VALIDATING ONLINE SERIAL HOLDINGS

with Help from a Worldcat API

James Van Mil – OVGTS, 2011
The Problem:

Keeping our online serials knowledgebase accurate
The Problem:

- Publishers!
  - With mergers, purchase model changes, and poor record-keeping, serial holdings are often in flux
  - It is time-consuming (and painfully boring) to check up on things
The Problem:

- But we’re part of the problem, too:
  - Serial title change cataloging rules make this task too complicated to hand off to a student worker
  - And we don’t have enough work-hours from trained folk to be able to keep up (even if it weren’t painfully boring...)
The Problem:

- And, to exacerbate the importance:
  - Increasing integration of our knowledgebase with other systems (ILLiad, Summon (or other discovery layers)) multiplies the impact of problem in the knowledgebase
A solution?

Some automation, using…
ACRONYM SOUP!!!

- KBART
- PIE-J
- APIs
- xISSN
KBART

Knowledge Bases And Related Tools

- The idea:
  - Consistent, clear, machine-readable lists of holdings information that can be uploaded to an institution’s knowledgebase.
  - Inspirational! (Hopeful?)
Presentation and Identification of E-Journals

- Really, title change problems
- Publishers are getting better, but this is still a problem in holdings data. A lot of progress has been cosmetic, not part of the infrastructure:

_Previously published as: Waves in Random Media (0959-7174, 1361-6676) until 2005_

(This is on the website, but not in the back-end data)
API

Application Programming Interface

- (Relatively) open sources of data that let programs exchange information
- For example!...
API

Application Programming Interface

Learning Python / Mark Lutz.
Lutz, Mark.
xliv, 700 p. : ill. ; 24 cm.

Book

(from the OhioLINK catalog)
Part of the WorldCat API
Gives you access to title data for serials
Fun example of the service in work...
Hercules's 13th labor (Serials cataloging)
The process

- Get some holdings information
- Get your knowledgebase holdings

- Need:
  - ISSN
  - Title coverage dates (from KB)
  - Your holding dates (from publisher)

- Run the Script...
Run the script!

- What it does...

(If everything goes well...)
Matching on ISSNs, the script updates all the knowledgebase holdings, using the data from the publisher.

For example:
Example, pt. 1

- Our publisher says we have:


- So the knowledgebase gets updated with this info...
2)

- The script now starts looking for problems.
- The default data in our knowledgebase is respectful of title changes, which means that the default coverage dates are an excellent match point for problems...
Example, pt.2


- But, our knowledgebase says that this title actually started in 2005, not in 1991!
- Something is Wrong!
TO THE xISSN API!!!!!
So the script has learned that this title actually has a preceding title, so it splits up the holdings between them:

-Waves in Random and Complex Media: 2005-present
-Waves in Random Media: 1991-2004
Having saved the day on all of the straightforward stuff, the script goes looking for other problems:

- Titles we no longer have access to.
- Titles with an unclear relationship…
Titels with an unclear relationship.

The Philosophical magazine: comprehending the various branches of science, the liberal and fine arts, geology, agriculture, manufactures and commerce

1941-5796
1941-580X

Philosophical magazine and journal comprehending the various branches of science, the liberal and fine arts, geology, agriculture, manufactures

Edinburgh journal of science
0367-0287
The philosophical magazine or Annals of chemistry, mathematics, astronomy, natural history and general science
1941-5850
1941-5869

The London and Edinburgh philosophical magazine and journal of science
1941-5966
1941-5974
Other sources of error:

- There are limitations to this approach:
  - Missing ISSNs
  - Bad xISSN data
At the end of the day

- You get two spreadsheets:
  - One to upload to your knowledgebase
  - And one full of problems to follow-up on...
At the end of the day

- This script really just generates more work...
  - Investigating lapsed access
  - Understanding complex titles
    - (and “helping” vendors to understand them, too…)
- But it’s more interesting work
Thank you!

James Van Mil
Collections and Electronic Resources Librarian
University of Cincinnati Libraries

james.vanmil@uc.edu