UIC UNIVERSITY LIBRARY
UNIVERSITY OF ILLINOIS AT CHICAGO

Measures of Journal Use: A Comparison of Vendor, Link-Resolver, and Local Citation Statistics

Sandra L. De Groote, Deborah D. Blecic, and Kristin E. Martin
CARLI Usage Statistics Forum
November 22, 2013

The Problem

- Use statistics provide a way for libraries to the assess value of electronic journal subscriptions, but what types of use should be examined and how can use statistics be quickly and easily collected?
  1. From the vendor/platform?
  2. From the link-resolver?
  3. From citations of research publications?

UIC UNIVERSITY LIBRARY
UNIVERSITY OF ILLINOIS AT CHICAGO

2

Vendor/Platform Statistics

- Most publishers/vendors are COUNTER-compliant (Counting Online Usage of NeTworked Electronic Resources)

- Advantages:
  – COUNTER allows for standardized reporting of use across different publishers and platforms.
  – Successful Full-Text Article Requests (SFTARs) provide an article-level metric of use

- Disadvantages
  – Statistics often need to be collected from many different vendors in an idiosyncratic manner
  – For journals available from multiple vendors, statistics must be tabulated and merged to get a complete SFTAR count
  – The platform design can affect usage counts so that there is inconsistent reporting
  – Standardized Usage Statistics Harvesting Initiative (SUSHI) should help with collection, but implementation has not been consistent

UIC UNIVERSITY LIBRARY
UNIVERSITY OF ILLINOIS AT CHICAGO

3

Link-Resolver Statistics

- Count the number of click-throughs to a journal title from a link-resolver service, including click-throughs from A-Z lists and MARC records
- Advantages
  – Easy to collect from single source
  – Use on different platforms is brought together and managed in knowledge base
- Disadvantages
  – Does not measure actual article use
  – Once user leaves link resolver, additional use is no longer counted
  – A lot of use may occur through alternate paths

UIC UNIVERSITY LIBRARY
UNIVERSITY OF ILLINOIS AT CHICAGO

4

Local Citation Data

- Reflects how many times a journal is cited by researchers at a specific institution
- Advantages
  - Use demonstrates clear value of content
- Disadvantages
  - Time-consuming to collect
  - Other uses, such as consultation, clinical use, student reference, etc., are not represented

5

The Questions

- Do link-resolver statistics positively correlate with COUNTER-compliant vendor SFTAR statistics?
- Do link-resolver statistics reflect use patterns seen in vendor statistics?
- Does local citation analysis demonstrate different patterns not reflected in vendor or link-resolver statistics?

6

Methodology: Data Collection (1)

- Study university (UIC) is large urban Research 1 university with 6 health sciences colleges, a large urban medical center, and 3 regional medical campuses.
- Journal list compiled from ERMS, Serials Solutions
  - Used HILCC (Hierarchical Interface to Library of Congress Classification) to identify journals in the health sciences
  - Identified list of 3496 current titles
  - List narrowed to 2619 after removing titles missing use data
- Collected vendor COUNTER-compliant SFTARs for 2010
  - Went to individual website and downloaded Excel reports from over 20 providers
  - All were COUNTER-compliant except *MD Consult*

7

Methodology: Data Collection (2)

- Obtained link-resolver data for titles for 2010
  - Click-through counts through the OpenURL resolver
  - Click-through counts through the e-journal A-Z list
  - Click-through counts through the library catalog MARC records
- Collected citation data for 2010
  - Searched ISI's *Web of Science* by author affiliation
  - Although study focused on health sciences, citations for all researchers affiliated with the university were included

8

Matching and Cleaning the Data

- Matched link-resolver data to citation data
  - This was easy: 1 line of data per journal
- Matched link-resolver and citation data to COUNTER SFTARs
  - This was the hard part
- Challenges
  - A title might be listed in up to 10 different databases in ERMS
  - Automated matching of data by ISSN, but titles still needed to be examined by hand
  - Matching process was not 1 to 1
  - Some platforms lacked vendor statistics (e.g., open-access)
  - Some databases were listed multiple times in ERMS, but only once in platform statistics (e.g., one title could be in multiple *EBSCOhost* databases)
  - Errors in knowledge base or changes in providers could lead to missed matches
  - Unmatched data need to be manually searched by title

UIC UNIVERSITY
UNIVERSITY OF ILLINOIS AT CHICAGO LIBRARY

9

Title coding scheme

Category	Description	Decision	Total Titles
Perfect Match	Each source in ERMS matches corresponding source for provider statistics	Used in analysis	1853
Perfect Match Plus Open Access	Each source in ERMS matches corresponding source for provider statistics AND title also available on OA platform	Used in analysis	392
Open Access Only	No vendor statistics available because only on OA platforms	Title removed	467
Perfect Match + Extra Data	Each source in ERMS matches corresponding source for provider statistics AND received additional provider statistics with no match in ERMS	Used in analysis	311
Perfect Match + Extra Data + OA	Each source in ERMS matches corresponding source for provider statistics AND title also available on OA platform AND received additional statistics with no match in ERMS	Used in analysis	63
Missing Data	Missing data in one or more sources in ERMS that was not found through individual title searching	Title removed	408
			3494

UIC UNIVERSITY
UNIVERSITY OF ILLINOIS AT CHICAGO LIBRARY
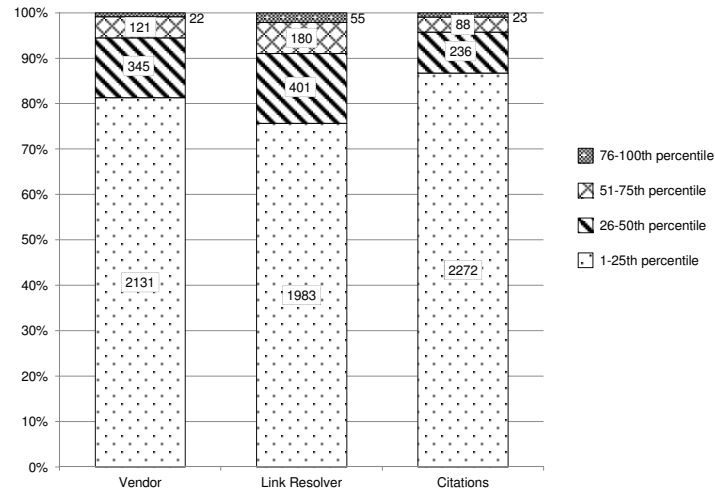
10

5

Statistical Analysis

- Spearman rank order

  – Looks only at the rank of use between the three use factors and does not take into consideration the actual number of uses

  – Best for non-parametric data

11

Results- Descriptive Statistics for the Three Datasets

	Vendor COUNTER Successful Full-Text Article Requests	Link-resolver Click – Through Statistics	Local Citations
Counts at the 25th Percentile mark of a ranked title list (663 titles)	68	26	0
Median of all data	283	81	3
Counts at the 75th Percentile mark of a ranked title list (1302 titles)	757	215	13
Highest Count for Each Measure	71,326	11,761	1,784

N=2,619

12

6

Number of Titles Per Usage Quartile

Legend:
- 76-100th percentile
- 51-75th percentile
- 26-50th percentile
- 1-25th percentile

Bar values:
- Vendor: 2131, 345, 121, 22
- Link Resolver: 1983, 401, 180, 55
- Citations: 2272, 236, 88, 23

13

Results – Correlation Coefficients

Data Subsets	No. of Titles	Link Resolver/ Vendor Spearman's Correlation	Link Resolver/ Citation Spearman's Correlation	Vendor/ Citation Spearman's Correlation
All Journals	2619	.843	.752	.726
1-25th Percentile of titles (Link-resolver use ≤26)	663	.454	.392	.311
25th-75th Percentile of titles (Link-resolver use ≥27-≤214)	1302	.563	.408	.455
75th -100th Percentile of titles (Link-resolver use ≥215)	654	.586	.511	.598
25th -100th Percentile of titles (Link-resolver use ≥27)	1956	.703	.581	.622

All correlations were significant at p<.01.

14

7

Scatter Chart 25th to 75th percentile

15
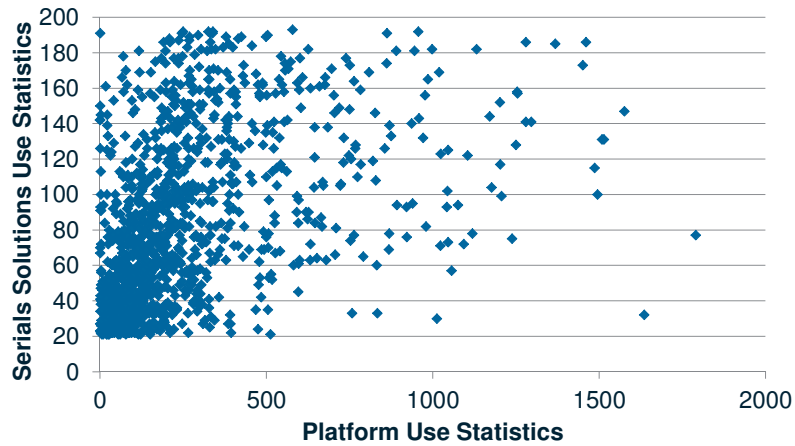
Sample of titles with vendor use higher than ERMS,ERMS higher than citations

Title	Link-resolver	Vendor	Citations
Academic Emergency Medicine	795	992	37
Annals of Internal Medicine	4,161	7,581	155
Circulation	2,545	14,134	472
Develop. & Comparative Immunology	12	4,069	0
Diabetes Care	1905	730	190
Journal of Biological Chemistry	2,280	45,526	1,784
Journal of Family Practice	808	1,365	8
Nature	4,420	41,428	1,015
PNAS	2,830	35,944	1,372
Science	5,706	40,155	997

16

Sample of titles with anomalous results

Title	Link-resolver	Vendor	Citation
Amer. Journal of Clinical Nutrition	1,295	0	140
Biochemistry (Easton)	935	0	364
Infection and Immunity	142	0	122
Internet J. of Peds. & Neonatology	0	301	0
J. For Specialists in Ped. Nursing	0	799	0
Critical Care	665	19	21
Diabetes Care	1,905	730	190
Journal of Studies on Alcohol	520	496	38
British Journal of Ophthalmology	508	512	35
Gerontologist	590	608	26

17

Anomalous Use Patterns

- Anomalous use patterns can point to problems with links or vendor use statistics accounts

- Anomalies can easily be discovered using Excel by ratios testing or subtraction to discover where the data for a title show unexpected results

18

"80/20" Calculations

- Link-resolver Data : 80/29
  - 0.8% of titles had no click-throughs
- Vendor Data: 80/24
  - 4.5% of titles had no SFTARs
- Citation Data: 80/17
  - 27% of titles received no citations in 2010
- Use is the most concentrated for citations

UIC UNIVERSITY
UNIVERSITY OF ILLINOIS AT CHICAGO LIBRARY

19

Conclusions

- Link-resolver data correlates with vendor use statistics, and thus can aid with collection development decisions, especially if vendor data cannot be obtained.

- Link-resolver data can be used to identify journals that need further evaluation for retention decisions. If a journal has a high-link resolver count, no further evaluation is necessary.

- Citation data, which identifies titles used in faculty research, displays a similar pattern of use compared with other measures.

UIC UNIVERSITY
UNIVERSITY OF ILLINOIS AT CHICAGO LIBRARY

20

Areas for further study

- Repeat the study in the humanities, social sciences, and basic sciences to see if use patterns differ.

- ERMS and vendor data were not always proportional in use, the ratio of ERMS to vendor data varied widely, even within expected parameters. Is that effect related to the platform the journal is on? The type of journal? Or is it related to user behavior?

UIC UNIVERSITY
UNIVERSITY OF ILLINOIS
AT CHICAGO LIBRARY

21

Next Steps at UIC

- UIC subscribed to *360 COUNTER*, to assist with collecting and merging vendor statistics

  - First year collected was 2012
  - Still clean-up/matching issues for the subscribing library to address
  - Data out is only as good as the data in
  - Need to evaluate SUSHI option

- Evaluate options to calculate cost-per-use

UIC UNIVERSITY
UNIVERSITY OF ILLINOIS
AT CHICAGO LIBRARY

22

Sample JR1 report from *360 COUNTER*

Journal Holdings Report (JR1)

Month (remove): Jan 2012 to Dec 2012   Source Counter Report: JR1, JR2   LibraryCode (remove): IAY   Uses (remove): IS GREATER THAN 0

Authority Title	Title Name	Provider/Platform	ISSN	Uses	Cost per Use
Acta radiologica (1987)	Acta radiologica (1987) total:			119	-
Acta seismologica sinica	Acta Seismologica Sinica	MetaPress	1000-9116	4	-
	Acta Seismologica Sinica	Springer-Verlag	1000-9116	4	-
	Acta seismologica sinica total:			8	-
Acta sociologica	Acta Sociologica	EBSCOhost	0001-6993	13	-
	Acta Sociologica	HighWire Press	0001-6993	36	-
	Acta Sociologica	JSTOR	0001-6993	239	0.07
	Acta sociologica total:			288	0.06
Acta theriologica	Acta Theriologica	MetaPress	0001-7051	30	-
	Acta Theriologica	Springer-Verlag	0001-7051	25	-
	Acta theriologica total:			55	-
Acta tropica	Acta Tropica	Elsevier	0001-706X	136	-
	Acta tropica total:			136	-
Acta veterinaria scandinavica	Acta veterinaria Scandinavica	EBSCOhost	0044-605X	2	-
	Acta veterinaria scandinavica total:			2	-
Acta zoologica (Stockholm)	Acta Zoologica	Wiley-Blackwell	0001-7272	23	-
	Acta Zoologica (Stockholm)	EBSCOhost	0001-7272	5	-
	Acta zoologica (Stockholm) total:			28	-

UIC UNIVERSITY
UNIVERSITY OF ILLINOIS AT CHICAGO **LIBRARY**

23

Thank you!

Questions?
Sandy De Groote sgroote@uic.edu
Deb Blecic: dblecic@uic.edu
Kristin Martin: kmarti@uchicago.edu

UIC UNIVERSITY
UNIVERSITY OF ILLINOIS AT CHICAGO **LIBRARY**

24