for occasions where your local system may be unavailable.

For example, the Discover display shown below is actually created from live circulation data, but we have most of the equivalent basic details on Discover in a 509 and 852:

```
509 || $a With an engraved portrait of Andrewes - one of 78 octavo copies printed. Bound in a volume lettered: Typographical Society's publications, 8. Biographical, I.
```

```
852 || $aUkLiU$b$ca$hSPEC H55.50(3)
```

Holdings information at the University of Liverpool Libraries for:

An exact narration of the life and death of ... Lancelot Andrewes, late Bishop of Winchester / [by Henry Isaacson; edited by J.T. Brockett].
Isaacson, Henry.
London : for John Stafford 1650?
Printed

| Location | Shelfmark | Availability/Copy Note |
|--|----------------|------------------------|
| Sydney Jones Library Special Collections/Archives (check locally for hours of opening) | SPEC H55.50(3) | USE IN LIBRARY |

Notes
With an engraved portrait of Andrewes - one of 78 octavo copies printed. Bound in a volume lettered: Typographical Society's publications, 8. Biographical, I.

2.12 Can you confirm the form in which your Institution/Library name should appear.

We hold the name of a contributing institution in multiple forms depending on where it appears within Discover and Compare.

So for the Library Information page, Holdings display and link names we need to know the full form of your institution/library name that you wish us to use. In certain circumstances we may need to use a shortened form of your name.

2.13 Please upload a CSV or Excel spreadsheet of your Location Codes.

Where your records contain location details in coded form then we need a translation table.

For example, here there is a single code for branch and shelving location name/collection in the 852 \$b:

| 852\$b code | Translation |
|-------------|---|
| jbir | <branch library name><shelving location name/collection > |
| jbp | <branch library name><shelving location name/collection > |

Whilst here there are two separate codes for branch and shelving location in the 852 \$b and \$c

| 852\$b\$c code | Translation |
|----------------|--|
| EDL BARC | <branch library name><shelving location name/collection> |
| EDL BCAR | <branch library name><shelving location name/collection> |

If you don't use the \$b\$c for location codes you can tell us the correct field/subfields to use in questions 2.6-2.8

Note:

The maximum file size is: 5MB

You can only upload the following formats: .csv, .xls, or .xlsx

2.14 'Contact...' name for any record without a location code

It is not unusual for some records to reach us that do not contain a location. Similarly, you may have locations you prefer us not to display. In this case we like to point the user towards the best source of information, generally the 'main library', and we would create 'Contact...' text along the lines of 'Contact the Main Library'. So we need you to confirm your preferred 'contact' text for this situation.

2.15 Comments & questions

Use this option for any comments or queries you have about the above questions or any other data related issues.

3. Timetable & Contact

3.1 Are there any dates in the next few months when you will be unable to supply your catalogue?

In terms of scheduling your data supply we will work round your local activity so that it isn't too much of an additional burden at a busy time. Similarly, if you are planning major data changes that are likely to affect a significant proportion of your data, it makes sense to do the load after the work is complete rather than to have to absorb large data changes shortly after the load.

3.2 Are you planning to move to a new library system in the next 12 months?

Within the NBK we rely on your stable local record ID to update your data. So any system change that results in a change to record IDs will require a reload, which is obviously work for both sides. So if you are planning a system move within a few months, where the record numbers will change, it is preferable to wait until that move is complete before we do the initial data load.

3.3 & 3.4 Technical & other contacts

We need contact details for the people who will be involved directly in the data supply to ensure good communication over the data load and subsequent updates. Contacts will be added to the NBK contributor mailing list. This is a Jiscmail list, with low levels of use which allows us to pass on service related information, e.g. new contributor loads or planned downtime, as well as providing us with a way of gathering feedback on specific issues.

3.5 Comments & questions

Use this option for any comments or queries you have about the above questions or any other timetable or contact related issues.

4. Updates

4.1 Please confirm your preferred update frequency

It is important that the records on the NBK be kept as up-to-date as possible, to ensure they are an accurate reflection of a library's holdings.

For an academic library we prefer weekly updates. For smaller institutions, with less fluid collections, less frequent updates may be adequate.

The only exception to the need for regular updates are those libraries with a static historic collection. These are generally small collections where a periodic reload, e.g. annually, is adequate.

4.2 In what form will you send the updates?

Updates need to include:

- » New records
- » Amended records
- » Deletions

We can accept updates in a variety of forms, but they must be consistent each month. New and updated records will be added to the NBK, with any pre-existing records being replaced by the updated record in the incoming file. Deletions are completely removed from the database.

- » If you send a single update file containing all record changes, any deletion records **must** be appropriately flagged in the leader.
- » Alternatively you can send two files:
 - one for new and amended records - labelled as 'new'
 - one just for deletions - labelled as 'deletions'.

In this case, as long as the deletions file is clearly labelled there is no necessity for indicating the deletion status in the leader of each record.

- » If you are supplying records in a non-MARC format we will need to confirm how to identify deletions - and a separate deletion file may be the best option.

Note: Any records flagged in the MARC Leader as deletions will be treated as such regardless of the file in which they are supplied.

4.3 Comments & questions

Use this option for any comments or queries you have about the above questions or any other timetable or contact related issues.

Submitting the questionnaire:

You do not have to complete the questionnaire all at one time. For example, if you don't have all the information you need at hand, you can use the 'Save' button to save the information you have entered so far and come back to the page later to complete it.

NOTE: The NBK team will **not** act on any information you have supplied until you select the 'Form complete' option at the end then 'Save' the form again. If there are things you are unsure about just include questions for us in the relevant sections of the form and select 'Form complete' then 'Save'. We will follow up with you any questions you may have when we review the form.

Next steps

Once we have received your questionnaire we will review the information and be in touch regarding any questions you may have included, or with follow-up questions of our own.

Data Management questionnaire: Notes

The following notes provide some background information regarding the questions we are asking in our Data supply questionnaire.

Introduction:

The NBK is created from data supplied by contributing institutions. If your institution will be contributing data it is important that we understand how you want us to manage your data. Specifically we need to know about your requirements relating to data exclusions and data sharing. These are actions we will incorporate into the pre-processing and standardisation activity that we carry out before loading your data. We may not be able to cater for every requirement but if there are any problems we will confirm our proposed actions with you.

1. Data pre-processing

1.1 Excluding records

We expect you to exclude from your export any bibliographic records or item branch locations that you don't wish to be included in the NBK. However, we know that this can sometimes be a problem so, if necessary, we can drop identifiable records or branch locations from your data before loading. Dropped records are removed from your data set completely and will not be available for use within Library Hub Discover, Compare, or Cataloguing.

If this is required, let us know how we identify the relevant locations/records that you need us to drop before we load your data.

If dropping branch locations leaves a record with no physical or electronic location, unless you tell us otherwise, we would then drop the record. If we drop records these will be reported in a rejects file.

1.2. Records with no locations

If we find records in your catalogue with no physical location (MARC 852 field or equivalent) AND no electronic location (MARC 856 field) we need to know if we should keep or drop these records before loading your data.

If we drop records these will be reported in a rejects file.

1.3 Rejects and warnings

From late 2019, when we pre-process MARC21 data we will create a rejects file for any records dropped, as well as a warnings file for potential record errors. **Note:** We do not supply rejects and warnings files for non-MARC data.

Many libraries find these files useful for identifying data problems which may be less easy to see locally. So we would like to confirm whether you wish to have access to these files. You can change your decision at a later date.

This is an area we will be developing in future as we explore mechanisms for supporting data quality enhancement.

1.4 Any general comments on Data pre-processing:

Use this option for any comments or queries you have about the above questions or any other aspect of the data pre-processing.

2. Data sharing

Within the NBK we want to make data available as widely as possible for resource discovery, collection management, and shared cataloguing activity, so we will be sharing data in both MARC (Cataloguing) and non-MARC formats (Discover and Compare). We encourage you to make your data available for sharing in MARC21 format through the Cataloguing service, but we are aware not everyone may wish/be able to do this, or that there may be sub-sets of your data where sharing in MARC21 format is not possible.

Many data suppliers have told us that they do not want their data to be available in MARC format, and we have global checks in place for this data so that it may be used for Discover and Compare, but not Cataloguing or WorldCat sync. Please see the NBK Data Sharing FAQs document for more information.

We want to ensure that you are comfortable with where your data flows within and out of the NBK.

2.1. Shared Cataloguing

The Library Hub Cataloguing service is created from data supplied by NBK contributors. We want to confirm whether your records can be included in this shared cataloguing service or whether you have records you wish to suppress. Suppressed records will be available for Discover and Compare in non-MARC format. (See the NBK Contributor Guide document for more information).

If you have sub-sets of data you wish us to suppress from Cataloguing we need to know how to identify those records. We are also interested in why you prefer not to share the specified records.

2.2. Data exposure to aggregator services: OCLC WorldCat

We will be making the NBK data available to data aggregators in MARC format (see the NBK Contributor Guide for more information). In the first instance this will be through data synchronisation with WorldCat, allowing records to be automatically incorporated into the WorldCat database. This option is available only to Jisc members who send data in MARC format. Only data that is allowed to be shared through Cataloguing will be synced to WorldCat.

NB data which is synced with WorldCat becomes subject to the WorldCat data agreement [<https://www.oclc.org/en/worldcat/cooperative-quality/policy.html>]

If you currently send data to WorldCat, once data synchronisation with WorldCat is established you will be able to manage this data flow through the NBK if you wish, removing the need for a separate data supply to OCLC.

We need to know whether you wish any or all of your data to be synced with WorldCat. If you wish us to suppress just some of your data from WorldCat we need to know how to identify the relevant records. If you wish us to suppress some or all of your data from data synchronisation with WorldCat we are interested to know your reasons for this.

2.3 Data exposure to aggregator services: other aggregators

As well as WorldCat, it is likely that other data aggregators will also wish to make use of the NBK data in future. We need to know if you are happy for your data to be shared with other data aggregators. **Note:** data made available to another aggregator is likely to be subject to that aggregator's data agreement.

We are interested to know whether you wish any or all of your data to be made available to other data aggregators in future. If this is something you would wish to consider depending on the nature of the aggregator, just let us know, we will consult you as/when the need arises.

If you definitely wish us to suppress some or all of your data from data synchronisation with other aggregators we are interested to know your reasons for this.

2.4 Any general comments on Data sharing: [text box]

Use this option for any comments or queries you have about the above questions or any other aspect of data sharing.

3. Non-MARC data exposure

We will be making your data available in non-MARC format (eg. MODS XML) for resource discovery, collection management services, and other NBK supported data exposure and data share.

We are interested to know of any questions or concerns you may have relating to data visibility and data sharing in non-MARC formats.

4. Additional Comments or questions:

Use this option for any comments or queries you have about the Jisc NBK development.

Submitting the questionnaire:

You do not have to complete the questionnaire all at one time. For example, if you don't have all the information you need at hand, you can use the 'Save' button to save the information you have entered so far and come back to the page later to complete it.

NOTE: The NBK team will **not** act on any information you have supplied until you select the 'Form complete' option at the end then 'Save' the form again. If there are things you are unsure about just include questions for us in the relevant sections of the form and select 'Form complete' then 'Save'. We will follow up with you any questions you may have when we review the form.

Next steps

Once we have received your questionnaire we will review the information and be in touch regarding any questions you may have included, or with follow-up questions of our own.

Library Information questionnaire: Notes

The following notes provide some background information regarding the questions we are asking in our Library information questionnaire.

1. Information page content

For each NBK contributor we create a Library information page on Discover so that a user is able to find basic information about your library. Each Discover holdings display links directly to the relevant Library Information page. You can see all the existing pages at <https://discover.libraryhub.jisc.ac.uk/about/libraries/>

1.1 – 1.2 Contact details & Library description:

It is helpful for researchers and others to have basic information about your library, including contact details.

1.3 Photograph:

We like to include a photograph, either of your library or part of your collection. Please upload your preferred image:

- » Image size: the maximum image size is 5MB.
- » Format: the image format needs to be .jpg, .png, or .gif

1.4 & 1.5 Inter-Library Loans

Library staff using Discover are keen to have access to information regarding Inter-Library Loans (ILL). So if some of your materials are available for ILL it is much appreciated if you provide your ILL library and policy codes. If you have details about your ILL policy/procedures on your website then this can also be very useful for both end-users and library staff.

1.6 & 1.7 Visitor information

We encourage potential visitors to check in advance regarding visiting arrangements and access to materials. It is helpful to be able to provide a link to relevant information for visitors on your web site. Providing your address and postcode means we can include a location map for your library.

2. Contact details

We are keen to promote contributor collections through Library Hub and one way of doing this is to invite libraries to write guest posts for our blog, for example, highlighting specific collections or forthcoming exhibitions. In order to do this we need contact with those in your library most likely to be interested in collection

promotion, eg. Special Collections staff. Contacts will be added to the NBK contributor mailing list. This is a Jiscmail list, with low levels of use which allows us to pass on service related information, e.g. new contributor loads or planned downtime, as well as providing us with a way of gathering feedback on specific issues.

3 Comments & questions

Use this option for any comments or queries you have about the above questions or any other timetable or contact related issues.

Submitting the questionnaire:

You do not have to complete the questionnaire all at one time. For example, if you don't have all the information you need at hand, you can use the 'Save' button to save the information you have entered so far and come back to the page later to complete it.

NOTE: The Copac & NBK team will **not** act on any information you have supplied until you select the 'Form complete' option at the end then 'Save' the form again. If there are things you are unsure about just include questions for us in the relevant sections of the form and select 'Form complete' then 'Save'. We will follow up with you any questions you may have when we review the form.

Next steps

Once we have received your questionnaire we will review the information and be in touch regarding any questions you may have included, or with follow-up questions of our own.

Circulation Data questionnaire: Notes

The following notes provide some background information regarding the questions we are asking in our Circulation data questionnaire.

1. Does your catalogue support access to circulation data via Z39.50?

Where possible we like to provide live circulation data for Discover users, showing the current availability details for materials in your collection. We do this using a Z39.50 link to your catalogue – when a user selects the Discover holdings display for one of your documents we send a record ID search to your local catalogue to pull back the current availability details and we include this in the holdings display.

Where your catalogue does not support this type of circulation data access the holdings display will show the local data included in the records you have sent to us. Similarly, where the Z39.50 link is temporarily unavailable, e.g. because of a network problem or local system downtime, we will display the location details we hold centrally. This means that when your local catalogue is unavailable because of system maintenance Discover can be used to provide details of your collections as well as basic location information.

2. If you are able to support Z39.50 access please provide:

If your system does support Z39.50 access we will need the following information to establish a search link to your local catalogue. We will test the link and will be in touch if we have any problems establishing access. You may need to give us access through your local firewall for this to work.

- » Z39.50 server IP address: e.g. z3950.jisc.ac.uk
- » Port number: e.g. 210
- » Database name: e.g. LIBRARY

3. Comments & questions:

Use this option for any comments or queries you have about the above questions or any other timetable or contact related issues.

Submitting the questionnaire:

You do not have to complete the questionnaire all at one time. For example, if you don't have all the information you need at hand, you can use the 'Save' button to save the information you have entered so far and come back to the page later to complete it.

NOTE: The NBK team will **not** act on any information you have supplied until you select the 'Form complete' option at the end then 'Save' the form. If there are things you are unsure about just include questions for us in the relevant sections of the form and select 'Form complete' then 'Save'. We will follow up with you any questions you may have when we review the form.

Next steps

Once we have received your questionnaire we will review the information and be in touch regarding any questions you may have included, or with follow-up questions of our own.