There are two components to this webinar:
1. The online presentation:
   Congratulations, you have logged in successfully!
2. The phone conference call: To join, phone into the conference at 888-858-6021 and enter PIN code 2583615155.

Communiqué Conferencing technical assistance is available by:
• Press 00 on your telephone or
• Call 1-877-283-7062 or +(973) 796-5047

Welcome to the NISO Performance Measures Webinar on November 14, 2008.

How it works:
• Let the presentation run; you do not have to interact with your computer.
• Your phone conference line is on listen-only mode. If you have a question during the meeting send a message to the presenter using the “Q&A” button at the top of the screen (more on this is on the next slide).

Communiqué Conferencing technical assistance is available by:
• Press 00 on your telephone or
• Call 1-877-283-7062 or +(973) 796-5047

Asking Questions During the Webinar

Click “Q&A” menu located at the top left of your screen to open the Question Panel.
Place your cursor into the blank field at the top, type in your question and press the ‘Ask’ button.
We will answer as many questions that we can during the presentation. Questions that we aren’t able to answer during the webinar will be answered and made available on the NISO website following the event.

Frequently Asked Questions

Will I be able to get a copy of these slides after the event?
Yes
Is this Webinar being recorded so that I can view it at a later time?
Yes

You will receive an e-mail with information on how to access these following the webinar.

NISO Education: Join Our Next Webinars

NISO/EDItEUR Webinar
ONIX for Serials: Case Studies of Use
Thursday, December 11th, 1:20 p.m.

2009 Q1 Webinars
- January 14: Preservation
- February 11: Authentication
- March 11: Data Movement & Management

Visit online to see our full 2009 education calendar:
ANSI/NISO Z39.7
Information Services and Use:
Metrics & Statistics for Libraries and
Information Providers

Dianne L. Carty
Chair, Z39.7 Standing Committee
Head, State Aid & Data Coordination,
Massachusetts Board of Library Commissioners

Z39.7—Brief Background

The Library Statistics Standard, ANSI/NISO Z39.7-1995, was first

When the 1995 edition of the Library Statistics Standard was released
the committee that developed it acknowledged that the Standard did
not address two important emerging areas: measurement of
electronic resources and performance measures.

It was recommended that these issues be examined at the next five-
year review point.

Z39.7—Brief Background

The NISO Forum on Performance Measures and Statistics for Libraries
held in February 2001 resulted in several recommendations to NISO.

National standards should move beyond defining data elements to
provide methodologies for qualitative and quantitative measures of
library service.

NISO should take a leadership role in supporting the development of
surveys/measurements for to gauge service quality and outcomes.

Z39.7—Brief Background

In 2002, the a NISO Standards Committee AY was appointed and tasked to:

- Revise ANSI/NISO Z39.7-1995 incorporating improvements based on the existing survey
  instruments in use by nationally recognized U.S. library data collection programs.
- Prepare for release a Draft Standard for trial use of evolving methods of measuring electronic
  network performance, vendor and publisher-based use statistics reporting methods.
- Review and incorporate as appropriate definitions for evolving methods of measuring electronic
  network performance, vendor and publisher-based use statistics reporting methods, and service
  quality measures.
- Based upon the results of the trial use of evolving e-metrics, advise NISO on how best to integrate
  new measures into Z39.7 and/or suggest other appropriate approaches (such as nationally
  supported best practices or guidelines) that over time might lead to the development of national
  consensus standards.

Z39.7—Brief Background

The committee developed an interactive data dictionary
rather than print standard.

The draft standard for trial use was released by June 2002.

There was a comment and trial use period of 12 months.

The web-based comments site was up and running in

Key elements of the 2002 revision
and in the Current Data Dictionary

Reporting Units
Describes types of libraries, the entity that has administrative or budgetary control, and the
populations served by each.

Human Resources
Describes all levels of staffing associated with each library type.

Collections
Describes broad collection categories in all formats.

Infrastructure
Describes facilities, including capacity, technology, and hours open.

Finances
Describes broad categories of revenue (e.g., income) and expenditure by type and source.

Services
Describes broad categories of services provided by libraries, as well as other metrics associated with
understanding library use both tangible and virtual.
Why a standard? Why Z39.7?

Key term is comparability—

- [ Across time
- [ Across libraries—both statewide and nationally

What does comparability do?

Allows libraries to:

- [ Advocate for increased funding
- [ Measure performance
- [ Assess services and collections

How data is used—Massachusetts

“I use it [data] annually to get raises for my staff! (I’m not always successful, but no matter what happens, the numbers help me to show our Town Fathers how successful the library is as compared to similar libraries.)”

How data is used—Massachusetts

“Because I have used this data so consistently during my 20 years as Director here, boards such as the Personnel Board and Advisory Committee know they can rely on me to answer their questions, provide transparency, and that the Library is highly accountable.”

How data is used—Massachusetts

“I often use the public library data to defend my budget to town officials.

It is also useful for making comparisons with other libraries. I like to use graphs and other visuals to show where the library stands in comparison with others, and to show how much the public is using us. When I first arrived in Millis I created a “State of the Library” PowerPoint which I now give each year to Fin Comm, Selectmen, and anyone else who will listen. The public library data is an important source of info for these presentations.”
At a Glance

Annuol's Public Library Services for FY2006

<table>
<thead>
<tr>
<th>Library</th>
<th>Memorial Library</th>
<th>Budget</th>
<th>Hours</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Statistics:

- Total Population: 6,288,736
- Total Libraries: 5,368

HOW DATA IS USED—MASSACHUSETTS

http://mblc.state.ma.us/

Friday, November 14, 2008

Next Steps

- Continuous maintenance of Z39.7 through standing committee review of comments posted to the data dictionary
- Promote local use of data dictionary through outreach to library data collectors and data users

Friday, November 14, 2008

Final Points

What is critical?

- Consistency
- Comparability
- Accuracy
- An understanding of the definition
- Vetted data
- An understanding of the context

Friday, November 14, 2008
Next Steps

Assess the need for performance indicators to use in tandem with the data dictionary.

Thank you.

Scholarly assessment in the digital era.

Johan Bollen
Digital Library Research & Prototyping Team
Los Alamos National Laboratory - Research Library

jbollen@lanl.gov

Acknowledgements:
Herbert Van de Sompel, Marko A. Rodriguez (LANL), Ryan Chute (LANL), Lyudmila L. Balakireva (LANL), Arc Hagberg (LANL), Luis Bettencourt (LANL)

Research supported by the Andrew W. Mellon Foundation.

Scholarly assessment

Determining the value/impact/influence of scientific activity

- Qualitative/subjective:
  - Peer review
  - Tenure committees
  - Networks

- Quantitative:
  - Mostly citation statistics
  - Many proposals, little clarity

Citations and the journal impact factor.

Citation data:
- Golden standard of scholarly evaluation
- Citation = network of scholarly influences.
- Extracted from published materials.
- Main bibliometric data source for scholarly evaluation.

Journal Impact Factor:
- Mean 2-year citation rate
- Published yearly by Thomson Scientific
- +/- 8000 journals tracked

Widely applied:
- Fair approximation of journal "status"...
- Authors -> users?
- Trust, Prestige, Popularity
- See MySpace and Facebook!

Scholarly assessment

Electronic paradigm changes everything
- New models of communication
  - Everything will be published
  - No central vetting authority
- New models of scholarship
  - Publish multimedia, raw data, software
- New data sources and new metrics of evaluation?
  - Authors, # users?
  - Trust, Prestige, Popularity
  - See MySpace and Facebook!

Articles, journals, Citation data

Citation metrics

IR, Pre-print, multimedia, raw data, software, etc
Citation data limitations

Citation data's limitations:
• Community: authors of journal articles.
• Artifacts: journal articles (8,000 journals?).
• Timing: +1 year publication delay.

Usage data's promise:
• Community: any user
• Artifacts: all that is accessible.
• Timing: recorded upon publication.

Hence, various initiatives focused on usage data: COUNTER, IRS, SUSHI, Citibase

Usage data and derived metrics: move to 21st century assessment paradigm suitable for e-science, open science:
1) usage data + 2) network metrics.

The MESUR project.

○ Johan Bollen (LANL): Principal investigator.
○ Herbert Van de Sompel (LANL): Architectural consultant.
○ Ari Hagberg (LANL): Mathematical and statistical consultant.
○ Lyudmila Balakireva (LANL): Database management and development.
○ Wenxiong Zhao (LANL): Data processing, normalization and ingestion.

“The Andrew W. Mellon Foundation has awarded a grant to Los Alamos National Laboratory (LANL) in support of a two-year project that will investigate metrics derived from the network-based usage of scholarly information. The Digital Library Research & Prototyping Team of the LANL Research Library will carry out the project.

The project’s major objective is enriching the toolkit used for the assessment of the impact of scholarly communication items, and hence of scholars, with metrics that derive from usage data.”

MESUR general approach

Generalizable, quantitative results
1. Create very large-scale reference data set (1B usage events)
   1. Usage, citation and bibliographic data combined
   2. Various communities, various collections
2. Investigate sampling issues:
   1. Effects of sampling on usage-based assessment
   2. Mapping and characterization of scholarly community
   3. Uncertainty quantification: noise, bots, …
3. Investigate validity of usage data and usage-based metrics
   1. Cross-validation: compare to existing, accepted journal-focused metrics and data
   2. Not selling 1 metric: exploring many possibilities, many facets of impact
   3. Exploratory approach: not top-down, bottom-up exploration
4. Lay foundation for scientific, generalizable study of usage data-based assessment

Presentation structure
1) Usage data acquisition
2) Science mapping from usage graphs
3) Metrics survey
4) Services
5) Discussion

Network-Based Metrics

We have 50 years of network science available to us

• A wide variety of metrics has been proposed to characterize networks, and to assess the importance of nodes in a network
  • E.g., social network analysis, small world graphs, graph theory, social modeling
  • So when defining metrics for scholarly communication (clearly a network), we should probably leverage network science
    • Cf. Google’s PageRank versus Alta Vista’s statistical ranking
  • A network (and hence a network-based metric) takes context into account; a statistical count does not.

• Readings:

Presentation structure
1) Usage data acquisition
2) Science mapping from usage graphs
3) Metrics survey
4) Services
5) Discussion
How to obtain 1,000,000,000 usage events?

Politely asked selected publishers, aggregators and institutional consortia for usage data:

- Scale: $> 1,000,000,000$ usage events and $>500,000,000$ citations
- Span:
  - $> 50M$ documents
  - $> 100,000$ journals (incl. newspapers, magazines, ...)
- Use reference data set for mapping, metrics and services

Presentation structure

1. Research Library, Los Alamos National Laboratory
2. Metrics survey
3. Services
4. Discussion
5. Network metrics from usage data: usage networks?

Usage map

- 200M usage events
- 2006 usage only
- JCR journals (~7600)

Citation map

- 1,000,000,000
- 2006
- 2002-2007

Data normalization and ingestion

Minimal requirements for all usage data

- Unique usage events (article level)
- Fields: unique session ID, date/time, unique document ID and/or metadata, request type
- Note difference with usage statistics
- Central role for link resolver data

Network metrics from usage data: usage networks?

Same session = documents relatedness
- Same session, same user: common interest
- Frequency of co-occurrence = degree of relatedness
- Normalized: conditional probability

Usage data is on article level:
- Works for journals and articles
- Anything for which usage was recorded

Note: not something we invented!
- Association rule learning in data mining
- Beer and diapers!
Presentation structure

1) Usage data acquisition
2) Usage data aggregation from usage graphs
3) Metrics survey
4) Services
5) Discussion

Structural metrics calculated from usage graph

- Classes of metrics:
  - Degree
  - Shortest path
  - Random walk
  - Distribution

- Degree
  - In-degree
  - Out-degree

- Shortest path
  - Closeness
  - Newman

- Random walk
  - PageRank
  - Eigenvector

- Distribution
  - In-degree entropy
  - Out-degree entropy
  - Bucket Entropy

Each can be defined to take into account weights by e.g.
means of weighted shortest path definition

Citation network rankings

- List of metrics:
  - Citation PageRank (CITE-PG)
  - Citation In-Degree (CITE-IN)
  - Citation Out-Degree (CITE-OD)
  - Citation Betweenness (CITE-BE)
  - Citation Entropy (CITE-ENT)
  - Citation pagerank (CITE-PG)
  - Citation In-Degree (CITE-IN)
  - Citation Out-Degree (CITE-OD)
  - Citation Betweenness (CITE-BE)
  - Citation Entropy (CITE-ENT)

- Usage network rankings

- List of metrics:
  - Usage PageRank (USES-PG)
  - Usage In-Degree (USES-IN)
  - Usage Out-Degree (USES-OD)
  - Usage Betweenness (USES-BE)
  - Usage Entropy (USES-ENT)
  - Usage pagerank (USES-PG)
  - Usage In-Degree (USES-IN)
  - Usage Out-Degree (USES-OD)
  - Usage Betweenness (USES-BE)
  - Usage Entropy (USES-ENT)

Note:
- Metrics can be calculated both on citation and usage data
- Structural metrics require networks:
  - E.g., JCR
  - Usage: created from MESUR data

Usage graph created: Shwering Filen
Metrics: Marko Rodriguez and Aric Hagberg

Usage network rankings

- List of metrics:
  - Usage PageRank (USES-PG)
  - Usage In-Degree (USES-IN)
  - Usage Out-Degree (USES-OD)
  - Usage Betweenness (USES-BE)
  - Usage Entropy (USES-ENT)
  - Usage pagerank (USES-PG)
  - Usage In-Degree (USES-IN)
  - Usage Out-Degree (USES-OD)
  - Usage Betweenness (USES-BE)
  - Usage Entropy (USES-ENT)

- Usage network rankings

- List of metrics:
  - Usage PageRank (USES-PG)
  - Usage In-Degree (USES-IN)
  - Usage Out-Degree (USES-OD)
  - Usage Betweenness (USES-BE)
  - Usage Entropy (USES-ENT)
  - Usage pagerank (USES-PG)
  - Usage In-Degree (USES-IN)
  - Usage Out-Degree (USES-OD)
  - Usage Betweenness (USES-BE)
  - Usage Entropy (USES-ENT)

- Usage network rankings

- List of metrics:
  - Usage PageRank (USES-PG)
  - Usage In-Degree (USES-IN)
  - Usage Out-Degree (USES-OD)
  - Usage Betweenness (USES-BE)
  - Usage Entropy (USES-ENT)
  - Usage pagerank (USES-PG)
  - Usage In-Degree (USES-IN)
  - Usage Out-Degree (USES-OD)
  - Usage Betweenness (USES-BE)
  - Usage Entropy (USES-ENT)
**Presentation structure**

1) Usage data acquisition  
2) Science mapping from usage graphs  
3) Metrics survey  
4) Services  
5) Discussion

**MESUR: conclusions.**

**After 1.5 year of MESUR:**

- Scientific exploration of metrics for scholarly evaluation  
- Creation of large-scale reference data set  
- Mapping science from the viewpoint of users  
- Metrics that cover various aspects of scholarly impact and prestige  
- Infrastructure for a continued research program

**Final deliverables:**

- A better understanding of the possibilities and pitfalls  
- Guidelines on metrics:  
  - Standardized definitions  
  - Semantics: aspect of impact and interpretation  
  - Community: proper usage
BUT, serious challenges

Scalability of this approach:
• 1.5 years of negotiation to obtain log data
• No standards: Recording, aggregating, normalizing, ingestion, deduplication,…
• No generally accepted principles: privacy, property,…
• No census data: when is a sample large and representative enough?

Quality control:
• Bots, Crawlers (detectable but never perfect)
• Cheating, manipulation (easier with usage statistics than network metrics)

Community acceptance and advocacy:
• Simple is not always better (but smart is)
• Fundamental issue of fairness in evaluation: one number to rule them all? Precisely what we hope to avoid.

Some relevant publications.


Questions?

All questions will be posted with presenter answers on the NISO website following the webinar:

www.niso.org/news/events/2008/webinars/measures

Thank You!

Thank you for joining us today. Please take a moment to fill out the brief online survey.

We look forward to hearing from you!