Integrating usage statistics into a collection assessment tool via Innovative’s ERM module

NWIUG 2007
Created by Hana Levay
levay@u.washington.edu
and Diane Carroll
carroldi@wsu.edu
Intro

- Brief history of statistics
- COUNTER
- SUSHI feeds
- ERM Beta: Cost per use
- ScholarlyStats: Autostat Configuration
- What’s next?
A Brief History of Statistics

-- How is this useful??
COUNTER

- Counting Online Usage of Networked Electronic Resources
- Standardized way to count usage
- Defined reports for Journals, Databases, and eBooks
SUSHI

- Standardized Usage Statistics Harvesting Initiative
- Model for automation of statistics harvesting
- SUSHI feed: usage data that magically appears!
ERM Beta

- Can now accept SUSHI feeds
- Easy to generate COUNTER-style reports within ERM
- Brings in data from order records to calculate cost per use
### Average Cost Per Title

Average cost per title = $26121/354 titles = $73.79
## A Close-Up View of Statistics Export

<table>
<thead>
<tr>
<th></th>
<th>Title</th>
<th>2007 CPU</th>
<th>2007 Use</th>
<th>2007-Jan</th>
<th>2007-Feb</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Africa Today</td>
<td>$0.47</td>
<td>156</td>
<td>25</td>
<td>21</td>
</tr>
<tr>
<td>8</td>
<td>Africa: The Journal</td>
<td>$2.04</td>
<td>36</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>African Studies R</td>
<td>$1.05</td>
<td>70</td>
<td>10</td>
<td>16</td>
</tr>
</tbody>
</table>

CPU per title = Cost per title / Use for title

CPU: Africa Today = $73.79/156

= $0.47
**Project Muse Order Record**

**Bib Record - o1319138x**

- **OCLC #**: 34935760
- **TITLE**: Project Muse [electronic resource]
- **LOCATIONS**: el

**Summary**

- **o1319138x**
  - **Last Updated**: 09-25-2007
  - **Created**: 11-25-1996
  - **Revisions**: 70

**SELECTOR**: lahi
**INPUTTER**: Pitt, Alice

**VEN. ADDR.**:
- MINITEX LIB INFO NETWORK
- 15 ANDERSEN LIB-U OF MN
- 222 21ST AVE S
- MINNEAPOLIS MN 55455-0439

**ERM Link**: SLKe1000037

<table>
<thead>
<tr>
<th>PAID</th>
<th>DATE</th>
<th>INVD</th>
<th>INV#</th>
<th>AMT</th>
<th>VOUCHER</th>
<th>COPIES</th>
<th>FOR CURR</th>
<th>SUB FRM</th>
<th>SUB TO</th>
<th>NOTE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>11-20-2006</td>
<td>11-01-2006</td>
<td>37098</td>
<td>$15,675.00</td>
<td>192647</td>
<td>001</td>
<td></td>
<td>01-01-2007</td>
<td>12-31-2007</td>
<td>1/1/07-12/31/07</td>
</tr>
<tr>
<td></td>
<td>10-01-2005</td>
<td>10-11-2005</td>
<td>25822</td>
<td>$11,457.00</td>
<td>181800</td>
<td>001</td>
<td></td>
<td></td>
<td></td>
<td>1/1/06-12/31/06</td>
</tr>
</tbody>
</table>
**A Close-Up View: Project Muse**

<table>
<thead>
<tr>
<th>COPIES</th>
<th>FOR CURR</th>
<th>SUB FRM</th>
<th>SUB TO</th>
<th>NOTE</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td></td>
<td>01-01-2007</td>
<td>12-31-2007</td>
<td>1/1/07-12/31/07</td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>---------</td>
<td>------------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>Gene Function &amp; Diseases</td>
<td>1438-7506</td>
<td>1438-826X</td>
<td>0.00</td>
<td>5</td>
</tr>
<tr>
<td>Genes. Chromosomes And C</td>
<td>1045-2257</td>
<td>1098-2264</td>
<td>0.00</td>
<td>322</td>
</tr>
<tr>
<td>Genes. Chromosomes And C</td>
<td>1045-2257</td>
<td>1098-2264</td>
<td>0.00</td>
<td>0</td>
</tr>
<tr>
<td>Genesis</td>
<td>1626-964X</td>
<td>1626-968X</td>
<td>0.44</td>
<td>439</td>
</tr>
<tr>
<td>Genetic Epidemiology</td>
<td>0741-0395</td>
<td>1098-2272</td>
<td>0.00</td>
<td>492</td>
</tr>
<tr>
<td>Genetic Epidemiology 1984-1</td>
<td>0741-0395</td>
<td>1098-2272</td>
<td>0.00</td>
<td>0</td>
</tr>
<tr>
<td>Geochronology</td>
<td>0833-6363</td>
<td>1520-6548</td>
<td>0.00</td>
<td>73</td>
</tr>
<tr>
<td>Geochronology 1986-1995</td>
<td>0833-6363</td>
<td>1520-6548</td>
<td>0.00</td>
<td>0</td>
</tr>
<tr>
<td>Geological Journal</td>
<td>0072-1050</td>
<td>1099-1034</td>
<td>0.00</td>
<td>15</td>
</tr>
<tr>
<td>Geological Journal 1964-1995</td>
<td>0072-1050</td>
<td>1099-1034</td>
<td>0.00</td>
<td>0</td>
</tr>
<tr>
<td>Geological Journal 1951-1995</td>
<td>0072-1050</td>
<td>1099-1034</td>
<td>0.00</td>
<td>0</td>
</tr>
<tr>
<td>Goman Research</td>
<td>0172-1625</td>
<td>1522-2322</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>Goma</td>
<td>0894-1491</td>
<td>1098-1136</td>
<td>0.00</td>
<td>862</td>
</tr>
<tr>
<td>Goma 1986-1995</td>
<td>0894-1491</td>
<td>1098-1136</td>
<td>0.00</td>
<td>0</td>
</tr>
<tr>
<td>Global Business And Organiz</td>
<td>1932-2054</td>
<td>1932-2062</td>
<td>0.00</td>
<td>4</td>
</tr>
<tr>
<td>Head &amp; Neck</td>
<td>1043-3074</td>
<td>1097-0347</td>
<td>0.00</td>
<td>385</td>
</tr>
<tr>
<td>Head &amp; Neck 1978-1995</td>
<td>1043-3074</td>
<td>1097-0347</td>
<td>0.00</td>
<td>0</td>
</tr>
<tr>
<td>Health Economics</td>
<td>1057-9230</td>
<td>1099-1050</td>
<td>0.00</td>
<td>255</td>
</tr>
<tr>
<td>Health Economics 1992-1995</td>
<td>1057-9230</td>
<td>1099-1050</td>
<td>0.00</td>
<td>0</td>
</tr>
<tr>
<td>Heat Transfer - Asian Research</td>
<td>1099-2871</td>
<td>1523-1496</td>
<td>0.00</td>
<td>3</td>
</tr>
<tr>
<td>Heat Transfer - Japanese Res</td>
<td>0096-0802</td>
<td>1520-6566</td>
<td>0.00</td>
<td>6</td>
</tr>
<tr>
<td>Helvetica Chimica Acta</td>
<td>0018-019X</td>
<td>1522-2675</td>
<td>0.00</td>
<td>134</td>
</tr>
<tr>
<td>Helvetica Chimica Acta 1919</td>
<td>0018-019X</td>
<td>1522-2675</td>
<td>0.00</td>
<td>0</td>
</tr>
<tr>
<td>Hematological Oncology</td>
<td>0278-0232</td>
<td>1059-1069</td>
<td>0.00</td>
<td>52</td>
</tr>
</tbody>
</table>

Sum = 1321.44
NUM =
### Genesis Statistics Detail

<table>
<thead>
<tr>
<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>Genesis</td>
<td>$4.44</td>
<td>439</td>
<td>66</td>
<td>71</td>
<td>46</td>
<td>68</td>
<td>71</td>
<td>55</td>
<td>62</td>
</tr>
</tbody>
</table>

2007 payment / 2007 Use = 2007 CPU

$1950.71 / 493 = $4.44
ScholarlyStats

- Does most of the work for us
- Consolidates and analyzes data
- One source for SUSHI feeds
ERM AutoStat Configuration
## Example of AutoStat Setup

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AccessPrdr</td>
<td><code>musep</code></td>
</tr>
<tr>
<td>CustomerId</td>
<td><code>SS00098-lawas@prom</code></td>
</tr>
<tr>
<td>Affiliatn</td>
<td></td>
</tr>
<tr>
<td>DayOfMon</td>
<td><code>20</code></td>
</tr>
<tr>
<td>From(#MonthsBack)</td>
<td><code>2</code></td>
</tr>
<tr>
<td>To(#MonthsBack)</td>
<td><code>2</code></td>
</tr>
<tr>
<td>WSDL</td>
<td><code>http://170.224.101.201/sushi_01/s</code></td>
</tr>
</tbody>
</table>
What’s Next?

- Now that we have accurate cost per use data, what do we do with it?
- Look at a collection assessment tool
- How can we get usage data in our tools?
Serial Collection Assessment

- The evolution of electronic journals has changed the way journal collections are managed and assessed.
- Usage statistics are just one of many variables that selectors review when making collection decisions.
Objective

- Provide an example of a report that would be useful for serials collection development (Serials decision database).
- Demonstrate how close you can get creating a serials decision database using only data in Millennium, including ejournal use and cost per use information.
WSU Serials Decision Database

- Database designed for selectors to make a decision to add, cancel or renew a title
- The complete serials decision database contains subscribed and unsubscribed titles
- This example contains only subscribed titles so the same set of titles can be compared to results obtained only from Millennium
- Handout 1: WSU Serials Decision Database
  - Row 1 – Source of the information
  - Row 2 – Label for data found in the columns
What sources were included for assessment of subscribed titles?

<table>
<thead>
<tr>
<th>Millennium</th>
<th>Subscription agent</th>
<th>Scholarly Stats and publisher website</th>
<th>Calculated</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Title, ISSN,</td>
<td>- Title, format purchased, subscription period, publisher, ISSN</td>
<td>- Ejournal use statistics</td>
<td>- Cost per use</td>
</tr>
<tr>
<td>- Order record #, location, vendor, fund code, cost</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Access provider Ejournal coverage,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Archival access and license restrictions</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
What sources included subscribed and unsubscribed titles?

- **ILLiad**
  - Interlibrary loans borrowed (ILLiad), ISSN
- **Web of Science**
  - Number of authored papers
  - Number of papers referenced by institution’s authors
- Calculated information on the priority of the titles
<table>
<thead>
<tr>
<th>Priority of the journal</th>
<th>High Use (1)</th>
<th>Medium Use (2)</th>
<th>Low Use (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of articles by WSU authors in the journal from 2004 to 2006</td>
<td>&gt;10</td>
<td>5-10</td>
<td>2-4</td>
</tr>
<tr>
<td>Number of articles cited by WSU authors from 2004 to 2006</td>
<td>&gt;40</td>
<td>10-40</td>
<td>5-9</td>
</tr>
<tr>
<td>Full text/PDF downloads</td>
<td>&gt;250</td>
<td>40-250</td>
<td>&lt;40</td>
</tr>
<tr>
<td>Interlibrary loan requests over the last 12 months</td>
<td>&gt;10</td>
<td>6-9</td>
<td>5</td>
</tr>
</tbody>
</table>
Creating a Serials Decision Database only from Millennium

- Handout 2 – Serials Decision Database created from Millennium
Create List challenges

- Can not do one create list that will pull data from the Bib, Order, Checkin, Resource and License records.
- Must do this report in stages and merge data to get information on each subscribed title on one line on a spreadsheet.
Tip about getting holdings for ejournals

The coverage database (ERM) contains the holdings information for ejournals.

- To include electronic holdings (or years covered) be sure to go to Admin – Settings and check “include coverage database information” before running a create list.
Step 1: Create a list of current subscriptions using the Checkin Record

Do you have a current subscription?

- 008 Marc Fixed-Length Data Elements
  - 06 Receipt or acquisition status (R/A Stat)
    - 0 Unknown
    - 1 Other receipt or acquisition status
    - 2 Received and complete or ceased
    - 3 On order
    - 4 Currently received
    - 5 Not currently received
Step 1: Create a list of current subscriptions using the Checkin Record
Export the data for Bib and Checkin
Step 2: Create a list of order records

Not all current subscriptions have order records
- Electronic – Aggregator collection of 10,000 titles (EBSCOhost, ProQuest, Wilson, etc).
- Consortial shared titles
- Print gifts

Not all order records have payment information
- Bundled titles in print
Step 2: Create a list of order records

- To accurately calculate cost per use, each title must have an order record with payment information.
- There is no direct link between order record and checkin records for print titles.
- There is a soft link between related holdings and related orders via the Resource Record. These must be linked manually.
Step 2: Create a list of order records

- The order record must indicate if the subscription is current.
  - Status = F or Serials, not encumbered
- The goal is to have one line of data for each order
- If a field is repeated, it must line up consistently so it can be edited on the spreadsheet (Paid)
- The results of the create list need to be limited to a specific time period. This can only be done in character based system.
Step 2: Create a list of order records
Step 2: Create a list of order records
Step 2: Create a list of order records

-----OUTPUT RECORDS-----
Your review file, Cost Per Use 3.06 cellis,
has 6378 records in it ready to be output
  PAID FIELD SELECTED
  Limit payments counted to just those made within a particular
  time period? (y/n) y
  Use PAID date or INVOICE date? (p/i)p
  Do a comparative report for two time periods? (y/n) n

  Include payments made beginning on: 07-06-2006
  and ending on: 07-11-2007
  Only payments made between 07-06-06 and 07-11-07
  will be included in this report
  Are these dates correct? (y/n) y
Step 3: Merge the order record data with the checkin record data

- Output the file and manipulate in EXCEL
- Total all of the Amounts Paid
- Remove duplicate bib records from both data sets (use the subtotal function on bib record)
- Use Vlink or a macro to merge data by Bib record
- Hand merge the duplicates
Step 4: Create list for Resource and License Record data

- **Resource Record**
  - Resource Name
  - Access provider
  - Resource ID

- **License Record**
  - Terms of Use
  - Archival information

- Match data to checkin record by resource ID
Step 4: Create list for Resource and License Record data
Step 4: Create list for Resource and License Record data
Step 5: Add use stats and cost per use

- Use data and cost per use are not part of the Create List function.
- Need to download spreadsheets loaded into ERM and merge the data using ISSN.
Making Millennium more friendly for journal collection assessment

Greater integration of data
  - Need to be able to download journal, provider, licensing restrictions, coverage and use data in a create list

Expanded use of import function introduced with ERM
  - Can already load coverage information
  - Need the ability to upload data such as Web of Science use data, priorities, other local use information and link it to the checkin records
Making Millennium more friendly for journal collection assessment

Creative uses of the ERM coverage loader
- unsubscribed titles and data uploaded and suppressed from the OPAC so both subscribed and non-subscribed titles can be reviewed.