Does COUNTER tell the whole story?

Case-by-case examples demonstrating the limitations of COUNTER, and suggestions for alternative evaluation metrics

CARLI Spring Forum on Collections Data Analysis and Maintenance
Governors State University, April 28th, 2017

Jonathan Shank
Acquisitions & E-Resources Librarian
Northwestern University Galter Health Sciences Library
Disclaimer
Institutional Context

Galter Health Sciences Library

• Serves Northwestern's Feinberg School Medicine in Chicago
• Administratively separate from University Library in Evanston
  - Cost sharing with Evanston on big deal agreements
  - Separate standalone subscriptions and a medical specific collection
• Member of CARLI, but not part of I-Share or union Voyager catalog
• Entire NU system migrated to Alma in Summer of 2015
• Galter maintains custom Primo front-end
• Currently in transitional phase for handling of COUNTER
  - No ERMS or usage client, efforts currently focused on JR1 stats
  - Usage functionality coming to Alma this summer
COUNTER usage statistics

What works well

• Standard format, impressive data set and BIG numbers
• “Consistency” across vendors
• Ease of utilizing for CPU analysis
• Increasing compliance among vendors
• Growing interoperability
• Iterative improvements with each new release
COUNTER usage statistics

What works well

Active and engaged community of librarians, publishers and vendors.
COUNTER usage statistics

What doesn’t work so well

• Merging multiple providers and platforms, unless you have an aggregator client (i.e. Ustat, 360 Resource Manager, CORAL, etc)
• Manual retrieval of reports
  - Still necessary despite major improvements from SUSHI
  - Login credentials must be stored & maintained, difficult with shared licenses
• Issues with accuracy and title consistency with historical titles and title changes, splits and merges
  - Stats may be inaccurate or useless as a result
COUNTER usage statistics

What doesn’t work so well

• Occasional issues with accuracy, compliance and reliability
• Overlapping accounts, IP ranges and multiple access points can inflate or deflate numbers
• Lack of distinction by location, school department, or affiliation
• Not available for some resources
COUNTER usage statistics

What doesn’t work so well

Individual usage is a relatively flat or static indicator of impact and value.

“Statistics are a measurement of users’ actions that we try to correlate to their intentions.”

Oliver Pesch, EBSCO Publishing
Specific Examples
Demonstrating the limitations of COUNTER
Example 1: Inflated numbers

Numbers can be inflated by a publisher’s interface & platform design

- Some platforms load HTML full text automatically, if user clicks PDF it can be counted twice
- Some linking mechanisms like CrossRef allow publishers to choose linking level, i.e. link to TOC, abstract, html, pdf
- COUNTER is continuously working to improve and resolve these issues
- Publisher interference, or at the very least, optimization for high stats, still possible
Example 2: IP issues

Incorrect IP information can distort figures

- On the vendor side, most usage in COUNTER reports is ultimately attributed to accounts based on IP addresses.
- According to a recent study/audit: 58% of IPs held by publishers to authenticate libraries are wrong (Spence, PSI Ltd).
Example 3: Problems distinguishing locations

COUNTER still has limitations with location or account specific reporting

- IPs often overlap between departments, schools and campuses, making usage indistinguishable by location
  - NU has campuses in Evanston, Chicago and Qatar with overlapping IPs
  - Content at NU is licensed by several different entities for different groups of users
- Accounts themselves also have overlap in locations and access entitlements, which are lumped together in COUNTER

“There is no single way [outlined in the COUNTER code of practice] for providers to categorize usage transactions to capture reporting by subsets.”

- Project COUNTER
Example 3: Problems distinguishing locations

Overview of NU’s Elsevier landscape
Example 4: Lack of context or normalization

Not all usage is created equal, but it’s treated equally

*Undergraduate student padding out works cited for English 101 paper*

Vs.

*Faculty conducting research for major grant or high impact publication*

Usage and information-seeking behaviors may vary widely by discipline, research area, or department
Example 5: False negatives

### Journal of Dermatological Science

<table>
<thead>
<tr>
<th>Journal</th>
<th>Publisher</th>
<th>Platform</th>
<th>Journal DOI</th>
<th>Proprietary Identifier</th>
<th>Print ISSN</th>
<th>Online ISSN</th>
<th>Reporting Period Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal of Dermatological Science</td>
<td>Elsevier</td>
<td>Default</td>
<td></td>
<td></td>
<td>0923-1811</td>
<td>0923-1811</td>
<td>1</td>
</tr>
</tbody>
</table>

- Journal is licensed by Galter Library through Elsevier’s ClinicalKey
  - Showed only 1 full text download in ClinicalKey’s 2016 JR1
- Citation analysis indicated journal was cited 46 times by NU scholars in same time period, obvious discrepancy
- Title is also available through NUL’s ScienceDirect Freedom Collection
  - 397 full text downloads in ScienceDirect’s 2016 JR1
Alternative usage metrics
Substitutes and supplements for COUNTER
Alternative usage metrics

Proxy logs

• Pros
  - Data is potentially stored in one place with a single access point
  - Possibility to capture user affiliation, domain or location
  - Integration with Google Analytics or other log analysis tools
• Cons
  - Initial set up is manual, and can be complicated
  - Some programming knowledge may be required
  - Not all traffic goes through proxy (on campus, VPN, etc.)
  - Not all institutions have a single proxy server
Alternative usage metrics

Link resolver logs, stats and analytics

• Pros
  - Can be much easier to retrieve, depending on your resolver
    • Alma has some functionality built in to Analytics, more coming with next release
  - Generally found to correlate closely with COUNTER stats
  - Potential to capture user affiliation, domain, and/or location
• Cons
  - Manual setup may be required
  - Does all of your traffic really go through the link resolver?
    • Galter routes PubMed traffic back to customized resolver
Alternative usage metrics

Citation data

• Pros
  - Identifies usage based on actual research output; demonstrates impact
  - Depending on how it’s collected, data can be normalized and contextualized by school, subject or research area
  - Could identify low use, high impact titles and save them from cancellation

• Cons
  - Not as useful for non-research oriented institutions (i.e. liberal arts & community colleges)
  - Doesn’t capture scholarly usage outside of publishing

• Galter team currently working on project in this area for NASIG, stay tuned!
Main takeaways

No single metric is a silver bullet

• Useful to have multiple evaluation metrics to check against
  - Outliers or anomalies from one metric can be investigated further with others
  - Different metrics for different titles
• Institutional context plays a large role
  - Systems, licensing, and locations
  - Mission of school, level of research activity
Questions?
Thank You!

j-shank@northwestern.edu

@ShankLib
References


