IOTA

NISO Update
January 22, 2012
Oliver Pesch
Chief Strategist, E-Resources
EBSCO Information Services
Overview

- Quick review
- Recent activities
- Current activities
Quick Review

• IOTA stands for:
  Improving OpenURL Through Analytics

• Offshoot of work done by Adam Chandler trying to answer the question:
  Why do some OpenURLs not resolve?
Quick Review

The theory

• If you know what data elements are required to create target links
• You should be able to predict the probability of success for any given OpenURL based on the elements included in the link
Quick Review

The Completeness Score

• A mathematical formula that quantifies the probability of success by scoring the OpenURL based on the data elements included, where the value of each element is weighted based on how often it is needed in a target link

• Group all the OpenURLs for a given Source and look at their scores in aggregate – you can rate the provider of the OpenURL
Quick Review

• Committee Members
  o Adam Chandler, Cornell University (Chair)
  o Aron Wolf, Serials Solutions
  o Clara Ruttenberg, University of Maryland
  o Elizabeth Winter, Georgia Tech
  o James Wismer, Thomson Reuters
  o Oliver Pesch, EBSCO Information Service
  o Rafal Kasprowski, Rice University
  o Susan Marcin, Columbia University

• Charge
  o Investigate the feasibility of creating industry-wide, transparent and scalable metrics for evaluating and comparing the quality of OpenURLs
Accomplishments

- OpenURL Quality Web Site
  - Processed over 21.5 million OpenURLs
  - Calculated Completeness Scores
  - Provides Completeness Scores by SOURCE (OpenURL Providers)
  - Provides metrics on elements frequencies

http://openurlquality.niso.org
Current Activities

- Conducting validation experiments
  - Confirm a correlation between the Completeness Score and the OpenURLs ability to create an actual full text or ILL link.
  - EBSCO and Serials Solutions participating
  - Using a sample size of 10s of thousands of OpenURLs
Next Steps

• Complete validation experiments
• Adjust element weights, if necessary
• Complete a final report for the committees work with recommendations for further work
• Continue to maintain the OpenURL Quality web site as a tool for librarians and content providers
Overview

• Quick review
• Current activities
  o Support for Release 4 of the COUNTER Code of Practice
  o SUSHI Server Test Mode Recommended Practice
  o SUSHI Server Status Report
  o COUNTER SUSHI Implementation Profile
Quick Review

- Standardized Usage Statistics Harvesting Initiative
- Addresses the labor intensive and repetitive task of retrieving COUNTER reports
Quick Review

• Committee Members
  o Bob McQuillan, Innovative Interfaces, Co Chair
  o Chan Li, California Digital Library
  o John Milligan, ScholarlyIQ
  o Paul Needham, Cranfield University
  o Oliver Pesch, EBSCO Information Services, Co-chair

• Charge
  o Standing committee responsible for the SUSHI standard
  o Maintenance Agency for the COUNTER XML Schema
COUNTER Code of Practice: Release 4

Addressing…

• New formats
• New devices used for access
• New access technologies
• Need for automation
• Need for machine processing
Changes from prior release...
- New reports
- New metric types
- New data elements

What this means for the SUSHI committee
- Update the COUNTER XML Schema
- Update the COUNTER Enumeration Schema
- Update the reference documents on the web site
SUSHI Server Test Mode: Recommended Practice

The problem
• Developers find it difficult to test their SUSHI client against production versions of SUSHI Servers
  o Complicated requirements for authentication
  o Limits to the number of requests per day
  o Test data may not represent reality

The solution
• A recommended practice
  o Lowers the barrier for the developer
  o Simplified registration
  o Removed authentication hurdles
  o Allows multiple re-tries in a day
  o Data representative of live data
The goal

- A simple report that can be requested via SUSHI
- Provides status information
  - Is the server able to deliver usage
  - What dates of usage are available
- Provides configuration information
  - Lists COUNTER reports that are supported
  - Instructions for configuring a client
  - Authentication information
  - Day of the month when usage available
Why

• SUSHI and COUNTER tend to be abstract
• Abstraction leaves the opportunity for concepts to be interpreted differently
• Different interpretations leads to inconsistencies
• Examples:
  o Different authentication methods that require custom development
  o Wrong capitalization on controlled data values
  o Including “Totals” in the XML
  o 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
COUNTER-SUSHI Implementation Profile

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
COUNTER-SUSHI Implementation Profile

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
COUNTER-SUSHI Implementation Profile

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
## 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_html, 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_html, 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_html, 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>turnaway, no_license, other</td>
</tr>
</tbody>
</table>
Next Steps

• Finalize and publish…
  o COUNTER SUSHI Implementation Profile
  o SUSHI Server Status Report
  o SUSHI Server Test Mode Recommended Practice
• Continue our work in supporting developers and librarians
• Continued promotion and education