5

BIBLIOGRAPHIC AND ITEM DATA

Kelley McGrath, Metadata Management Librarian, University of Oregon

M igrating data to a new ILS is a massive undertaking during which you will inevitably encounter problems with bad data and data mapping. In this chapter I discuss common pitfalls to be aware of when preparing bibliographic and item data for migration and suggest methods for mitigating potential problems.

SYSTEM ARCHITECTURE

One of the most difficult, but most important, tasks for a successful migration is learning as much as you can about the structure of the system that you're migrating to and how its parts work together. Understanding its workflows and how they use different pieces of information, as well as what data points affect the discovery interface, is key for making migration-related decisions. These differences can be at a very fundamental level, such as the way different types of records are related to each other. In some systems, both holdings and item records are attached directly to bibliographic records, whereas in others items are attached to holdings records, which are then connected to bibliographic records (figure 5.1).

FIGURE 5.1

Bibliographic, holdings, and item records structure
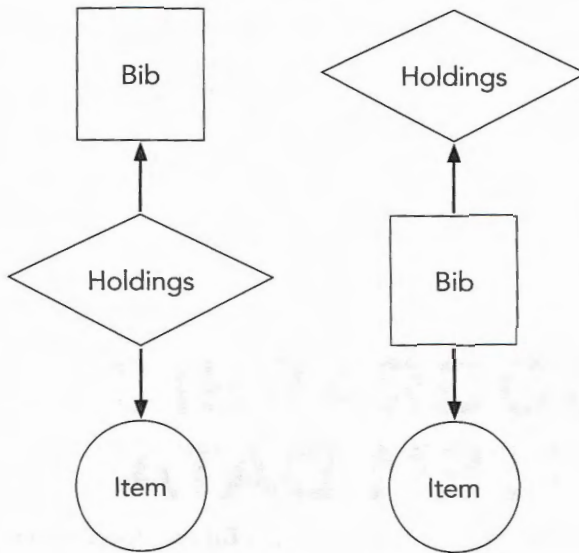
If you are migrating from a system where items are attached to bibliographic records to one where they are attached to holdings, there is a potential for problems connecting the migrated items to the correct holdings records. In addition, assumptions that work in your current system may not hold in the new one. For example, your current system may not require item records in situations where your new system does, such as when receiving new items on continuing resource records or for proper display to the public. It may be easier to add these records in your current system prior to migration. An additional benefit of making these kinds of adjustments pre-migration is that things will immediately function correctly in your new ILS.

One of the most difficult, but most important, tasks for a successful migration is learning all you can about the structure of the system that you're migrating to and how its parts work together.

It is also important to be cognizant of the structure of your current ILS. It is easy to assume that the way things work in your current ILS is just the way things work. However, your new system may have been constructed with a different underlying model. It is also important to be aware of this dynamic when communicating with your new vendor. Vendors are also prone to think their worldview is the only one. The staff supporting migrations don't necessarily have much knowledge about how their competitors' systems work. When migrating to

a system from another vendor or to one structured differently than your current system, contact institutions that have made the same transition to find out what surprised them or those to which they had to adapt.

> **If you are migrating to a system built by a different vendor or that is structured differently than your current system, contact institutions that have made the same transition to find out what things surprised them or those to which they had to adapt.**

BIBLIOGRAPHIC DATA QUALITY

Moving data from one system to another exposes data-quality problems, such as missing, malformed, or incorrect data. For a smoother migration, identify, and if possible, remedy these problems pre-migration. At the most basic level, verify that your records are in structurally valid MARC. This can be done by exporting your records and using an external program, such as MarcEdit or MARC Report. Corrupt records can cause problems with both export from your current system and import into the new one.

MARC records can be structurally correct while still violating basic constraints of the format. Check for problems, such as:

- Records with more than one instance of non-repeatable fields, such as 100 or 245.
  - » 245 00 ‡ a Rashōmon
  - » 245 00 ‡ a 羅生門
  - » (the non-Roman title should have been in a linked 880 field and not a second 245)

- Records where explicit subfield markers have been entered in a field, such as 001 or 007, that should not have subfield markers in the underlying MARC record
  - » 007 ‡ a c ‡ b r ‡ d c
  - » (underlying MARC should have 007 cr\c with no subfielding)

- Records where nonstandard fill characters have been used
  - » =008 080326p20081951xx■083■g■■■■■■■■■■■■vljpn■d

Look at the quality of coded data in fixed fields, which consist of characters that derive their meaning from their position in the string that makes up the field (see the 007 and 008 fields in the list above for examples of fixed fields). Many newer discovery interfaces use this information in facets, thereby exposing missing or badly coded data. This coded data may also be used in the migration process. For

example, rather than mapping our existing item material types to item material types in the new system, our vendor generated material types based on the coded value for the record type (e.g., j for musical sound recordings) in the related MARC bibliographic records. In the past, fixed fields were often neglected, and many records retrospectively converted from catalog cards fail to code most of them. Your system may have internal non-MARC codes for things like format that are more obvious to people working in the system than the actual MARC coding. Even though these internal fields may be correct, it is the underlying MARC that will migrate. Check to see that these two values agree with each other and resolve discrepancies.

Be aware of diacritics and character-encoding issues. Know whether your records export in MARC-8, UTF-8, or some mixture as well as what kind of imports your new system accepts. Look for corrupt diacritics and be sure to verify that diacritics have transferred correctly during your test load. See chapters 2 and 4 for more help with character encodings.

Consider the option of sending your records to a vendor rather than under-taking bibliographic data cleanup in-house. As an alternative to having staff clean the data, consider having your vendor do it. Vendors have extensive experience with data cleanup and will have more sophisticated, comprehensive, and auto-mated processes for identifying and correcting bad data than you will be able to develop on your own for a one-off project. If your local records will be loaded into your new system, rather than just matched against existing records, you may also want to think about combining RDA enhancement or authority control updating services with your data cleanup project. Vendors may also be able to address some deficiencies in fixed fields and update records to current standards (e.g., update the 008 form byte to use for online resources).

> **Consider the option of sending your records to a vendor rather than undertaking bibliographic data cleanup in-house.**

SYSTEM CONTROL NUMBERS

System control numbers play a critical role during migrations. Record numbers from bibliographic utilities play an important role in record sharing and are often used for record matching. The OCLC consortium and OCLC numbers play a unique role in library systems, and OCLC numbers have become an important identifier in many shared library systems and projects. Because these issues are so common, this chapter discusses OCLC numbers, but the issues are applicable to other kinds of control numbers.

If you are migrating to OCLC's WorldShare or to a shared consortial catalog where OCLC numbers are used as the primary match point, it is important to have accurate OCLC numbers in all your bibliographic records. Many things can go wrong with control numbers, including records with leading zeros stripped off, overlapping control numbers that are used for different purposes, as well as ones that are just plain wrong. Verify that your vendor uses a normalization process when matching control numbers or else make sure that all your control numbers are formatted the same way.

If you haven't looked before, you may be surprised by how many of your bibliographic records lack numbers corresponding to a bibliographic utility such as OCLC. If there are many records like this and you need OCLC numbers for your migration, the most efficient solution is to undertake a reclamation project by sending your bibliographic records to OCLC. In addition to synchronizing your holdings, a reclamation project can match your numberless records with existing OCLC records and add new records to WorldCat when a match is not found. You will then receive a copy of your records with the new numbers added that you can load into your local system. OCLC will perform reclamation for a library once for free, but after that there is a fee. If your records already have OCLC numbers and you only want to check the accuracy of your holdings, you could try the method described by Johnston (2015) (http://journal.code4lib .org/articles/10328). You may have brief records that include unverified OCLC numbers that you do not wish to use for matching purposes. Consider removing these numbers before export.

Export your ILS's internal record identifier and the field in your records that contains the OCLC number to an external file. Use this file to identify problems with records with more than one valid OCLC number (e.g., a record with two 001 or two 035$a fields, each containing a different OCLC number). For example, you might have a single record that includes both of these 035 fields:

```
035 $a(OCoLC)681501143
035 $a(OCoLC)756452689 $z 904378234
```

The only way to resolve these is to manually review the records to determine the correct number. Information in chapters 2 and 3 may help you identify and fix these problems.

Use this same file to find instances where more than one record has been assigned the same OCLC number. These also need to be examined individually and compared with the master record to determine the correct course of action. Do the records need to be merged or should one of the OCLC numbers be changed? This can be complicated if a library has locally edited master records for different editions.

ILS SYSTEM CONTROL NUMBER	OCLC NUMBER
1	681501143
1	756452689
2	681501143

The table shows that record 1 has two different OCLC numbers and that the same OCLC number is associated with both record 1 and record 2.

In addition to obvious duplicates, there are many cases where OCLC has merged records that the institution downloaded separately when they were two different records. These are more difficult to identify. If you are an OCLC member, it can be done by using WorldShare Collection Manager to create a local collection with all your holdings by searching for records that are held by your OCLC symbols. You can use the li command for this (li:oru). Request a download of these records and use a program like MarcEdit to extract all the 035 $a and $z fields from these files. Split out all the 035$z numbers and pair each one with the $a number from the same record. Make a list with all the $a/$z pairs as well as a $a/$a pair for each record. Export all the valid OCLC numbers (001 or 035$a, but not 035$z) from your ILS. Normalize all the OCLC numbers by removing "(OCoLC)," "ocm," and "ocn" and deleting any leading zeros. Match the OCLC numbers from your ILS against the second number in the giant list of pairs of 035 fields that you downloaded from WorldCat. If any 035$a from the downloaded file matches more than one OCLC number from your local file, you have two separate records that have been merged into one in WorldCat. You may wish to verify that the merge is correct and either merge your own records or change the OCLC number on one of them.

WORLDCAT 035$A	WORLDCAT 035$Z
1000	3000
2000	

035 $a and az exported from WorldCat.

WORLDCAT 035$A	WORLDCAT 035$A/035$Z
1000	1000
1000	3000
2000	2000

WorldCat 035$a numbers associated with each 035$a and 035$z from the same record.

BIBNO IN ILS	OCLCNO IN ILS	WORLDCAT 035$A/035$Z	WORLDCAT 035$A
Bib001	1000	1000	1000
Bib002	2000	2000	2000
Bib003	3000	3000	1000

Bib001 and Bib003 have different OCLC numbers in the local ILS, but because they link to the same 035$a in the WorldCat record, the two records were merged in OCLC sometime after they were downloaded to the local ILS.

If a system number that is being used for matching is in the 001 field and your new system takes the corresponding 003 field into account when matching, verify that the 003 field is present and correctly formatted. For OCLC numbers, this value is OCoLC. Any other system numbers (e.g., YBP numbers or MARCIVE shipping list numbers) that are used as match points for loading records should migrate such that they can still be used in the new system.

> You may have other system numbers that are used as match points for loading records, such as YBP numbers or MARCIVE shipping list numbers. Make sure these migrate in such a way that they can still be used as match points in your new system.

LOCAL DATA AND LOCAL EDITS

It is important to determine what will happen to any local data during migration or in your new system. For example, a local field that can be used to trigger reminders for actions in one ILS may do nothing in another.

One challenge facing institutions migrating to a shared catalog where their original records will be matched to a master record and will not be retained is the potential loss of many years of local data enhancements. There are several scenarios where this becomes a problem.

Local Record Has a Bibliographic Utility Number That Describes a Different Resource

Especially in the early days, many catalogers used OCLC master records as a base for a description of a different edition in the same manner they used to modify purchased catalog cards. Catalogers would download a near-match from Connexion and modify it locally to describe the item that they actually had. In the worst-case scenario, they downloaded any record on the same record format to modify locally. This means that the local record is now linked to an inappropriate

OCLC record. As well, in some cases there may be two records with the same OCLC number in the catalog that do not describe the same thing because one or both have been locally edited. There is no easy way to identify these.

Local Edits to Enhance or Correct Master Records

Even when using a matching record, libraries have often corrected errors in records or added additional information, such as contents notes, only in their own local catalog records. In the past, it was not possible for many institutions to make these changes in the OCLC master record, and even now, when OCLC has significantly expanded the types of users who can edit the master record, some institutions continue to make only local edits. It is not easy to identify these. You may also have information in your local bibliographic record, such as tables of contents that cannot be added to the master record for licensing reasons. This type of data must be handled using the methods described below for true local data.

Institution-Specific Local Data

A third category is truly local data. If you are migrating to a system that uses a master record approach, the system should provide some way to record local data. You should be able to migrate your existing local data to these new local fields if you can identify and mark it in some way. The challenge is identifying it. Data that really is only locally applicable include the names of special collections or information describing characteristics of the item that only apply to the library's copy. It also includes things like local headings for theses and advisor's names, which are not allowed in WorldCat master records. In the past, it may not have been considered important to identify this local information in your local catalog by using 590 notes rather than 500, or by using subfield 5 with your institution's MARC organization code. In some cases, local practices have been documented, making the related local fields easier to identify. You may be able to use information about the local collection to come up with likely keywords that can be used to find local data. You may also wish to export notes and headings fields from the records in collections that are likely to contain local notes and analyze them externally. For example, notes for review can be selected using keywords that often represent local information, such as "copy," "sign*," "donat*" or the name of a library. Looking at the initial list of local notes will likely suggest additional keywords to search. This review can also be accomplished with direct searches in the ILS, but an external list prevents redundant review. Another strategy is to export the bibliographic records for the items you hold from WorldCat (using

the technique described in the earlier section on system control numbers) and compare the number of headings or notes in the current master record with the number in your local records. Focus on those record sets that are likely to contain local or value-added data. In cases where your local record has more headings or notes, there is additional information in your local record that should either be marked as true local data or added to the master record. OCLC's batch loading process may be able to add some data, such as contents notes and summaries, to master records that do not already contain those fields. Chapters 3 and 4 include information that will help with analyzing your data and marking local information. However, all methods for retrospectively identifying and retaining local data are labor-intensive and unlikely to be complete.

Whether you are migrating to a system that uses a master record rather than your local data or not, you may also find that you have some nonstandard practices or uses of fields that improved indexing, display, or functionality in your old system, but that cause problems or don't work in your new system.

ONLINE RESOURCES

You may need to identify all your records for online resources separately so that the URLs will migrate properly. Now is a good time to verify that all your online records include links and that those links work. In some cases, the new system includes a knowledge base of information about online resources, such as Alma's Community Zone. Consider whether it will be easier to re-create your online holdings using the new vendor's system or to try to convert your existing data. If you are currently managing some records in a stand-alone ERM, there may be some duplication between records in your ILS and in your ERM. If you are migrating to a system with an integrated ERM, you will want to identify duplicate records and remove the extra copies from the data being migrated.

In the past, it was common to combine print and electronic holdings on a single record. It is now more common to use separate records for maintenance purposes. You should investigate the effects of your existing practice in the new system. Does your new system require separate records for accurate faceting and limiting in the discovery interface? If the new system relies on information in the bibliographic record to distinguish print and e-resources, the single record approach will provide incorrect information about one or the other format. In this case, you may wish to split the records before your migration. Separate print and electronic records can be easier to maintain over time, but the single-record approach may result in a more user-friendly display. It is also possible that your new discovery interface will merge the print and electronic records for display even if you have separate records on the back end. See chapter 9 for more information about migrating electronic resources.

OTHER CLEANUP

Think about whether you have records in your current catalog that it would be better not to move to your new ILS, such as brief records for equipment, reserves, or other records created for circulation purposes, and delete them if necessary. Look for anomalous situations where record statuses are out of sync or related records are missing. It is a good idea to:

- Identify brief records and consider upgrading or deleting them.
- Check for unsuppressed bibliographic records that have no associated holdings and should be suppressed or deleted.
- Check for suppressed bibliographic records that do contain items and should display to the public.
- Look for bibliographic records with no holdings or other associated records. If you cannot figure out what they're for, you may want to delete these.

It is often easier to do cleanup in your existing system where you are familiar with the search options and tools for batch editing. In addition, once your new system goes live, you will be very busy developing and implementing new workflows and are unlikely to have time to focus on data cleanup. Take advantage of any error-checking reports offered by your current vendor or export your records and examine them with an external program. If you have a small collection and database maintenance has been neglected, taking an inventory to identify items to remove may be helpful

ITEM RECORDS

MARC bibliographic and holdings records use a standard format and, in theory, should be able to migrate losslessly, although some systems store holdings records in nonstandard ways that don't support the export of accurate MARC. Systems also vary in the degree to which they respect some parts of the MARC standard, for example, constraints on the length of records and number of allowable fields. However, where there is no standard format, such as for item records, data loss is almost inevitable. There are often similar fields that are defined differently from system to system. Fields may be present in one system but not in another. There may also be data that you can't export from your old system (assuming you want to). Although you would expect any ILS to be able to do a complete data dump of the library's information, unfortunately this is not possible in all systems. This is especially a problem for item records or non-MARC fields associated with bibliographic records.

Make sure that you know what information you can export and what you can import into your new system. One library found that an item record field in its old ILS that it used for managing an off-site storage facility would not be

imported into its new system. Staff had to find a work-around so that the 100,000 items in the storage facility would not need to be re-inventoried.

Find out what fields the item record in your new system contains and determine how to map your existing data to values that will produce the functionality you need in the new system. Migration presents a good opportunity to rethink and possibly simplify your circulation policies, location codes, and material types. Find out how the fields in the item records function in your new system. Ask questions, such as:

- Is there a hard-coded list of values or does each library create its own?
- What processes use a given field?
- Where and when does each field display?

For example, in our former system, material type influenced circulation policy, whereas in our new one it is only used for display and statistics.

There are several reasons why it might be difficult for you to migrate data in item records that you'd like to keep. There may be no corresponding field in the new system. Alternatively, there may be a field that exists in both systems, but can't be mapped for some reason. For example, the field could be free text in your old system, but the new system uses a controlled vocabulary. Sometimes data that has no equivalent in the new system, such as loan history, can be mapped to note fields where it can still be searched or used for statistics. In some cases, your only option may be to export and store this data in a separate file.

Some item record fields have lists of controlled terms associated with them. Common examples include status, location, and material type. The available values for the equivalent fields in your new ILS may be different. You may also wish to change the list of values that you use for a field if that field functions differently in your new system or because you have rethought your workflows. In these cases, you will need to map your existing values to new values that will work in the ILS to which you are migrating. Whether mapping or migrating codes as-is, be sure that they conform to the new system's requirements in terms of length and presence or absence of non-alphanumeric characters and spaces.

If an item record field is supposed to be limited to a controlled list of values, search for null values in your current system and, if they are allowed by your system incorrect values that are not on the list of valid terms. This will prevent erroneous or unusable data from migrating.

Fields that affect loan policies vary from system to system. In some systems, a variety of fields may combine to determine the conditions under which an item circulates. In others, circulation may be dependent on a single variable such as a location code or a lending policy. Systems may also have ways to override the default value, which should be taken into account. Systems also mark statuses and processes differently.

Export data from various item record fields and look for outliers that might represent miscoded data, such as barcodes in volume fields. If your item records include "variable" fields that are repeatable (such as v for volume), check to see if any of these are repeated in an item where they shouldn't be.

Export a list of all your barcodes and identify any duplicates for cleanup. You can also use this list to identify incorrectly formatted barcodes, barcodes with the wrong number of characters, or barcodes that begin with the wrong characters. Chapters 2 and 3 address this kind of cleanup.

You may have many items in a temporary status, such as in-process or at the bindery, which should have been removed from that status long ago but were somehow missed. Resolve as many of these as possible prior to migration. You may have lost or withdrawn items that are being kept for historical reasons. Now is a good time to export these to external files if they must be kept and then delete them from your database rather than migrating them. If your new system requires items be attached to bibliographic or holdings records for some functions, identify records that lack the items, and add them.

BOUNDWITH RECORDS

In the past, it was common for libraries to bind a group of pamphlets or other separate publications into a single volume that is circulated as a unit. However, to provide access to the contents, bibliographic records were created for each of the individual publications. These individual bibliographic records must somehow be linked to the single barcode and item record information for the bound volume. These record groupings are commonly known as "boundwith records" and can be tricky to migrate. Systems handle boundwith records differently. Some systems link a single item record to multiple bibliographic records, but others link items to many holdings records or chain item records together (figure 5.2).

If your catalog includes boundwith records, find out early on how your new system handles these and how the vendor plans to migrate them. Test the completeness and functionality of migrated boundwith records thoroughly. Check whether the export process is associating the correct call number with boundwith items. We often only had call numbers in the individual bibliographic records for each part of the boundwith, so we had to add a call number to the item record that represented the range of volumes present in the boundwith. Otherwise the export process would just pick up the call number from the first bibliographic record and interpret it as the call number for the whole resource.

Bibliographic records for boundwith resources can be linked by shared holdings or item records on linked item records
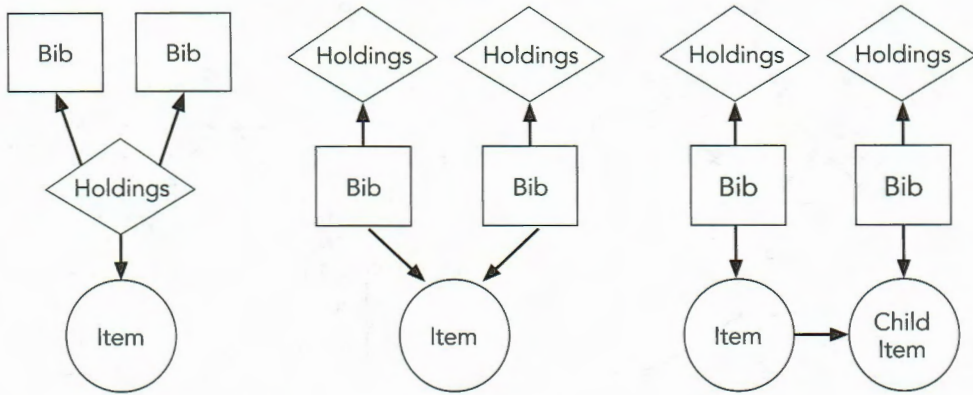
CALL NUMBERS

It is extremely important to migrate call numbers correctly or you'll never find anything again. Different systems handle call numbers in different ways. Call numbers associated with an item may be stored only in item records, in bibliographic records with an item-level override, or in holdings records with an item-level override (figure 5.3).

Call numbers stored only in items, in bibliographic records with an item-level override and in holdings with an item-level override
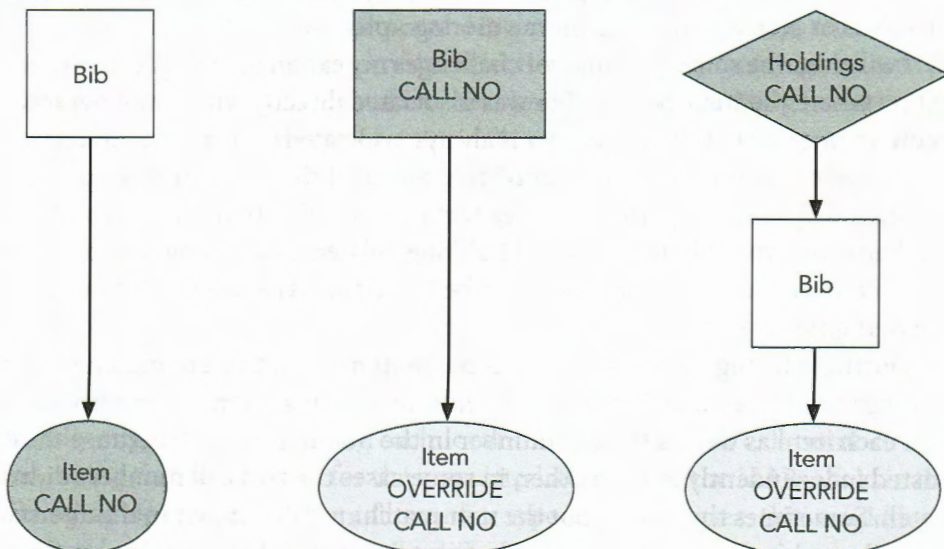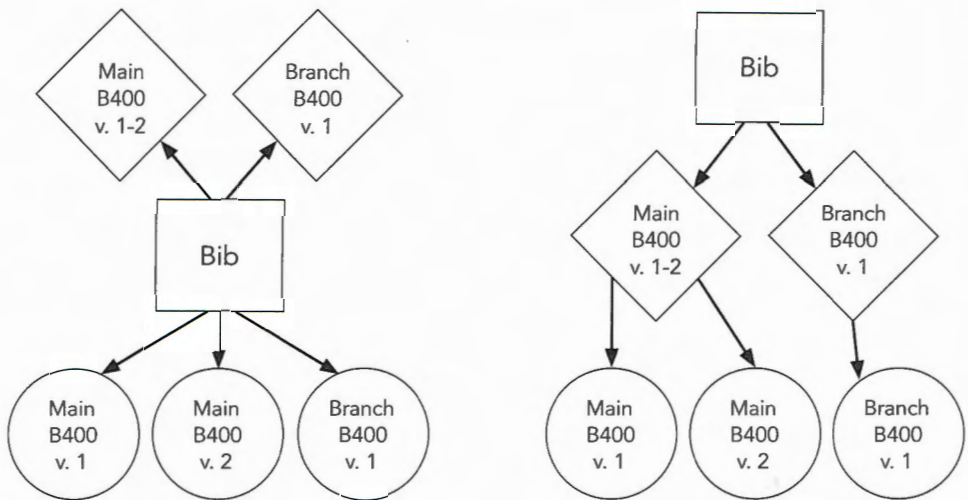
FIGURE 5.4

Item and holdings call numbers that are linked directly to bibliographic records may need to be linked to each other*

*As shown on the right of the diagram.

When migrating from one model to another, there is a potential for mismatches. Migrating from a system where holdings and items are linked to the bibliographic record independently to one where items are linked to the bibliographic record via the holdings record can cause problems. A MARC holdings record is designed to record a summary of the specific volumes held by a library, along with their associated location and call number. Generally, a separate record is made for each location and call number. Libraries where the call number is stored in the item or the bibliographic record may only create holdings records for serials or serials and multivolume monographs.

I will describe some examples of challenges my experience migrating from a system where the item call number was associated directly with a bibliographic record to one where the call number is always associated with a holdings record. When there is not a holdings record related to a bib, the migration process automatically created a brief holdings record using the item location and call number. However, if both items and holdings related to a bibliographic record exist in the old system, the items should be associated with the correct holdings record (figure 5.4).

During our migration, in most cases the item records were matched to a holdings record based on location. We sent the vendor a call number associated with each item as well as the call number in the holdings record. Because these existed independently of each other, in some cases the two call numbers didn't match. Sometimes the vendor put the non-matching call number in the alternative call number field in the item. Unfortunately, the vendor's matching process

did not normalize the call numbers before comparing them, so we ended up with tens of thousands of unnecessary alternative call numbers in item records because of differences in spacing or in the use of periods. It is very important to make sure that your vendor is normalizing any values, such as call numbers or system numbers, that are used for matching. If your new vendor will not properly normalize such data, try to normalize as much as possible yourself. Normalization will take data that has been input in different formats and put them into a standard format. Although a human can easily see that the call numbers in the first column in the table below are all meant to be the same number, the literal-minded computer will look at them character-by-character and see three different strings. Removing all the spaces and periods before the comparison creates strings that even a computer can match, as shown in the second column. See chapter 2 for help on normalizing data.

CALL NUMBER ENTERED IN ILS	NORMALIZED CALL NUMBER
Z678.93.I57 B6	Z67893I57B6
Z678.93 .I57 .B6	Z67893I57B6
Z678.93.I57B6	Z67893I57B6

In other cases, we ended up with two holdings records: the migrated holdings record with no items attached and a system-created holdings record with a different call number and the migrated items attached. Although we already had call numbers in our holdings records, some libraries migrating to a system that is looking for call numbers in holdings records have had to use a script to add them.

In a system where there is no direct link between holdings and item records, it is possible to list the volumes in multiple locations, such as general stacks and current periodicals, in a single holdings record. However, when the holdings record is the link between a bibliographic record and its items, then each holdings record can contain only one call number and one location. If you are migrating from the first scenario to the second, you will want to split your holdings records pre-migration.

Locations in holdings and items can also get out of sync if one is updated and the other isn't. Look for holdings records with no items in the same location. Check for items where you expect a matching holdings record in the same location, but one doesn't exist.

In our previous system, call numbers were often stored in the bibliographic record. The field being used to generate the call number for the associated items was marked as such with a non-MARC tag that wouldn't export. Our new vendor initially suggested associating each location with a specific call number type and taking the call number for the associated items from the MARC field in the bibliographic record for that call number type. In many cases, this would have resulted in an item being assigned a different call number than what was on the

piece. This would have been disastrous. However, as you are migrating your call numbers, verify that they are being mapped in a way that item records remain associated with their correct call numbers.

We ended up exporting all our call numbers as part of the item record. However, our old ILS could not output the subfielding that delineated the class and Cutter numbers in our call numbers nor could it export the tag that identified the type of call number (i.e., LC or Dewey). We therefore lost all our subfielding. To assign a call number type in our new system, we had to associate each of our locations with a specific call number type and clean up the exceptions post-migration. This is an example of data that exist in the current system and have an equivalent in the new system, but cannot be migrated because of the export limitations of the current system. Make sure you understand how information about call number types (e.g., LC or Dewey) is being migrated.

Our volume designation information was stored in a single item record field in the form "v. 2, pt. 1." However, our new system expects each part designation to be entered separately, which prevents an easy mapping. If your volume designation cannot be separated from your call number information, but your new system expects them to be separate, this is also a problem. These are examples of a mismatch in data format between two systems, which is always a problem when moving from a less granular system to a more granular system.

We used "input stamps" at the beginning of call numbers in our old ILS to indicate information like oversize. These displayed, but were not indexed and did not affect call number sorting. Because our call numbers were migrating as un-subfielded strings, these input stamps would have become part of the actual call number in our new system. We moved this text to the end of the call number pre-migration and slated it for post-migration cleanup.

Finally, check to see if you have records with empty call number fields and identify the appropriate call number where possible. Identify improperly constructed call numbers or call numbers associated with the wrong call number type and fix them.

These are some examples of the many things that can go wrong with call number migration. Be sure to ask:

- Is the link between items and their call numbers being properly maintained during the migration process?
- How is call number type being associated with call numbers?
- Do you have subfielding between the classification and shelflisting parts of call numbers and, if so, can you migrate this information? Will this affect shelflisting in your new system?
- If you have call number prefixes, do you know how these will migrate? How will they be indexed in your new ILS and in your discovery system?
- Are call numbers being properly normalized if they are being used for matching purposes, such as matching item and holdings records?

- If item call numbers are being linked to existing holdings records during migration, what information is being used to make the connection—location alone or both location and call number?
- Do you have call numbers in holdings records if your vendor requires them?
- Do you have holdings records that contain multiple locations or multiple call numbers? Will this be a problem?
- Do you have discrepancies between the call number or location in the holdings record and the call number or location in the associated item records?
- Do you have records with empty call numbers?
- Do you have records with improperly constructed call numbers?
- Do you have records with call numbers that are associated with the wrong call number type?
- How is volume information structured and how is it mapping to the new system?

GENERAL ERROR CHECKING

Run as many reports as you can in your current system. These can be used both as tools for identifying problems to fix and as resources to refer to when trouble-shooting your migrated data later.

Counting is important when determining the completeness of migrated data. Count the number of records of any given type exported, as well as the number that show up in your new ILS. Counting is especially important for mapped data. Count the number of items where a given field, such as location, contain a certain code in your old system before you migrate. If possible, reconcile this with the number of items that appear in your new system. If you are mapping data from more than one code to a single code, match the total number of items with the original codes to the total number of items with the new code.

Counting is an important tool for checking the completeness of migrated data.

EXPORTING DATA AND AFTER

If you have a large quantity of bibliographic or item records, the size of the exported files can become unwieldy or may exceed the limits of your ILS export function. Consider how to split up your records to make smaller files. Can you use record numbers, date ranges, locations, or some other method? Can you be sure that you aren't missing some records? For example, if you group records by location, are there some records that don't have locations?

It is a good idea to keep a backup of your exported data at the point of migration. Although large files can be cumbersome, exported MARC records can be viewed and searched with tools such as MarcEdit. Item data and non-MARC data associated with bibliographic records can be exported in tabular form for use with database or spreadsheet software. If something goes wrong, you will have a record of the state of your data at the point of migration to which you can refer. If it is financially feasible or if your old ILS is locally hosted, you may wish to maintain access to it for some time post-migration for problem solving.

When importing data, your new system must have some way to associate item and holdings records with bibliographic records. There are two main strategies for doing this. Item and holdings record information can be exported in 8xx and 9xx fields in your MARC records. This can potentially make individual records containing many items very long. In fact, it may be impossible to generate a valid MARC record when large numbers of items are involved because of the MARC format's limit of 99,999 bytes. You may need to distribute the items across multiple copies of the bibliographic record and recombine them later. Alternatively, item records can be exported in a delimited text format, such as CSV, with an identifier that links them to the appropriate bibliographic record. Holdings records can also be exported separately with an identifier, such as the OCLC number or ILS system number, for the corresponding bibliographic record.

If you are able to run a trial import, check as many different types of data and functions as possible. Keep a list of data-migration problems that can't be resolved and will need to be cleaned up post-migration.

RESOURCES

Tools for analyzing and editing MARC records:
- MarcEdit. http://marcedit.reeset.net/.
- Marc Report/Marc Global. www.marcofquality.com/soft/softindex.html.
- The Library of Congress. MARC Specialized Tools. https://www.loc.gov/marc/marctools.html.

Vendors for record cleanup:
- Backstage Library Works. www.bslw.com/.
- Library Technologies, Inc. (LTI). www.authoritycontrol.com/.
- MARCIVE. www.authoritycontrol.com/.

OCLC holdings reconciliation:
- OCLC Reclamation Project. www.oclc.org/support/help/batchload/Content/004_Filling_out_form/003g_Select_one_time_project_type/Reclamation_Project.htm.
- Johnston, Sarah. "Homegrown WorldCat Reclamation: Utilizing OCLC's WorldCat Metadata API to Reconcile Your Library's Holdings." Code{4}Lib Journal 27, 2015. http://journal.code4lib.org/articles/10328.