

Making up the Difference

Using custom reporting to identify inaccuracies in link resolver serial metadata

Serial Metadata Landscape

The difference between catalogue data and link
resolver data



Incorrect link resolver metadata = failed link resolving and the inability to connect users to your content

The Difference

Catalogue (PCC records)

National Law School of India Review

- ▷ Date1 2008, Date2 9999

Student Bar Review

- ▷ Date1 2002, Date2 2007

The Student Advocate

- ▷ Date1 1989, Date2 2001

Link Resolver (Publisher X Collection)

National Law School of India Review

- ▷ None (no access tied to latest title)

Student Bar Review

- ▷ 1988-2001, 2002-2007

The Student Advocate

- ▷ 1988-2001, 2002-2007,
2008-Current

PubDiff report process (simplified)

Compile SFX
thresholds

Take the threshold (coverage) statements for all active SFX object portfolios for a title and create a cumulative coverage statement representing all access points.

Compare
against ILS

Compare start and end dates of the cumulative coverage statements to Date1 and Date2 fixed fields in the corresponding catalogue record in the ILS.

Organize errors
by collection

Report titles where the SFX start date is too early, and/or the SFX end date is too late. Affected titles are organized by SFX Collection (target) so cleanup work can be prioritized.

Compile SFX thresholds

Title: *Student Bar Review*

Publisher X Journal Library

Available from 1988 volume: 1
issue:1 until 2001 volume: 13

Available from 2002 volume: 14
issue:1 until 2007 volume: 19

JSTOR Arts & Sciences XV

Available from 2002 volume: 14
until 2007 volume: 19 issue:2

Becomes:

1988-2007

Compare against ILS

SFX Cumulative Threshold

Start Date: 1988

End Date: 2007

Matchpoint used: SFX Object ID

The screenshot shows a library catalog record for 'Student bar review'. The record is displayed in a table format with the following fields:

Rec_Type	a	Bib_Lvl	s	TypeCtrl		Enc_Lvl	
Desc	a	Entrd	120105	Dat_Tp	d	Date1	2002
Date2	2007	Ctry	ii	Frequ	f	Regulr	r
Ser_Type	p	Phys_Med		Repr	o	Pub_Type	

Below the main table is a list of tags and their contents:

Tag	Ind.	Contents
245	00	Student bar review.
246	1	iCite as: aStud. bar. rev
246	1	Student b. rev
260		Bangalore : bStudent Bar Association, National Law School of India University
310		Semiannual, b2006-2007

Red arrows and boxes highlight the following fields in the screenshot:

- A red arrow points from the text 'Start Date: 1988' to the 'Date1' field (2002).
- A red arrow points from the text 'End Date: 2007' to the 'Date2' field (2007).
- A red box highlights the 'Date1' and 'Date2' fields.
- A red box highlights the 'Date1' field and its value '2002'.

Report: pubDiff.txt

Grand Total: 10549

Target ID	Target Name	Count	Priority
1000000000001505	EBSCOHOST_ACADEMIC_SEARCH_COMPLETE	990	<input type="checkbox"/>
111037301464000	FACTIVA	708	<input type="checkbox"/>
1000000000000630	EBSCOHOST_BUSINESS_SOURCE_COMPLETE	668	<input type="checkbox"/>
2550000000000007	ABI_INFORM_COLLECTION	463	<input type="checkbox"/>
4910000000000075	SPRINGER_NATURE_OA_JOURNALS	438	<input type="checkbox"/>
110976638852340	MISCELLANEOUS_FREE_EJOURNALS	418	<input type="checkbox"/>
1000000000001874	CRKN_SPRINGER_LINK_CURRENT	387	<input checked="" type="checkbox"/>
110976114395369	EBSCOHOST_MASTERFILE_PREMIER	290	<input type="checkbox"/>
111058563444000	PUBMED_CENTRAL_JOURNALS_FREE	284	<input type="checkbox"/>
111061245432000	DOAJ_DIRECTORY_OPEN_ACCESS_JOURNALS_FREE	217	<input type="checkbox"/>
1000000000001468	IEEE_XPLORE_ALL_CONFERENCE_SERIES	208	<input type="checkbox"/>
2400000000000007	ABI_INFORM_TRADE_AND_INDUSTRY	195	<input type="checkbox"/>
111021672379000	PUBLISHER_X_JOURNAL_LIBRARY	195	<input checked="" type="checkbox"/>
1000000000002296	CRKN_SPRINGER_LINK_ARCHIVE	193	<input checked="" type="checkbox"/>

Organize errors by collection

111021672379000	PUBLISHER_X_JOURNAL_LIBRARY	991042724308644	Canadian journal of criminology and corrections	2021-03-26	6821846	SFX start date too early: SFX-1970 Sirsi-1971^ SFX Holdings: 1970-1977^	<input type="checkbox"/>
111021672379000	PUBLISHER_X_JOURNAL_LIBRARY	991042726969234	Manitoba Law School journal	2021-03-26	6843020	SFX end date too late: SFX-1992 Sirsi-1965^ SFX Holdings: 1962-1965, 1992- ^	<input type="checkbox"/>
111021672379000	PUBLISHER_X_JOURNAL_LIBRARY	1000000000299016	The Georgetown law journal ... annual review of criminal procedure	2021-03-26	6836835	SFX start date too early: SFX-1971 Sirsi-2003^ SFX Holdings: 1971- , Most recent 1 year(s) not available^	<input type="checkbox"/>
111021672379000	PUBLISHER_X_JOURNAL_LIBRARY	1000000000324009	Student Bar Review	2021-03-26	7891777	SFX start date too early: SFX-1988 Sirsi-2002^ SFX Holdings: 1988-2007^	<input type="checkbox"/>
111021672379000	PUBLISHER_X_JOURNAL_LIBRARY	111021054530258	Benchmark	2021-03-26	9459651	SFX start date too early: SFX-1983 Sirsi-1984^ SFX Holdings: 1983-1993^	<input type="checkbox"/>

PubDiff report limitations

Multi-year issues

First Issue = 1988/1989

Catalogue Date1 = 1989

Catalogue Date 2 = 2001

Link Resolver Threshold =
1988-2001

Mismatch between CONSER
cataloguing rules and common
link resolver practices
incorrectly flags some titles as
problematic

Embargos

Last Issue = 2015

Catalogue Date2 = 2015

Link Resolver Threshold =
1992 - Most recent 6 year(s)
not available

All embargo statements are
assigned the value of 2021

Complex coverage statements

SFX threshold syntax
complexity can be difficult to
accommodate and can result in
inaccurate cumulative
coverage statements

**Using a cumulative coverage
statement as the comparator
eliminates the 1-1 connection
between problematic vendor
access points and date
misalignments**

PubDiff report advantages

Provides issue scope

Highlights vendors that consistently misrepresent title histories. Allows us to follow up with these vendors about a collection before working to manually correct data in the KB.

Can also identify issues in cataloguing contributed by consortium members to our shared ILS.

Allows for prioritized cleanup

Report organized so that priority collections/titles can be identified and worked on by Serials Cataloguing Specialists.

Potential for wider re-use

Some vendors and link resolver knowledgebases have access to quality PCC records (ex. OCLC, SFX (via MARCIt!)) that could potentially be leveraged to perform similar quality control.

Future planning

- ▷ Separate the PubDiff report from our e-journal catalogue record updating process
- ▷ Investigate potential to compare individual provider thresholds for a title against the corresponding catalogue record dates

Thanks!

For any follow up questions or comments, email ajsparli@ualberta.ca

Credits

- ▷ Thanks to Erika Banski, Scott Davies, Sam Popowich, Jeremy Hennig and Jim Hamilton who designed and implemented the PubDiff report
- ▷ Presentation template by [SlidesCarnival](#)