Lessons Learned from a decade of system migrations

Presented by:
Wes Osborn
Central Library Consortium

NOTSL Dec. 2022
Wes Osborn
Central Library Consortium
https://clcohoio.org/category/conferences/
AGENDA

BACKGROUND
What I do and don’t know

TEAM
Who should be on your team?

DATA
What is important and what is difficult

TIMING
What you should do when

TOOLKIT
What are some useful tools
25 years in Public Libraries

Worked with structured data

Migrated dozens of systems

I suck at authority control
Let’s find out about you
Join at slido.com #4697 326
You know your data best
You aren’t the first
You are in control
Your team

Someone who knows your power users
Someone who is familiar with your new system
Someone who knows spreadsheets or loves to comb through data
Great opportunity for growth
Working with others in the organization

https://www.worldcat.org/title/1290489949

Have a heart.
What do you really want?
Rational (the planner)
Weak
Logical
Thoughtful
Analytical
Requires direction
Clarity is key
Gets stuff done

Direct the rider

Sounds good

Motivate the elephant

Feels good

Shape the path

Behavior change
Jonathan Haidt

Emotional (the doer)
Strong
Passionate
Impulsive
Requires moral purpose
Stories are key
Overthinker
Maintain relationships with current vendor

- You’ll need their help to extract data
- Check licensing requirements if you need old system for reports
- The world needs more kindness
Understand your new vendors preferred communication method

If they prefer email, batch similar issues together – don’t send ONE massive email
Use good subject lines
Do NOT send patron data via email
What data to migrate
Spotting high complexity low migration value records
Ask your New Vendor
Request a copy of the data migration guide for the new system

Even if you personally don’t use it, the fact that they have one matters.

Ask for it before buying the new system.
Don’t fret over data you can’t move

If your new system doesn’t support it, you don’t have to move it!
The more connections, the more difficult
A Patron Record

- Items out
- CKO History
- Notice History
- Payment or fine balance history
A serials record

- Bibliographic Record
- Item Record
- Serials Holding Record
- Issue Record
- Publication Pattern
- Purchase Order
- Purchase Order Line Item
- Invoice
- Invoice Line Item
If it is user data, make it a migration priority.

For example: Properly migrating how patrons are notified about account transactions.
Focus on the data more people interact with
How many people manipulate the records?

The fewer the more likely the data is specialized to the current system and may NOT feasibly migrate.
Quiz Time – Where should you focus?

Customer data priority

More staff interactions = higher priority
Some data might get left behind
Don’t delete it yet

Even if you aren’t keeping it, that info could still be helpful in mapping data from the old system into the new one
Manual input is a teaching tool

Spending time learning how the new system works

Instead of cleaning up a messy import
Your new system may have better cleanup tools
Public Library Migration Cautions

Holds / Reserves / Requests

Fine and Payments

Lost titles
What to do with holds

Don’t migrate them

Special word of caution for consortia
What to do with fines

Is it time to go fine free?
How to handle lost items

Ask lots of questions about how the new system works
Why are you keeping the lost items?
Could you not migrate them?
When to migrate?

- Consider your typical annual cycles
- Academic Library: Slow Summer Session
- Public Library: Busy Summer Reading
Other migration date considerations

When do you run stats & reports?

When do contracts expire?
Slowdown activity prior to migration
Sample timeline – Oct. 1st go-live

- **June 1st**: Extract Data
- **July 1st**: Test Load
- **August 1st**: Perform 2nd extraction and test load
- **Sept. 1st**: Stop loading non-critical records and no more system changes or data cleanup in OLD system

Additional:
- **June**: Perform Mapping
- **July**: Fix Mapping perform source cleanup
- **August**: Finalize Mapping
- **Sept. 30th**: Run final reports, backups, etc.
If you can’t do two test data loads

Ask about version control process on data import scripts
What to do with old data? Keep It!

<table>
<thead>
<tr>
<th>Good</th>
<th>In plain text export files.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Better</td>
<td>In the old system.</td>
</tr>
<tr>
<td>Best</td>
<td>In the new system.</td>
</tr>
</tbody>
</table>
Benefits of having old data in new system

- Quickly reprocess old data if needed
- Running Reports from previous time periods
- Answering “Why did this migrate wrong?”
What to watch for in the data

- Non-printable characters
- Spaces / Line Breaks
- NULL things that shouldn’t be (barcode fields, etc.)
- Unexpected delimiters
Join at
slido.com
#4697 326
Special considerations
Be mindful of call numbers

Some systems use the MARC display

Some systems use the item display
If you are joining a consortium

When de-duping MARC records, always keep the EXISTING record – this makes migrating holds HARD

If you are a small library, it may be easier to manually attach add your items to existing records
If you are leaving a consortium

Attach consortium records to a single “unknown” title/patron
Special tools
Regular Expressions (AKA regex)

Near universal system for finding and replacing text
A life-long skill
Tutorial: https://regexone.com/
Playground: https://regex101.com/
A good text editor

- Should support regex
- Should support large files
- Should support line wrap / un-wrap
EmEditor

Windows
Fast
Large Files
Not Free (Trail)
Notepad++

Windows
Free
Find in Files
Slow on large files
BBEdit

MacOS
Fast
Large Files
Free & Paid Version
Other ways to manipulate text

- **Powershell (Windows):**

- **Grep (Linux & Mac):**
Your favorite spreadsheet
Expand your spreadsheet knowledge

Filter columns to look for unusual data
Use pivot tables to group and spot weird data
Be VERY careful with the ways it might “mangle” data – use text export
New (AI) tools on the horizon

https://flatfile.com/products/workspaces/

https://dromo.io/pricing
OpenRefine

- Can handle large tabular datasets
- Can parse different types of data

Get started:

https://www.youtube.com/watch?v=nORS7STbLyk
https://multimedia.journalism.berkeley.edu/tutorials/openrefine/
MARCEdit

- REGULAR EXPRESSION SUPPORT
- DE-DUPING SUPPORT
- TREATS MARC FILE AS INDIVIDUAL RECORDS
Use tools to spot problems

Do NOT cleanup data during migration
Clean up data in OLD system
OR alert data migration person of issues WITH EXAMPLES
You know your data best
You aren’t the first
Join at slido.com #4697 326
Wes Osborn
Central Library Consortium
https://clcohoio.org/category/conferences/