Connect NY Unique Item Project

Tom Keays, Electronic Services and Resources Librarian
Noreen Reale Falcone Library, Le Moyne College
Connect NY Annual Meeting, June 3, 2011

Goal of Project

To free up space in our stacks, Le Moyne College Library wanted to identify withdrawal candidates from the circulating collection acquired earlier than a given date that had no record of circulation since 1994, when we acquired Innopac (and subsequently Millenium).

Requirements of Project

- To work from withdrawal candidate lists that were produced from Le Moyne's Millenium catalog.
    - Note: the lists were to be used to identify withdrawal candidates only. They were not intended to be "pull" lists.
- To maximize the effectiveness of Connect NY's Inn-Reach book lending network, we wanted to make sure that we were not withdrawing the last copy in the system.
    - After some discussion among the librarians, it was thought that preserving two copies in the system, when they existed, would make for a more robust sharing network.

Methodology of Project

1. Produce lists of items from Le Moyne's circulating collection acquired before a given date that had zero circulation recorded in Millenium.
2. Obtain lists of items from the Connect NY union catalog that were either:
    - Uniquely held by Le Moyne (we owned the only copy)
    - Held by Le Moyne and exactly one other library
3. Produce from these, lists of items that had not circulated AND which were
   held by two other Connect NY libraries
     OR (fallback)
   held by at least one other Connect NY library.

Steps

1. Using Create Lists in Millenium, identify a list of titles from your library's collection that will serve as a starting point for the withdrawal project. The STORE RECORD TYPE for the search will be BIBLIOGRAPHIC since corresponding lists that come from Connect NY can only be of this type.

   Le Moyne created a list of titles from our circulating collection identified by the locations **lbks** and **lbkso** (Bookstacks and Oversized, respectively) that had a **cat date** prior to a specific cutoff date and which had never circulated since 1994, when we first began using Innopac.

2. Export the preliminary candidate lists from Millenium to a spreadsheet. Decide which fields you will want to use in your final spreadsheet. They will necessarily have to also be present in your preliminary spreadsheet.

   At a minimum, you will want to export the **RECORD #**, **TITLE**, and **PUBLISHER** fields, which are all type BIBLIOGRAPHIC. For Le Moyne, the **CALL NO.** field is type ITEM (this may vary by institution). Le Moyne's export template is shown below. The BIBLIOGRAPHIC RECORD # will serve as the unique identifier throughout and allows us to link and compare records between several different spreadsheet tables that we will produce during the course of the project.

3. One of the search criteria used to produce Le Moyne's preliminary Create List candidate list was a **TOT CHKOUT = 0** ITEM parameter. This search matches on any BIBLIOGRAPHIC record which has one or more attached ITEM records, at least one of which has zero circulation transactions.

   Le Moyne's preliminary spreadsheet appeared as follows.

	A	B	C	D	E	F	G	H	I
1	RECORD	TITLE	PUBLISHER	ISBN/ISSN	OCLC #	CAT DATE	TOT CH	CALL NO.(ITEM)	RECORE
2	b1000001x	Sprachdunnheiten / W	Berlin : W. de Gruyter, 1949		10700151	1/8/2003	0	PF3460 .W97 1949	i1000001x
3	b10000033	South moon under, by	New York, Charles Scribner's So		10700368	1/8/2003	0	PS3535.A845 S68 1933	i10000033
4	b10000082	Accessible storage of r	Phoenix, AZ (22:089774084X (pa		10696542	1/8/2003	0	Z685 .W36 1984	i10000082
5	b10000719	Lazzaro. Come tu mi vu	[Milano] A. Mondadori, 1951		10720811	1/8/2003	0	PQ4835.I7 L3 1951	i10000719
6	b10000756	Abbotsford / painted b	London : Adam and Charles Blad		10717382	1/8/2003	0	PR5338 .C938	i10000768
7	b10000847	Science and imaginatic	New York, King's Crown Press, 1		10655619	1/8/2003	0	PR3327 .M57	i10000859
8	b10000860	Trois piÃ¨ces: Le regarc	Paris, Plon, 1931		0	10655707	1/8/2003	0	PQ2625.A755 T7 1931	i10000872

4. When there are multiple ITEM records attached to a BIBLIOGRAPHIC record, Millenium will export multiple instances of each ITEM field specified in the spreadsheet export template, causing the final output to misalign with the column headings. This is the reason we grouped all the fields of type ITEM last in our export template; BIBLIOGRAPHIC fields only appear once, so they should come first.

   For Le Moyne's list, it was simplest to remove rows which represented BIBLIOGRAPHIC records with multiple ITEMs from the spreadsheet. These excluded rows can be saved and revisited, for instance, as part of a separate withdrawal project centered around books with multiple copies.

	A	B	C	D	E	F	G	H	I	J
1	RECORD	TITLE	PUBLISHER	ISBN/ISSN	OCLC #	CAT DATE	TOT CH	CALL N	RECORE	(ITEM)
26	b10003769	Isidori Hispalensi	Oxonii : e Typographeo Clarenc		10741286	1/8/2003	0	0	PA6445.I3 E 1911	
27	b10003800	Christian mysticis	New York : Pue	916134636	10739187	6/29/1994	6	0	BV5082.2 .E32	
28	b10003873	Vatican II : Open	Wilmington, D	0894533665 : $9.	10745120	6/29/1994	9	8	0	
29	b10003885	Schillers Werke.	Stuttgart und Leipzig, E. Hallber		10745347	1/8/2003	0	0	0	
30	b10003964	Gift of communit	Wilmington, D	0894532286 (pbl	10745622	6/29/1994	10	0	BV811.2 .M3 1984	
31	b10004063	The Powers / Wal	Philadelphia :	080061786X (pbl	10752664	1/8/2003	0	0	0	
32	b10004671	The Acts of the A	London : Longmans, 1958		10828411	1/8/2003	0	0	BS2625 .M38	
33	b10004828	Elizabeth I and he	New York, Norton [1958-		10846524	1/8/2003	0	16	DA356 .N42	
34	b1000483x	The Jesuits : histc	New York, N.Y.	0688028616 : $1:	10847770	1/8/2003	28	0	BX3715 .B2913 1984	
35	b10005237	The origins of the	New York : Fre	0029221307 (pbl	10849780	1/8/2003	0	6	E185.61 .M845 1984	
36	b10005286	What they don't t	Toronto New	553050613	10850393	1/8/2003	0	1	HF5386 .M474 1984	

5. The next set of Create List spreadsheets that we need for this project have to be supplied by the III liaison to Connect NY as they cannot be produced by the local library. What we needed was a way to identify items from our collection that were unique to Le Moyne or which were held by Le Moyne and exactly one other Connect NY library. We intended to use these two lists – what we called our "unique" and "unique plus one" item lists – to identify and remove corresponding records from our preliminary candidate list. Any item not removed in this process should therefore have holdings in two or more other Connect NY libraries.

   To produce these two lists, Le Moyne contacted Bart Harloe who arranged for us to work with Jean Berard from III. To get started, I arranged a phone conversation with Jean where I described what we wanted her to produce. With only a few follow-up emails to clarify the process, Jean created and loaded her tables directly into Le Moyne's Create List space (rows 4-6

below). Since the "unique" item list that Jean produced exceeded the 60,000 Max Record limit of Le Moyne's largest Create Lists slots, it had to be split into two parts. The "unique plue one" list fit in one of our 40,000 slots.

Note: These two "unique" item lists are not tied to a specific Le Moyne location or material type, so they conceivably could be used in similar projects for other types of circulating materials – e.g., videos, teaching materials, etc.
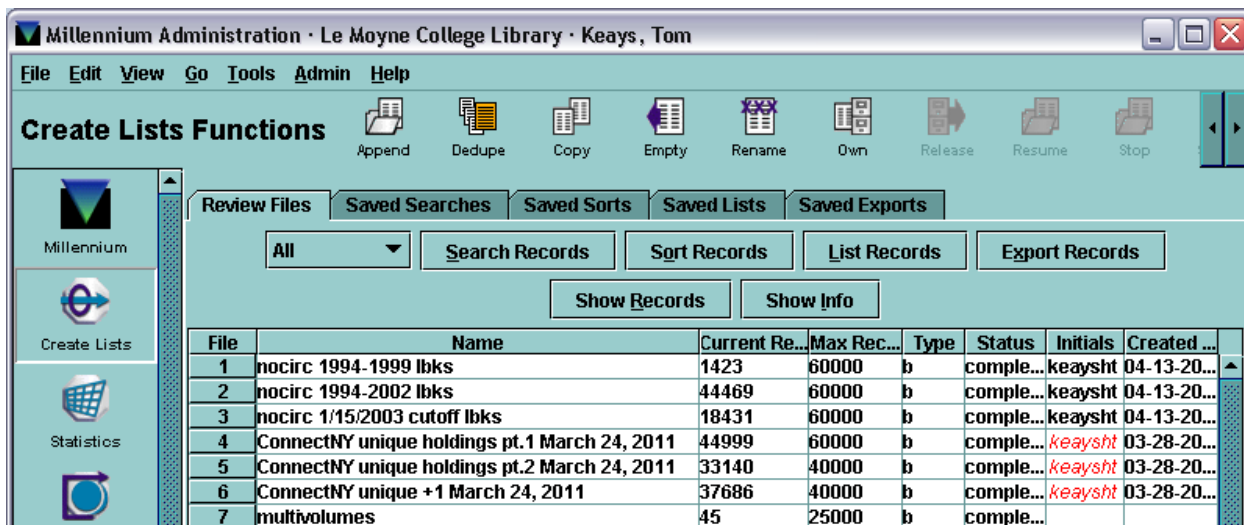
6. Export the "unique" and "unique plus one" lists from Millenium as Excel spreadsheets.

   Le Moyne exported these speadsheets using the same export template we used in step 2 to produce our preliminary withdrawal candidates list. Strictly speaking, the BIBLIOGRAPHIC **RECORD #** field is the only field we needed to export to accomplish the goal of our project – i.e., removing those rows from the preliminary candidate list that were also in either the "unique" or "unique plus one" lists. However, since librarians wanted to see what was represented in the "unique" and "unique plus one" lists, it was easiest just to export them in the same format as the preliminary candidate list.

7. At this point, there are three Excel spreadsheets: the preliminary withdrawal candidate list (produced locally) and the "unique" and "unique plus one" lists (produced by III). The preliminary list is our starting point and, using the BIBLIOGRAPHIC RECORD #, which is common to all three spreadsheets as the hook, we want to eliminate all items from the preliminary list that are found in either the "unique" and "unique plus one" lists. This operation, expressed in database searching terminology, is known as a "NOT" operation.

   There is no way to accomplish this procedure in Excel. Fortunately, it can easily be done using Microsoft's Access database program.

8. Open **Access** and create a blank database – i.e., a database with no tables. (Actually, it will create a table named "Table1" which you can either ignore for now or delete.)

9. Under the "**External Data**" tab, under the "**Import**" section, click the "**Excel**" option.
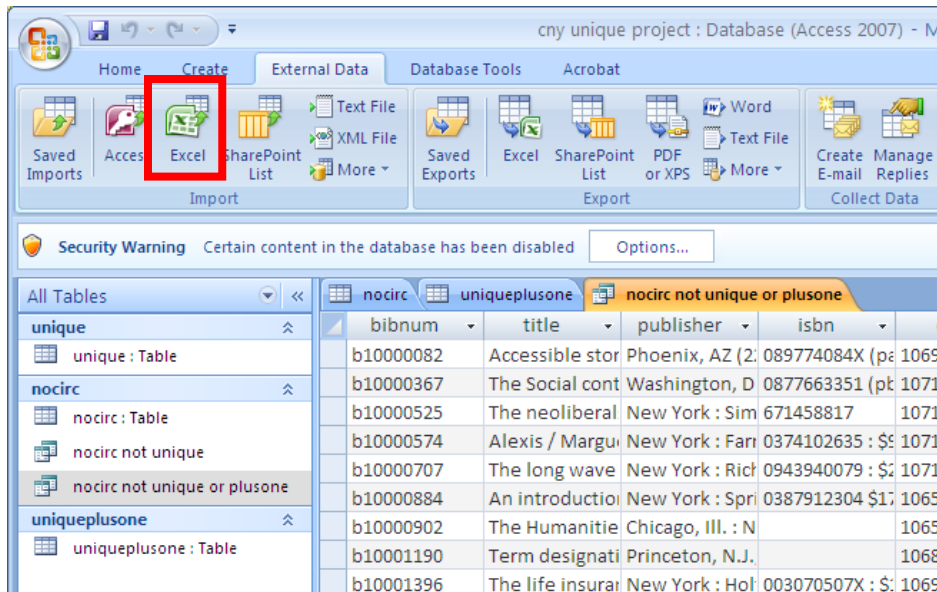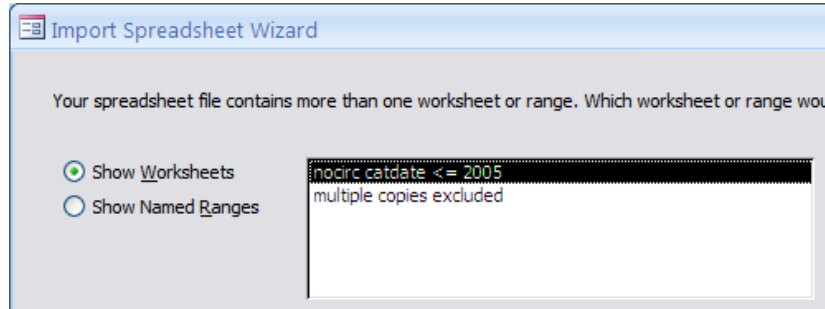

10. That will open a dialog box to "**Import the source data into a new table in the current database**."

   Click the **[Browse]** button, locate the spreadsheet you want to import, and click **[OK]**.
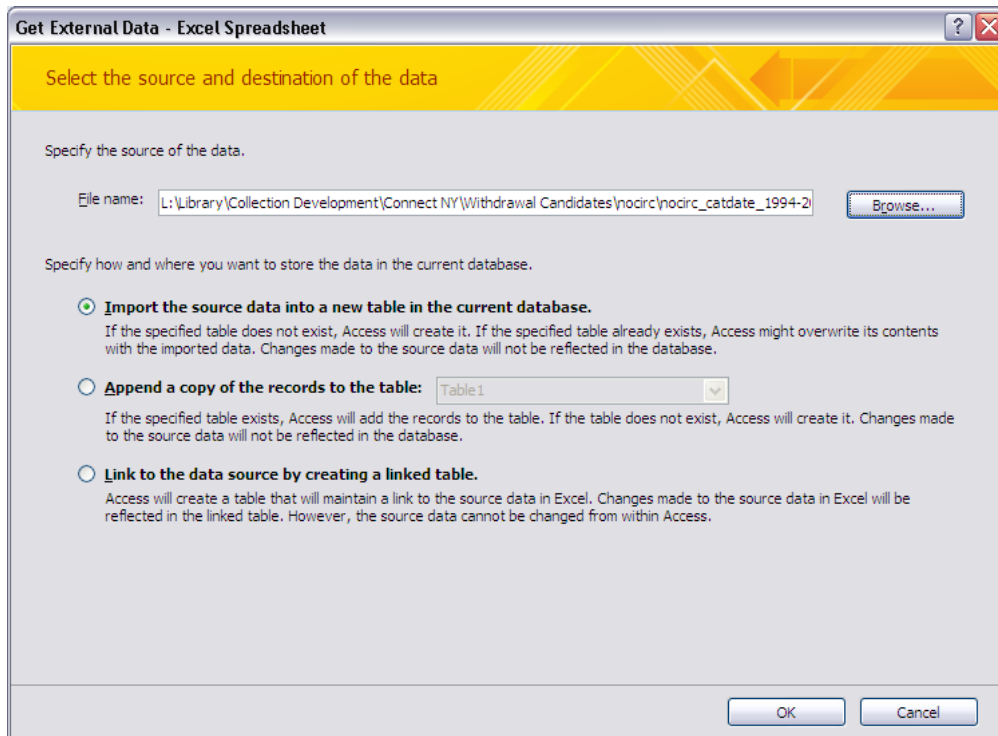
11. Follow the instructions provided by the import wizard.

    a. If your spreadsheet has more than one worksheet, select which one you want to import and click the **[Next >]** button to continue.

    b. If your spreadsheet has column headings, checkmark the "**First Row Contains Column Headings**" option and click the **[Next >]** button to continue.
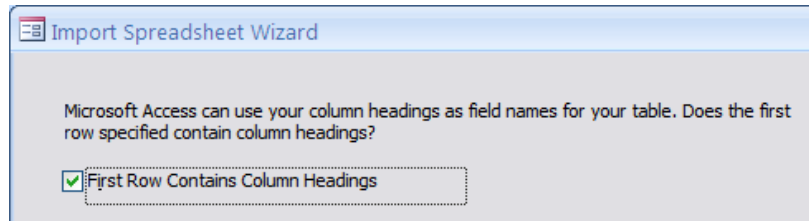
    c. Each column of the spreadsheet will be represented in the next dialog box. If you didn't rename the column headers created when the Excel spreadsheet was exported from Millenium, this is your opportunity to rename the fields in the final version of the spreadsheet you are working to create. E.G., you could change "**RECORD # (BIBLIO)**" to something like "**bibnum**". This is optional.

    However, it is important, since we want to use the "**RECORD #**" field as the unique identifier for the NOT operation, to change the "**Indexed**" field option from "**No**" to "**Yes (No Duplicates)**". The "**Data Type**" can be left as "**Text**".
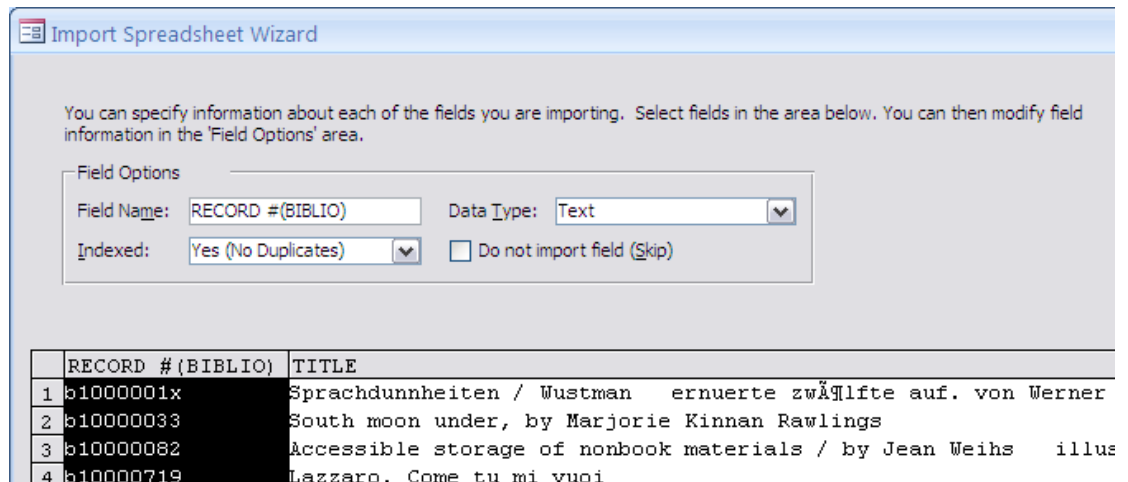
Scroll horizontally to examine all the columns. It matters less what "**Data Type**" you assign to the other columns, but there's no reason not to set them correctly, as it will simply things later. E.G., change the "**Data Type**" of the "**OCLC #**" column to "**Integer**", "**CAT DATE**" to "**Date/Time**," etc.

When you've finished fixing the columns, click the **[Next >]** button to continue to the next screen.

d. Select "**Choose my own primary key**" rather than the letting Access assign one. Choose "**RECORD # (BIBLIO)**" (or whatever you have renamed it to) as your primary key. This will let Access more quickly perform comparisons across tables in the database. Click the **[Next >]** button to continue to the next screen.
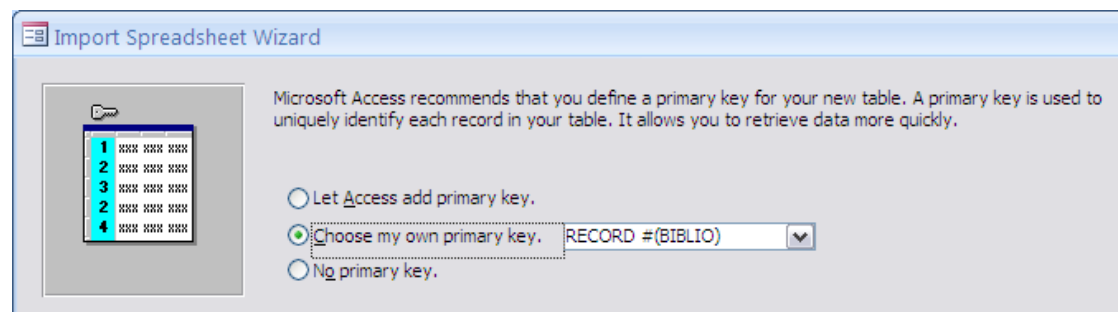

e. Give the table a name and click **[Finish]**. It is unnecessary to "**Save import steps**" when prompted.

12. If you see an extra table created that contains the name of the table you just imported plus "**Import Errors**", examine it and see if any of the errors were crucial. Very likely you can just delete this table without having to do anything. Errors will be less likely to occur if you set up the Data Types for each column ahead of time (see step 11c).

13. Repeat steps 9 – 11 for all three of your Excel spreadsheets, creating three corresponding tables in Access, each with "**RECORD # (BIBLIO)**" (or equivalent) as the primary key.

14. With all three tables in place, you are now ready to perform the NOT operation. You do this by building a "query" where you compare rows in one table to rows in another table. Access then defines a relationship between the tables and creates a new table that displays a new set of results based on your query. Like most "relational" databases, Access writes queries in a language known as "Structured Query Language" or "SQL". You will be happy to know you do not have to learn any SQL, but can instead rely on a simple to use "**Query Wizard**".

15. Access calls the "**NOT**" operation an "**Unmatched Query**," which it defines as "find records (rows) in one table that have no related records in another table."

Go to the "**Create**" tab in Access and click the "**Query Wizard**" button.
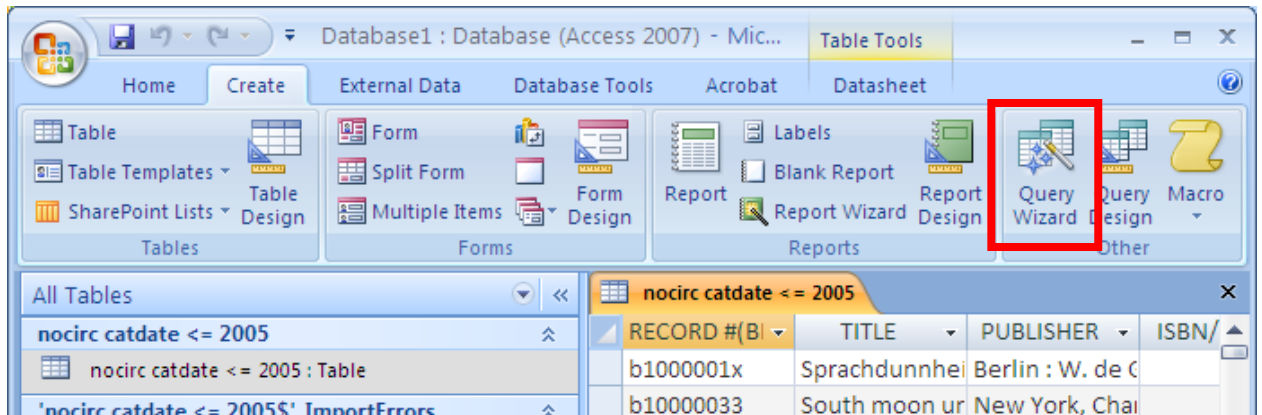
16. This will open the "**New Query**" dialog box. Select "**Find Unmatched Query Wizard**" from the choices and click the **[OK]** button.
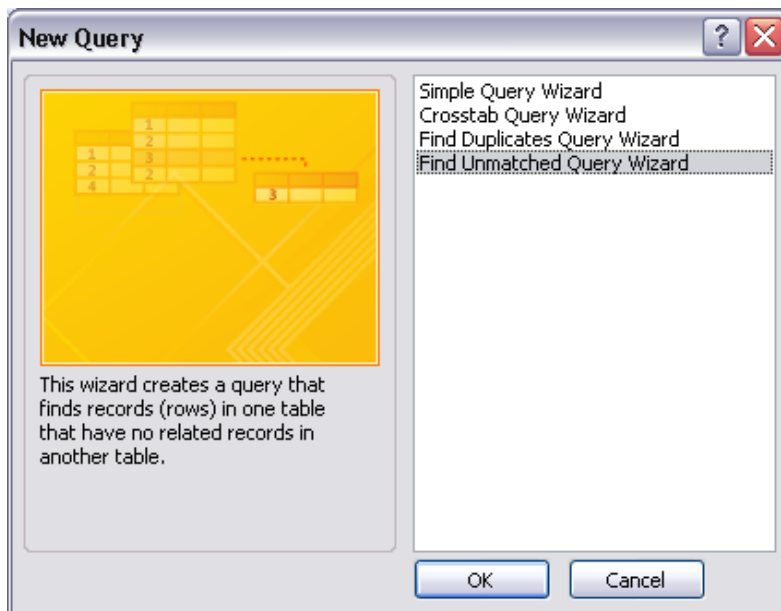

The "**New Query**" wizard will prompt you for the following information.

   a. **Which table or query contains records you want in the query results?**
      Select the table that contains the preliminary withdrawal candidate list. In the example, it is named "**nocirc**". Click the **[Next >]** button.
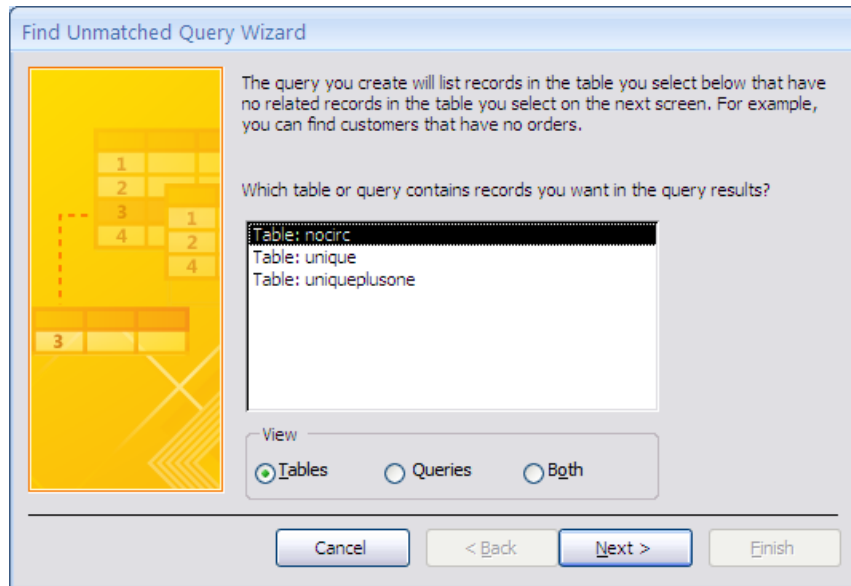
b. **Which table or query contains the related records?**
In the example, we selected the "**unique**" table. Click the **[Next >]** button.
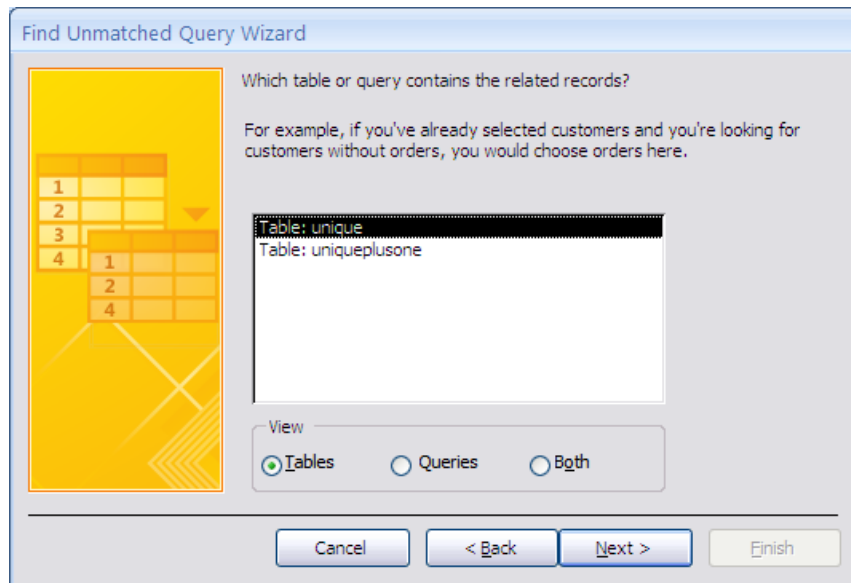

c. **What piece of information is in both tables?**
In the example, we selected "**bibnum**" (which is how we renamed "**RECORD # (BIBLIO)**" as it originally was labeled during the Millenium export). This is the primary key we defined earlier and it is perfect for doing this sort of relational comparison between the tables. Click the **[Next >]** button.
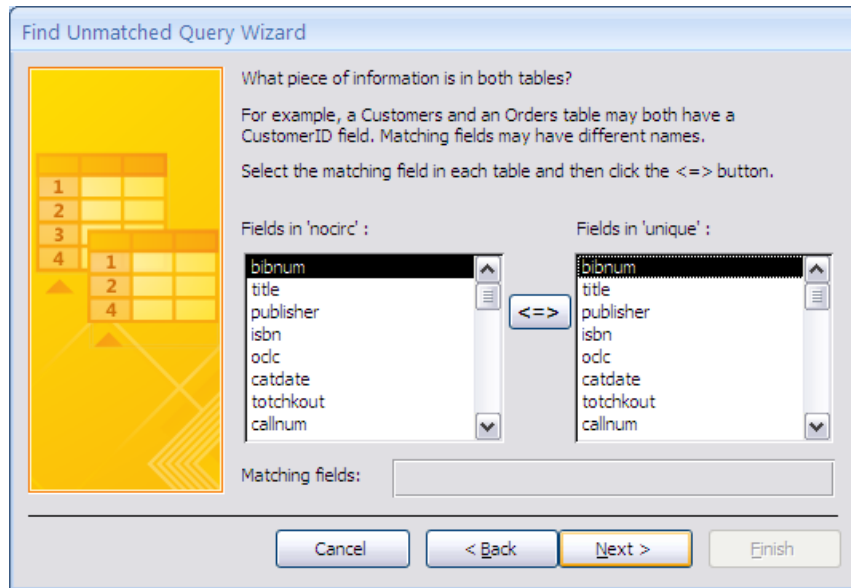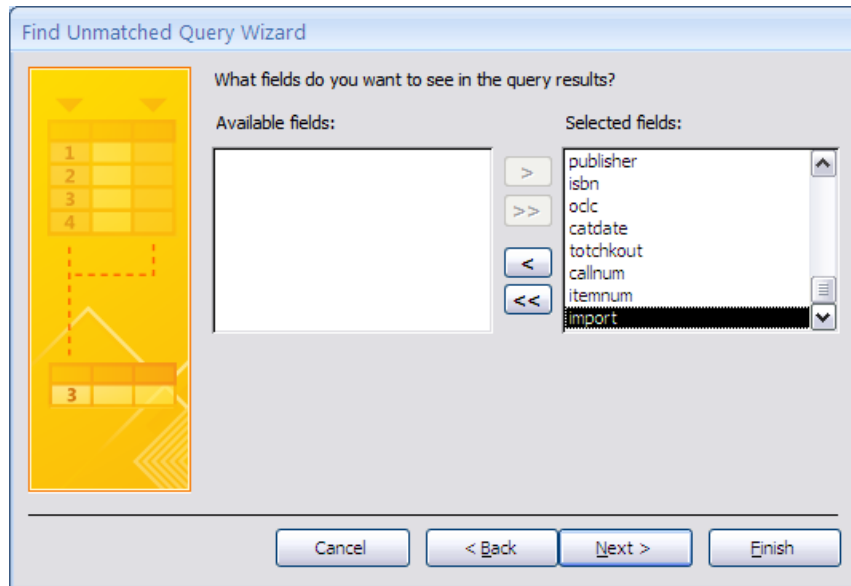
d. **What fields do you want to see in the query results?**
Because we want all the columns from the original "**nocirc**" table to carry over to our output, click the **[>>]** button to bring in all the fields.
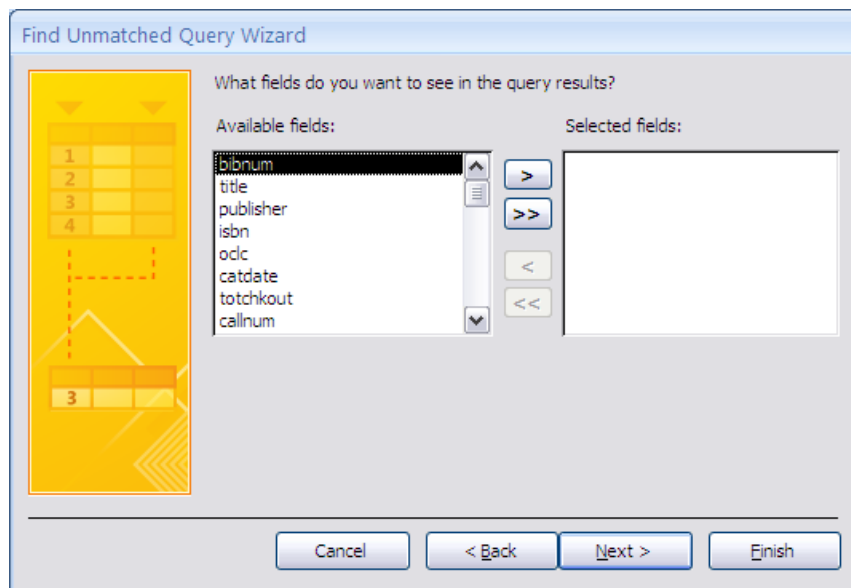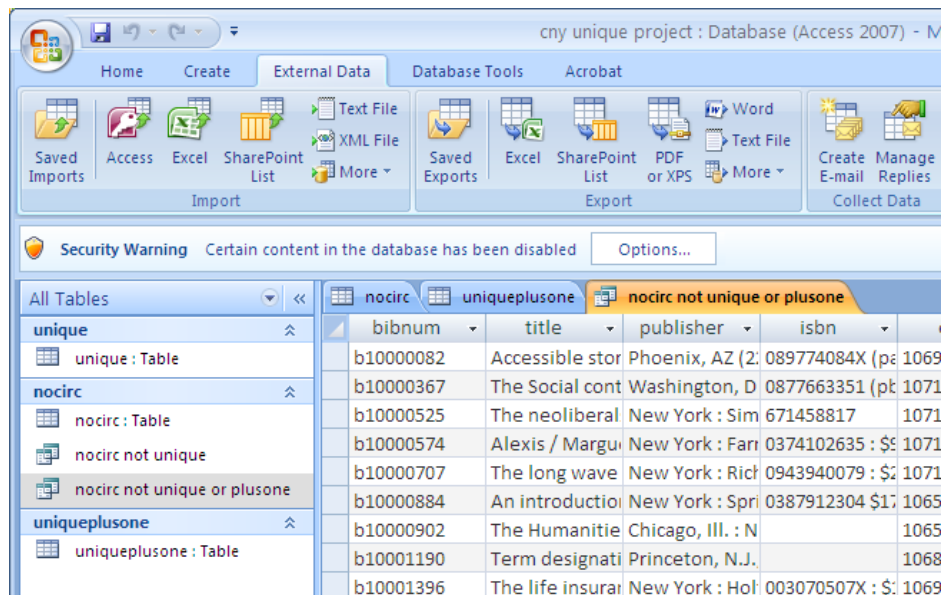

The fields should have all moved from the "**Available fields**" pane to the "**Selected fields**" pane.

Click the **[Next >]** button.

e. Name the new table and click the **[Finish]** button. The new query table will appear in the list of tables. Query tables can be identified by their icon of two overlapping tables.

17. The query table we just created, named in the example "**nocirc not unique**", can itself be used as the source data for a second "NOT" query.

Follow steps 16a – 16e differing in detail only in that you will be using "**nocirc not unique**" query as the **Which table or query contains records you want in the query results?** starting point and "**uniqueplusone**" table will be used as the **Which table or query contains the related records?**

If you don't see the "**nocirc not unique**" query, make sure the **View** "**Queries**" or "**Both**" radio button is selected.
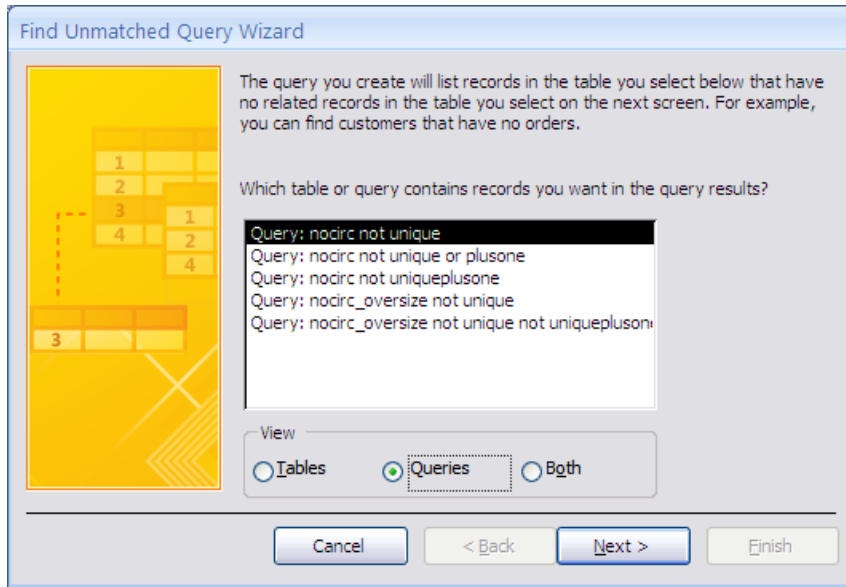

18. If we've done everything correctly, we'll end up with (at least) two new query tables in addition to the tables we imported from Excel.

    The final query table we created, labeled in the example as "**nocirc not unique or plusone**" is our desired end product. In it, we've taken a preliminary list of titles with no circulation record and removed all the records that were unique to Le Moyne and then further removed all the records that were owned by one other Connect NY library. This leaves a list of titles that are owned by Le Moyne and two or more other Connect NY libraries.

19. Because we will not need to run any further queries on the Access database, it makes sense to export the "**nocirc not unique or plusone**" table back to Excel, rather than expecting people to use data directly in the Access database.

    Under the "**External Data**" tab, in the "**Export**" section, click the "**Excel**" button. Follow the dialog box instructions to save the spreadsheet to your hard drive.
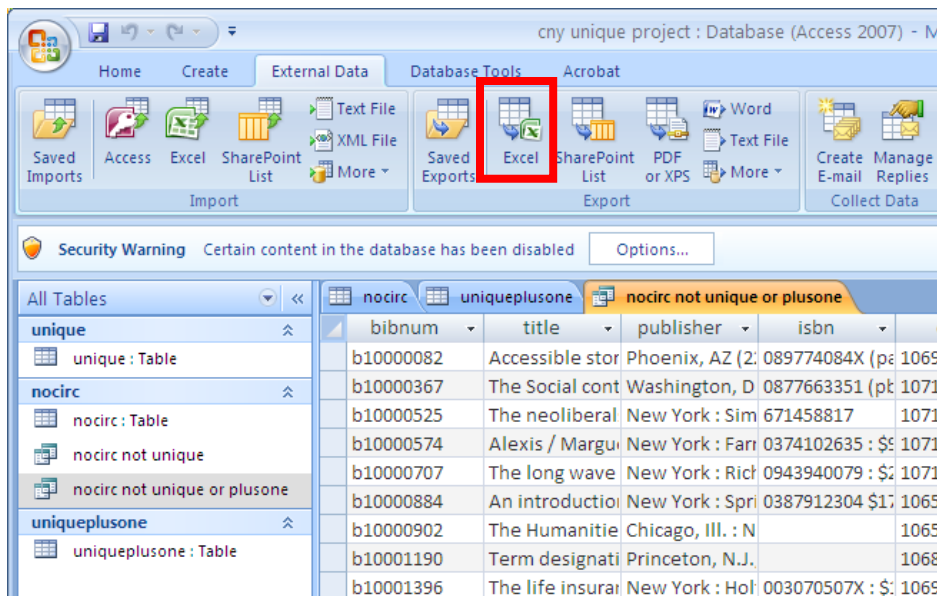
Notes

1. Copies of the final "**nocirc not unique or plusone**" spreadsheet can be further reduced by individual librarians who want to create smaller lists tailored to their subject areas.

2. In hindsight, given that Excel cannot properly sort by call number, I would have done a call number sort of the initial Create List (step 1) in Millenium. The sort would have created duplicate lines for bibliographic records with multiple attached items (copies or volumes), but this is easily de-dupped in Excel. Look for the "**Remove Duplicates**" button in the "**Data**" tab in Excel. The "RECORD #" column is all you really need to use for the de-dupping.