Assessment Metrics

December 14, 2011

Speakers: Oliver Pesch, Robin Kear, Tim Jewell and Corey Murata

Agenda

• What’s new with Release 4 of the COUNTER Code of Practice
• How the COUNTER and SUSHI XML Schemas are affected
• Facilitating implementation with the COUNTER-SUSHI Implementation Profile
• Ensuring Compliance through a more robust audit
• Timing

COUNTER and SUSHI: What’s new with Release 4 of the COUNTER Code of Practice

NISO Webinar: Assessment Metrics
December 14, 2011
Oliver Pesch
Chief Strategist, E-Resources
EBSCO Information Services

Release 4 of the COUNTER Code of Practice

What’s New
Single Code of Practice

One Code of Practice
- Covers books, databases, journals and multimedia collections
- Easier to manage for COUNTER
- Easier to manage for content providers – particularly those offering multiple resource types.

New Required Reports

Multimedia Report 1
- Reports on use of items in multimedia collections
- Applies to platforms like ArtStor

Consortium Report 3
- The multimedia report for consortia

Journal Report 5
- Usage of journal articles broken out by year of publication
- No longer optional

New Data Elements

DOI and Proprietary Identifier
- Improves identification of items listed in the reports
- Improves matching usage to the correct items in a knowledge base
- Involves coordination with KBART so the same identifier is on both the vendor’s Title List and their COUNTER report
- Applies to most COUNTER reports

Institutional Identifier
- Utilizes the NISO I2 as the standard identifier for the institution (being implemented as ISNI for organizations)
- Improves comparative analysis– particularly for usage processed by consortium
- Applies to all reports
New Data Elements

Result Clicks and Record Views
• Result Click counts each time a user clicked a link from the search result to the detailed view, a link resolver, etc.
• Record View counts the abstract/detailed records explicitly viewed from the database
• Both indicate an explicit expression of interest in the result
• Provides a more direct measure of database usage, in particular for A&I databases
• Applies to Database Reports 1 and 3

New Metric Types
Gold Open Access
• Covers articles accessible without subscription
• Included as a separate count in Journal Report 1 and Journal Report 1a

Access Denied: content item not licensed
• In addition to Turnaways (simultaneous user limit)
• Records a user’s attempt to access full text for which their institution does not have a subscription
• Applies to “Access Denied” reports:
  o Journal Report 2
  o Database Report 2
  o Book Reports 3 and 4
• Useful input for collection development
Metric Types Removed

Sessions
- Federated Search and Discovery products have made “Sessions” an unreliable measure of use
- Reports affected:
  - Database Report 1
  - Database Report 3
  - Consortium Report 2

COUNTER and SUSHI XML Schemas

COUNTER Schema
- Minor changes only
- ItemIdentifier, ItemPublisher now optional elements
- PubYr attribute (used by Journal Report 5) now defined as a year.
- PubYrFrom and PubYrTo attributes added to allow Journal Report 5 to report ranges of years.

COUNTER Enumeration Schema
- Contains the controlled vocabulary for elements like ItemDataType, Category, MetricType, etc.
- Added new values as required by the new and updated reports
COUNTER and SUSHI XML Schemas

SUSHI Schema
- No changes required

COUNTER-SUSHI Implementation Profile

Facilitating adoption of SUSHI

COUNTER-SUSHI Implementation Profile

Why
- SUSHI and COUNTER tend to be abstract
- Abstraction leaves the opportunity for concepts to be interpreted differently
- Different interpretations leads to inconsistencies
- Examples:
  - Different authentication methods that require custom development
  - Wrong capitalization on controlled data values
  - Including “Totals” in the XML
  - Including invalid ISSNs (e.g. “N/A”)

COUNTER-SUSHI Implementation Profile

Authentication for SUSHI Client
- IP and use of requester and customer IDs allowed
- Use of SOAP extensions not allowed
COUNTER-SUSHI Implementation Profile

Report Names
• Must exactly match the value in the “Name” column on the SUSHI Reports Registry

Date Ranges for SUSHI Request
• Format must be yyyy-mm-dd
• Begin date must be first of the month
• End date must be last day of the month
• End date must be later than begin date

SUSHI Exceptions
• Standard error numbers must be used for standard error conditions (see Table 17 in the SUSHI standard)

SUSHI Server Registry
• SUSHI Server must be registered on SUSHI Server Registry

Data to Return in the COUNTER Report
• Only the months of usage asked for
• Include exception if some or all usage not available
• Only include items appropriate to the report (e.g. don’t return book usage in a Journal Report)
• Do not include totals for the report or an item

Data Elements Values
• Use registered values with proper casing for enumerated values (Item Data Type, Category, Metric Type, etc.)
• Use values appropriate for the report
COUNTER-SUSHI Implementation Profile

Designed for use by implementers

- Prescriptive in nature
- Lots of tables and charts

---

COUNTER Audit

Monitoring compliance

---

Appendix A:
Summary of Data Element Usage by Report

Table 5 summarizes the possible enumeration values that are allowed for each report.

<table>
<thead>
<tr>
<th>Report Name</th>
<th>Description</th>
<th>Identifier Type</th>
<th>ItemDataType</th>
<th>Category</th>
<th>MetricType</th>
</tr>
</thead>
<tbody>
<tr>
<td>BR1</td>
<td>Number of Successful Title Requests by Month and Title</td>
<td>Online_ISBN Print_ISBN DOI Proprietary</td>
<td>Book</td>
<td>Requests</td>
<td>ft_ps ft_pdf ft_image ft_total</td>
</tr>
<tr>
<td>BR2</td>
<td>Number of Successful Section Requests by Month and Title</td>
<td>Online_ISBN Print_ISBN DOI Proprietary</td>
<td>Book</td>
<td>Requests</td>
<td>ft_ps ft_pdf ft_image ft_total</td>
</tr>
<tr>
<td>BR3</td>
<td>Access Denied to Content Items by Month, Title, and Category</td>
<td>Online_ISBN Print_ISBN DOI Proprietary</td>
<td>Book</td>
<td>Requests</td>
<td>ft_ps ft_pdf ft_image ft_total</td>
</tr>
<tr>
<td>BR4</td>
<td>Access Denied to Content Items by Month, Service, and Category</td>
<td>Proprietary</td>
<td>Platform</td>
<td>Access_denied</td>
<td>turnover no_licenseculture</td>
</tr>
</tbody>
</table>
Timing

Trial use, comments and final implementation

January 20, 2012
• Last day for comments on the draft of release 4 of COUNTER Code of Practice and the COUNTER-SUSHI Implementation Profile

March/April 2012
• Final publication of Release 4 of COUNTER CoP and the COUNTER-SUSHI Implementation Profile

December 31, 2013
• Deadline for compliance with Release 4 of the COUNTER CoP

References

COUNTER CoP R4
http://www.projectcounter.org/code_practice.html
COUNTER-SUSHI Implementation Profile
COUNTER Schema Date Element Values
http://www.niso.org/workrooms/sushi/values_R4
KBART
http://www.uksg.org/kbart
NISO SUSHI website
http://www.niso.org/workrooms/sushi
SUSHI Reports Registry
http://www.niso.org/workrooms/sushi/reports_R4
SUSHI Server Registry
https://sites.google.com/site/sushiserverregistry/

Thank you
Why measure journal impact?

- Collection development, value for money
- Advising faculty and students of the best journals
- Journal economics, pricing over time
- Journals use impact to promote themselves

Databases that Measure Journal Impact

Journal Citation Reports (Thomson Reuters; a standalone database that is also part of ISI Web of Knowledge)

- There is currently no overarching tool across vendors.
- Potentially, any database with citations could create bibliometric measures.
- Both companies have given their data to research labs to create new metrics that are freely available online from the labs and are also included in the company's products for subscribers.

Scopus (Elsevier)

Journal Citation Reports (Thomson Reuters)

- 5,900 journals in the areas of science and technology and of 1,700 journals in the social sciences.
- Years included vary by subscription.

Journal Citation Reports (Thomson Reuters)

- Impact Factor (Thomson Reuters; in Journal Citation Reports) 1961
- Eigenfactor (uses Thomson Reuters citation data) 2007
- Article Influence (uses Thomson Reuters citation data) 2007
Impact Factor

The ratio of the number of citations to the previous 2 years of the journal divided by the number of articles in those years — this is essentially the average number of recent citations per article.

- First proposed by Dr. Eugene Garfield in 1955 while at ISI
- First published in 1961 in Science Citation Index
- Oldest, most well-known measure with the most criticism
- Completely fee-based measure only available through Journal Citation Reports (Thomson Reuters)

Journal Citation Reports

- the journal’s immediacy index, the number of citations that year to articles published the same year
- the journal’s citing half life, the median age of the articles that were cited by the articles published in the journal that year
- the journal’s cited half life, the median age of the articles in the journal that were cited by other journals during the year

Eigenfactor

Created by the Metrics Eigenfactor Project, a bibliometric research project conducted by Professor Carl Bergstrom and his laboratory at the University of Washington.

- Eigenfactor Score is essentially a ratio of number of citations to total number of articles. (similar to Impact Factor)
  However, the score:
  - Counts citations to journals in both the sciences and social sciences.
  - Eliminates self-citations. Every reference from one article in a journal to another article from the same journal is discounted.
  - Weights each reference according to a stochastic measure of the amount of time researchers spend reading the journal

- Uses Thomson Reuters JCR citation data and social network theory.
**Eigenfactor**

Eigenfactor scores are scaled so that the sum of the Eigenfactor scores of all journals listed in Thomson's JCR is 100.

The top thousand journals, as ranked by Eigenfactor score, all have scores above 0.01.

Freely available at [http://www.eigenfactor.org](http://www.eigenfactor.org) AND/OR through the Journal Citation Reports subscription database available from Thomson Reuters.

**Article Influence**

Created out of the same project by Professor Bergstrom's team and uses Thomson Reuters JCR citation data.

It is the journal's Eigenfactor Score divided by the fraction of articles published by the journal.

That fraction is normalized so that the sum total of articles from all journals is 1.

**Example**

- A score greater than 1.00 indicates that each article in the journal has above-average influence.
- A score less than 1.00 indicates that each article in the journal has below-average influence.

Freely available at [http://www.eigenfactor.org](http://www.eigenfactor.org) AND/OR through the Journal Citation Reports subscription database available from Thomson Reuters.

Eigenfactor™ Score: A measure of the overall value provided by all of the articles published in a given journal in a year.

Article Influence™ Score: A measure of a journal's prestige based on per-article citations and comparable to Impact Factor.

<table>
<thead>
<tr>
<th>Journal Name</th>
<th>Eigenfactor™ Score</th>
<th>Article Influence™ Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADVANCES IN BIOENGINEERING / BIOTECHNOLOGY</td>
<td>35.10</td>
<td>0.033351</td>
</tr>
<tr>
<td>ISSN: 0724-6245</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHEMICAL ENGINEERING SCIENCE</td>
<td>14.65</td>
<td>0.044504</td>
</tr>
<tr>
<td>ISSN: 0017-5209</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Using citations from nearly 18,000 titles from 5,000 international publishers, the Journal Analyzer goes back to 1996.

- Developed by Félix de Moya-Anegon, University of Granada, SCImago Journal Rank (SJR) is a prestige metric.
- The subject field, quality and reputation of the journal has a direct effect on the value of a citation.

SJR

Is weighted by the prestige of the journal
‘Shares’ a journal’s prestige equally over the total number of citations in that journal
Normalizes for differences in citation behavior between subject fields
Is international
**SJR**

- SJR starts with a basic per peer-reviewed document measure that is the number of citations received by a journal divided by the number of citations in that journal for the three previous years.
- SJR then weights the citations with a prestige metric that normalizes across the citation behavior of different journals.


**SNIP**

Created by Henk F. Moed of Leiden University, Source Normalized Impact per Paper (SNIP) measures contextual citation impact by weighting citations based on the total number of citations in a subject field.

Freely available at [http://www.journalindicators.com](http://www.journalindicators.com) and [http://www.journalmetrics.com](http://www.journalmetrics.com) AND/OR through the Scopus subscription database available from Elsevier

**Example**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Annual Review of Immunology</td>
<td>16.204</td>
<td>179</td>
<td>26</td>
<td>81</td>
<td>4,256</td>
<td>3,406</td>
<td>81</td>
<td>19,459</td>
<td>177.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Cell</td>
<td>12.802</td>
<td>231</td>
<td>135</td>
<td>137</td>
<td>18.726</td>
<td>22.308</td>
<td>1,023</td>
<td>31.46</td>
<td>33.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Annual Review of Biochemistry</td>
<td>11.402</td>
<td>215</td>
<td>31</td>
<td>31</td>
<td>4,666</td>
<td>3,026</td>
<td>91</td>
<td>20,25</td>
<td>150.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 CA Cancer Journal for Clinicians</td>
<td>111,21</td>
<td>24</td>
<td>34</td>
<td>116</td>
<td>1,048</td>
<td>4,772</td>
<td>62</td>
<td>76.78</td>
<td>54.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Nature Genetics</td>
<td>10.400</td>
<td>311</td>
<td>327</td>
<td>962</td>
<td>7,048</td>
<td>17,407</td>
<td>638</td>
<td>29.07</td>
<td>21.59</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example

Summary

Databases:
• Journal Citation Reports (Thomson Reuters)
• Scopus (Elsevier)

Measures:
• Impact Factor
• Eigenfactor
• Article Influence
• SJR
• SNIP

Tools:
• Journal Citation Reports (Thomson Reuters)
• Journal Analyzer (Elsevier)

http://pitt.libguides.com/bibliometrics

LibGuide: Citation Searching and Bibliometric Measures

• A discussion on topics such as the h-index, Eigenfactor, Impact Factor, Journal Citation Reports, and other tools.
• Further reading
• Citation count instructions

Thank you!

Contact:
Robin Kear, Reference/Instruction Librarian, University of Pittsburgh
rlk25@pitt.edu
Using Journal Metrics for Decision-Making

Tim Jewell
Director, Information Resources and Scholarly Communication

Corey Murata
Information Resources and Collection Assessment Librarian

A “Culture of Assessment”

• Assessment and Planning Program
  – Steve Hiller, Director
  – Assessment and Metrics Team

• Surveys
  – Triennial faculty & student surveys since 1992
  – In-Library Surveys
  – Diversity and Organizational Climate, etc.

• Qualitative work
  – Usability
  – User-centered design, etc.

Collection Assessment

• Collection Management Services unit responsibility
• Have been looking at
  – Circulation data
  – Student demographics
  – WebBridge log analysis
  – Serials evaluation

STUDENT DEMOGRAPHICS

Data on the changes in campus demographics can help to make allocation decisions based on shifting needs and demands.

Data on number of majors by department is put into a database for selectors or administrators to easily view on the web. Data is also extracted and analyzed to show departments with high growth or decline in number of majors.
**Circulation Snapshot**

A snapshot report of all monographic items checked out is downloaded once every quarter from the ILS. Information includes title, call number, language and patron department and status. This data is loaded into a database that we use to produce web reports and run custom queries for additional analysis.

Custom queries provide additional insight into factors such as circulation by language.

The web report allows selectors to identify patterns of circulation behavior among the faculty and students in their department.

**Serials Worksheet**

By pulling together data from various spreadsheets and reports into one database, we were able to create easy-to-read PDFs customized for each fund manager. These handouts contain payment information, usage data and other bibliometric values, cancellation restrictions from the related license documents and perpetual access as available through Portico.

**WebBridge Log Analysis**

By analyzing our WebBridge link resolver logs, we are able to identify patterns of user behavior that is not available from vendor supplied usage statistics.

A script is used to parse the log file data and upload into a database.

Sample of WebBridge Log File

<table>
<thead>
<tr>
<th>Resource Name</th>
<th>Origin</th>
<th>Placement Date</th>
<th>Time Clicked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic search complete</td>
<td>ISI includes Web of Science (V 1.0)</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Academic search complete</td>
<td>ISI includes Web of Science (V 1.0)</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Academic search complete</td>
<td>OCLC Worldcat Local</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Academic search complete</td>
<td>OCLC Worldcat Local</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Science Direct</td>
<td>ISI includes Web of Science (V 1.0)</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Science Direct</td>
<td>ISI includes Web of Science (V 1.0)</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Science Direct</td>
<td>ProQuest (All)</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Science Direct</td>
<td>ProQuest (All)</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Science Direct</td>
<td>Gale</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Science Direct</td>
<td>Gale</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Science Direct</td>
<td>Google</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Science Direct</td>
<td>Google</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Science Direct</td>
<td>Google</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Science Direct</td>
<td>Google</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Science Direct</td>
<td>Ovid</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Science Direct</td>
<td>Ovid</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Science Direct</td>
<td>PubMed</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
<tr>
<td>Science Direct</td>
<td>PubMed</td>
<td>Resolution server</td>
<td>06012008</td>
</tr>
</tbody>
</table>

**Collection Development Context**

- **Financial**
  - Chronic budget pressure
  - Increased importance of accountability and evidence
- **Market**
  - E-journal “packages”
  - Licensing issues
    - Cancellation limits
    - Long-term access/archiving
  - Publisher consolidations
- **Organizational**
  - the end of “rugged individualism?”
2009 Financial Crisis and UW Libraries

- UW budget cuts
  - 13.5% reduction (1.5% rescission + 12% cut)
  - Branch libraries closed, staff reductions, etc.
  - Materials budget cuts (~20% purch. power loss)
- Complex, 3-campus system serving wide range of research, teaching (biomed to int’l studies)
- UW Resources Budget: 79% Serials/21% Monographs
- Shift to e-access (UW now @ ~70%)

Complex Serials Review Process

- Phase One
  - Reduction targets set, subscription lists generated
  - Departments contacted, cancellation lists identified
  - User input on proposed cancellations via survey
- Phase Two (overlapped with end of Phase One)
  - Publisher negotiations via consortia
  - Cancellation implications for “bundles” assessed
- For both, supplemental data made available:
  - Article downloads, cost per use
  - ISI Impact Factor, Eigenfactor
  - UW-authored publications & citations in published work

Phase One: Fund-Level Review

Using Access to Create a Serials Decision Database

Presented at the Charleston Conference, November 6, 2009
Hana Levay
Information Resources Librarian
University of Washington

Project Steps

- Gathered data (cost, use, impact factor, package info, license terms, etc.)
- Exported table, added cost per use column, imported back to Access
- Created final table
- Generated spreadsheets by fund code
  - .xls format
  - .pdf (printable) format
Phase 1 Problems and Issues

- Too many metrics
- Interpretation and comparison
- Which titles are “in play” for cancellation?
Phase 2: Data for Bundle/Package Value Review

- With and without “unsubscribed” titles
- Article downloads
  - Cost per use
  - “Adjusted” cost per use
- ISI Impact Factor (Sciences & Social Sciences)
- Eigenfactor (ditto, mostly)
- UW-authored publications (ditto)
- UW citations in published work (ditto)

Bundles We Compared

- American Chemical Soc.
- American Inst. Of Physics
- Cambridge Univ. Press
- Elsevier
- Johns Hopkins Univ. Press (Project Muse)
- Nature, Science
- Oxford University Press
- Sage
- Springer/Kluwer
- Taylor & Francis
  - No UW “bundle” purchased, so individual titles can be freely cancelled
- University of Chicago Press
- Wiley/Blackwell

Profit/Non-Profit Journal Publisher Pricing Comparisons (source: JournalPrices.com)

<table>
<thead>
<tr>
<th></th>
<th>Non-Profit</th>
<th>For-Profit</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Journal Price</td>
<td>$642.78</td>
<td>$1,837.88</td>
<td>2.86</td>
</tr>
<tr>
<td>Mean Price per Article</td>
<td>$7.98</td>
<td>$24.20</td>
<td>3.03</td>
</tr>
<tr>
<td>Mean Price per Citation</td>
<td>$12.73</td>
<td>$39.55</td>
<td>3.11</td>
</tr>
</tbody>
</table>

Evaluating Bundles: First Steps

- For each metric:
  - Determine total for each bundle
  - Determine total across bundles
  - Divide expenditures by metric for each bundle to determine a “Cost per” amount
  - Divide total expenditure for all bundles by metric total to determine average “Cost per” amount
Bundle-level Comparisons

- “Cost per” figures across bundles
- “Cost per” figures to averages (“mean”)
  - What is “estimated value” of a bundle, using overall averages as a baseline?
  - How much more or less than estimated value do we pay?
- Compare relative standings across metrics

Example 1: Cost Per Use (CPU)

- In 2008, UW Paid $4.4M for 2,686,489 article downloads
  - Avg. CPU = $1.64
  - Est. value of a bundle w/100K downloads = $164K
- % of estimated value paid
  - If paid $100K, $100K/$164K = 61% (“good”)
  - If paid $200K, $200K/$164K = 122% (“bad”)

Example 2: Eigenfactor

- In 2008 UW Paid $4.4M for 14 bundles that provided a total of 63.9 “Eigenfactors”
  - Avg. Cost/Eigenfactor = $88,955
  - Est. value of a bundle w/10 “eigenfactors” = $889.5K
- % of estimated value paid
  - If paid $500K, $500K/$889.5K = 57% (“good”)
  - If paid $1M, $1M/$889.5K = 113% (“not good”)

Applied Same Process To:

- Cost per use
- “Adjusted” cost per use
- Impact Factor
- Eigenfactor
- UW Faculty Publications
- UW Faculty Citations
- Average % of value paid
Bundle Averages: 2008

% of Value Paid

Practical Implications?
• Confirmed informal assessments
  – Value of some bundles was high
    • Retained
    • Disallowed individual title cancellations
  – Value of others was low
    • Talking points with publishers
    • One bundle dropped
    • Focus on others for future renegotiation

Further Questions
• Assess bundles first to determine which titles are “in play” for cancellation?
• Disciplinary Differences and Metrics?
• Metric inter-correlations
• Replication/validation
• “Local” Eigenfactor?
• Variation or Stability over time
  – use 5 years of data?
• Construct a composite index using selected metrics?