

More books are [not always] better: Managing the complex ebook purchase model landscape at a small university

Melissa Belvadi
Rosie Le Faive
University of Prince Edward Island

Ontario Library Association
SuperConference 2019



UNIVERSITY
of Prince Edward
ISLAND

Context

- Quick questions about the audience
- Institutional context - UPEI
- Library context - size, budget, ILS

UPEI Ebook involvement

- Firm order - many platforms mostly thru Gobi 1U,3U,NLL/Concurrent/UU
- Subscriptions - EBSCO, Proquest, desLibris
- Collection Purchases - ACUP/Ebound, Duke
- DDA - EBSCO, Proquest (Gobi)
- EBA / EBS / UBCM

EBA_s

EBA model:

- basic concept
- why (publisher, library)

EBA - Big Issues

Concerns about EBA as a model

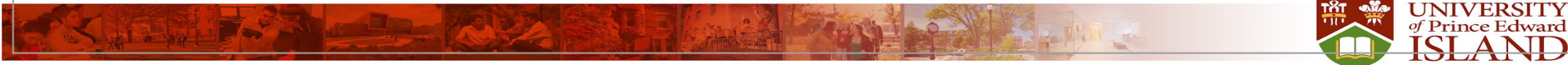
- Sustainable / Cost-effective? (Library)
- Long term impact on book market/small/non-profit publishers?

UPEI EBA experience

Vendor	Titles	Used 1+	Used 10+ (BR2)	Other access	Spend count	Spend lowest use	Decision
JSTOR	35,000	166	35	83	52	1	renewed same terms
year 2	40,000	702	39				
Project Muse	12,000	178	13	47	25	3	renewed same terms
year 2	12,400	58	8				
PM- Only 3 titles used in both years							

UPEI EBA experience

Vendor	Titles	Used 1+	Used 10+	Other access	Spend count	Spend lowest use	Decision
Sage	4,800	168	17	0	40	4	Renewed same terms
year 2	5,100	228	29				
Just 12 of spendout used in 2018							
Wiley	261	136	29	0	61	turna ways	renewed same cost, expansion to all collections
year 2	21,600	188	107 (BR1)				



UPEI EBA experience

Vendor	Titles	Used 1+	Used 10+	Decision
Elsevier	403	9	3	rolled over, total coll avail
year 2	29,400	331	34	usage is 2-year
Cambridge	120	3	1	rolled over, expanded list
year 2	2,700	37	8	usage just 2018

Cataloguing's task: Load the records

- **ILS:** Evergreen 3.0.3 (Open Source, yay!)
- **ERM:** None
- **Acquisitions module:** Nope (our choice)
- **Discovery Layer:** EBSCO EDS, populated weekly from Evergreen
- **Cataloguing Librarian:** 0.2
- **Cataloguing Technician:** 1.0



Challenges specific to our ILS (?)

Vandelay import/export (batch load):

- tends to fail if batches are larger than 500 records.
- is really slow at matching (>3 hours to process 500, no progress bar).
- can't do conditional logic → matches must be processed individually.
- *can* match efficiently if incoming record contains existing bib id

XML directly into database: terrifying, time consuming

Deleted records stay in the catalogue.

Why bother matching?

Why bother 1: Accidental DDA Triggers

We can't "afford" to firm-order books so we don't want to accidentally let DDAs get purchased if we have the resource in a different ("better") license.



UNIVERSITY
of Prince Edward
ISLAND

Why Bother 2: Minimizing duplicate records



5. **Field Guide to Mushrooms of Western North America.**

Davis, R. Michael,



E-book (2012.)

- Full text via EBSCOhost
- Full text via JSTOR

Publisher: Berkeley : University of California Press, 2012.



1. **Beyond the cyborg adventures with Donna Haraway**

Grebowicz, Margret, 1973-



E-book (2013.)

Full text via Proquest Ebook Central

Publisher: New York : Columbia University Press, 2013.

2. **Beyond the cyborg : adventures with Donna Haraway**

Grebowicz, Margret, 1973-



E-book ([2013])

Full text via EBSCOhost

Publisher: New York : Columbia University Press, [2013]



UNIVERSITY
of Prince Edward
ISLAND

Step 2: Design some logic?

MARC Record Load Matrix: Assumes new record is different platform from old one

	DDA	EBA	Subscription	Purchased
DDA	prefer plan X	discard DDA	discard DDA	discard DDA
EBA		keep both*	keep both*	keep both*
Subscription			keep both	keep both
Purchased				keep both

* Have to watch for these during spendout

Step 1: Document the Collections

	Access Engineering	Adam Matthew	Alexander Street Press	ANSI	ASME	CABI	California Digital Library	Cambridge	Credo Reference	Curio	desLibris	Duke	EBSCOhost	Films on Demand	Gale Reference	InteleX	JSTOR	JoVE	Kanopy	Knovel	NFB Campus	Oxford Reference	Oxford Scholarship	Project Muse	Proquest	Sage	Scholars Portal	ScienceDirect	SpringerLink	Taylor & Francis	Wiley
Purchased																															
Subscription																															
EBA																															
DDA																															
Open Access																															

Surprise challenge: MARC record delivery

- Platforms usually distribute MARC independently (only 4 use OCLC)
- Identifiers (020, 035) are consistently inconsistent
- Constant updates, additions, deletions (for DDA, EBA, Subscription)
- Can't assume the platform took care of removing purchases / other entitlements
- Entitlements frequently range in size from 1000 - 200,000





Marc-a-roni

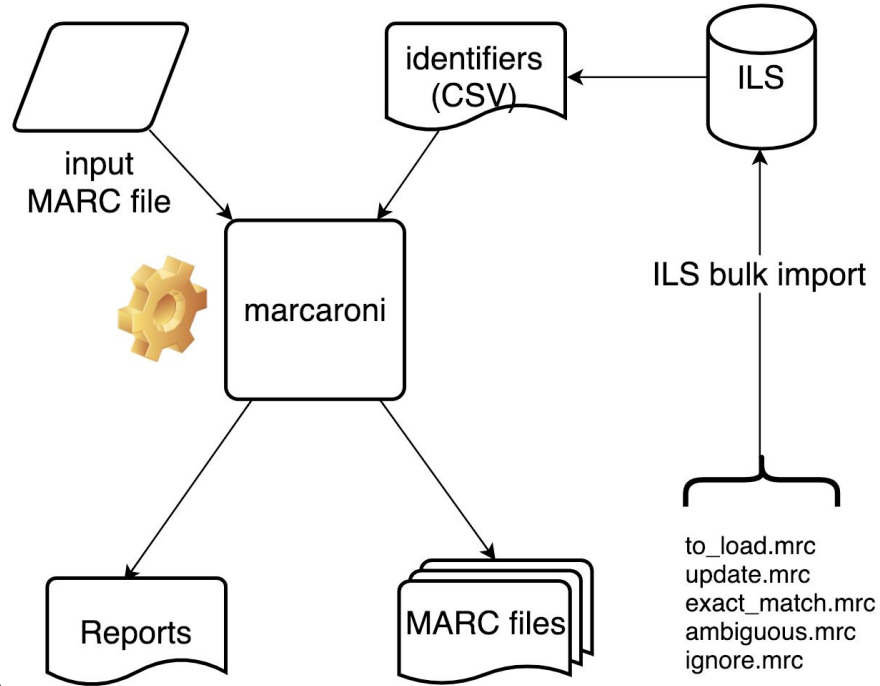
Python 3 + pymarc + isbnlib

Command Line interface, still in proof of concept (come play?)

<https://github.com/mutronic/marcaroni>



Marcaroni Workflow



Making Pasta by Andrea Parrish - Geyer CC-by-nd 2.0
<https://www.flickr.com/photos/tinytall/6867506889>



UNIVERSITY
of Prince Edward
ISLAND

What it does

- User provides it with a .mrc file and “bib source” (code denoting a collection/platform)
- Marcaroni evaluates “matches” in the catalogue using 020 and 035 tags
- Marcaroni splits the records in the input file into (up to) five separate files:
 - New records to add
 - Records to upgrade (matched same platform, lesser license)
 - Records to ignore (matched same platform, better license)
 - Records to optionally update (we already have this record, same license)
 - Ambiguous records (multiple matches, probably need to fix)



Step 3: Create some rules

- If the new record is DDA and we have it any other way,
 1. “**Ignore**” this record (don’t load)
 2. Tell the vendor to disable DDA on the title
 3. Fret about potential loss (if it drops out of the other subscription/eba)
- If it’s DDA, there’s only one match, and it’s on the same platform/license:
 - Then mark it as an “**exact match**”
- If the new record is DDA and all other matches are also DDA:
 - Then mark it as “**ambiguous**” - determine preference manually.
- If there aren’t any matches on the same platform:
 - Then “**add**” it.
- If there are multiple matches on this same platform:
 - Mark this record as ambiguous - the other records need to be sorted
- If there is only one match on this platform:
 - If it’s the same license, mark as “exact match”
 - If the existing record has a “better” license, mark it as “ignore”
 - If the existing record has a “worse” license, “**update**” the existing record.
- Also, if the new record isn’t DDA but some matches are:
 - Remove the DDA record, tell the vendor, and fret
- If none of the above apply, mark as “ambiguous”



Starting Marcaroni: 2018-10-10 10:44:08.081339

Bibsource: DesLibris-subscription

Record count: 748

Number of records to add: 142

Number of records to update: 0

Number of exact matches: 565

Number of records to ignore: 4

Number of records to figure out: 37

DDAs that need to be deleted: 1

Matches per Bibsource:

source	name	count(records)
53	DesLibris-subscription:	628
11	Proquest: All Purchases [Purchased]:	611
68	EBSCOhost: NA Acad [Subscription]:	139
1	oclc:	80
71	Proquest: Academic [Subscription]:	55
75	ScholarsPortal: ACUP CRKN [Purchased]:	40
44	desLibris: PurchasedCRKN:	20
22	OpenAccess/Free:	7
86	ProjectMuse: EBA [EBA]:	7
54	EBSCOhost: DDA from Gobi [DDA]:	1
51	EBSCOhost: All purchased [Purchased]:	1
7	GOVDOC:	1
2	System Local:	1

Results of doing things this way

Pros	Cons
• No accidental DDA triggers • No duplicate records on same platform	• Potential DDA lost opportunities • Many duplicate records across platforms

Results of doing things this way

Pros	Cons
● No accidental DDA triggers ● No duplicate records on same platform ● Cataloguer can keep up with demand	● Potential DDA lost opportunities ● Many duplicate records across platforms ● Deep understanding of context required ● Increased work & busy work for cataloguing technician ● Troubleshooting required (ambiguity?) ● Potential for false positive matches ● Technical Debt



Thanks for the help

- Alex Carmel-Veilleux
 - Object-oriented programming, clean code, help with unit tests
- Equinox Library Services (esilibrary.com)
 - Clarifying Evergreen's strange behaviours
 - Tips on programmatically overlaying/ingesting records from XML
- stolenpencil on Spoonflower
 - Pasta fabric design





Melissa Belvadi
mbelvadi@upei.ca

Rosie Le Faive
rlefaive@upei.ca



UNIVERSITY
of Prince Edward
ISLAND