WE’VE GOT THE WORLD COVERED WITH RESEARCH EVALUATION METRICS, TOOLS AND EXPERTISE FROM THOMSON REUTERS

Why do over 100 countries and all major funding bodies worldwide turn to Thomson Reuters when they need to make decisions that impact resources, careers and future research directions?

Because they know that Thomson Reuters delivers objective data and reliable tools based on established principles of bibliometrics and citation analysis. Backed by a long history of leadership and innovation in research evaluation. And enhanced with powerful functionality that lets you easily manipulate, summarize and graphically display your findings as you:

- **Track** researcher output and impact
- **Set** measurable goals
- **Allocate** funds accurately and intelligently
- **Attract** top-notch faculty and funding sources
- **Pinpoint** emerging trends and promising avenues of research
- **Identify** collaboration and tech transfer opportunities
NEWS

3 NISO Elects New Chair, Vice Chair, and Directors to Board
3 Open Archives Initiative Announces Public Beta Release of Object Reuse and Exchange Specifications
3 DAISY Issues 2008 Revision of Structure Guidelines for Digital Talking Books
4 NISO/ALPSP Committee Releases Best Practices for Journal Article Versions
4 Microsoft ‘Save as DAISY’ Plug-In Integrates NISO Digital Talking Book Standard
5 Comments Invited on Interactive Data Dictionary for Information Services Metrics
5 Proposed NISO CORE Standard to Form Information Bridge Between ILS and ERMS
5 New Versions of ONIX for Serials SRN and SPS Message Formats Piloted
6 AIIIM Offers PDF/Archive Training Courses
6 Techstreet and Intertek Launch New Standards Store
7 Web Portal Created for Case Studies on University Digital Records
7 ISBN Standard Undergoes Systematic Review
7 PDF format becomes ISO standard

CONFERENCE REPORT

Todd Carpenter reports from Stockholm on ISO International Committee on Information and Documentation

EDITORIAL

Mining for meaning

FEATURES

12 P—the Institutional Identifier
Helen Henderson asserts the time has come to take control of institutional identification

16 LIBX: Using Standards to Embed the Library into the User’s Webflow
Annette Bailey explains how LibX embeds standards into the library workflow

18 Progress on the DLF ILS Discovery Interface API
Marshall Breeding brings back news from Berkeley on the new DLF ILS API

20 State of the Standards

Ad Rates

<table>
<thead>
<tr>
<th>WINTER/SPRING 2008</th>
<th>SUMMER/FALL 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>FULL PAGE</td>
<td>$350</td>
</tr>
<tr>
<td>HALF PAGE</td>
<td>$225</td>
</tr>
<tr>
<td>QUARTER PAGE</td>
<td>$175</td>
</tr>
<tr>
<td>1/8 PAGE</td>
<td>$125</td>
</tr>
<tr>
<td>1/16 PAGE</td>
<td>$100</td>
</tr>
<tr>
<td>BACK COVER 4C</td>
<td>$800</td>
</tr>
<tr>
<td>COVER 4C**</td>
<td>$450</td>
</tr>
<tr>
<td><strong>(INSIDE FRONT OR BACK)</strong></td>
<td></td>
</tr>
</tbody>
</table>

CONTACT: MARYANN KARINCH, 970-577-8500, MARYANN@KARINCH.COM

Subscribe to ISQ

1 Year Subscription Rates
US: $126 / International: $160

Order Information:
Place your order with NISO, One North Charles Street, Suite 1905, Baltimore, MD 21201

For questions, call 301-654-2512 or e-mail nisohq@niso.org

Reuse

FOR PERMISSION TO PHOTOCOPY OR USE MATERIAL ELECTRONICALLY FROM INFORMATION STANDARDS QUARTERLY, ISSN 1041-0031, PLEASE ACCESS WWW.COPIRIGHT.COM OR CONTACT COPYRIGHT CLEARANCE CENTER, INC. (CCC)
222 ROSEWOOD DRIVE, DANVERS, MA 01923, 978-750-8400. CCC IS A NOT-FOR-PROFIT ORGANIZATION THAT PROVIDES LICENSES AND REGISTRATION FOR A VARIETY OF USERS.
Is there a standards based way for representing a location?

We’ve all had the experience of looking for a book in a library or book store and not finding it on the shelf. In part, the varying practices of putting the final resting location in various electronic places, particularly in the MARC record, have led to confusion, anything but interoperability, and difficulty migrating from large business systems. This is a vexing problem that everything from the NISO RFID best practices document to the OpenURL standard to better data cleanup promises to solve.

I found myself very interested in Marshall Breeding’s description of the DLF ILS API meeting in Berkeley, CA. The DLF ILS API seeks to provide four basic function. The GetAvailability function “provides location as text strings as an optional response format,” as well as status, which should make it easy to build new interfaces which utilize this information, from AJAX-enabled sites to mobile interfaces.

Though this may seem obscure, it promises to open up the discovery process exponentially, as Jon Udell’s Library Lookup project has proved with its easy linking from Amazon into library catalogs. Annette Bailey describes her project LibX, which offers a standards-based method of embedding research services directly into the browser in an amplified extension of Udell’s implementation.

In other disambiguation news, Helen Henderson writes of a new NISO project to bring together Institutional Identifiers, much like Dun and Bradstreet have been doing in the corporate world.

Finally, Todd Carpenter, Managing Director of NISO and publisher of ISQ, provides an excellent conference report on the Stockholm ISO conference, as well as the ALA preconference.

I look forward to your comments and questions—and if you can find that copy of Revolutionary Road I was looking for on the shelf yesterday, can you text the location to me? Thanks.

Jay Datema, Editor-in-Chief
editor@niso.org
NISO Elects New Chair, Vice Chair, and Directors to Board

NISO voting members have elected new leadership for the 2008-2009 terms beginning on July 1, 2008. Oliver Pesch has been elected to serve as Chair of the NISO Board of Directors. Currently the Chief Strategist for EBSCO Information Services in Birmingham, Alabama, Pesch helps set direction for EBSCO’s E-Resource Access and Management Services initiatives. Chuck Koscher has been elected to serve as Vice-Chair, and will become chair in 2009, when Pesch’s term expires. As Director of Technology at CrossRef, Koscher has been actively involved in improving the linking infrastructure for scholarly publications. Both will serve on NISO’s Board and Executive Committees.

“I am very excited about the opportunity to serve the community in leading NISO,” said Pesch. “The Organization has made enormous strides over the past three years, transforming itself internally and externally and is poised to effect great change in the community.”

The following industry leaders were also elected to seats on the NISO Board of Directors. They will join other currently serving Directors in managing and setting strategic direction for the organization.

• Nancy Davenport, President, Nancy Davenport & Associates
  • Janice Fleming, Director of Business and Planning, PsycINFO, American Psychological Association
  • Bruce Heterick, Director, Library Relations, JSTOR
  • Barbara Preece, Executive Director, Boston Library Consortium
  • Bruce Rosenblum, CEO, INERA, Inc.
  • Mike Teets, Vice President, Global Engineering, OCLC

Members of the NISO Board who will continue their terms through 2008-09 are:
  • James Neal, Vice President for IS & University Librarian, Columbia University
  • Winston Tabb, Dean of University Libraries, Johns Hopkins University
  • John Erickson, Principal Scientist, Hewlett-Packard Laboratories
  • John Harwood, Professor, Pennsylvania State University
  • Michael Jensen, Director of Publishing Technologies, National Academies Press

“We are very pleased to have such a diverse and esteemed group of Directors joining the Board this year,” said Todd Carpenter, Managing Director of NISO. “These industry leaders represent the diversity of our membership from libraries, publishers and systems providers and we trust that they will deftly guide NISO in the coming year.”

Open Archives Initiative Announces Public Beta Release of Object Reuse and Exchange Specifications

The beta version of Object Reuse and Exchange (ORE) specifications and implementation documents were released to the public in June by the Open Archives Initiative.


The DAISY Consortium has issued a 2008 revision of the Structure Guidelines that provide information on the correct usage and application of DAISY XML (the DTBook XML element set) in the creation of DAISY publications. The DAISY/NISO Standard (ANSI/NISO Z39.86), Specifications for the Digital Talking Book, is known as DAISY 3.

The DAISY 3 Structure Guidelines adds two new main sections (Poetry and Mathematics), and an Index of Elements, Alan alphabetical listing of the DTBook elements with each element linked to the point in the Guidelines where the element is explained. Also included are page example images with corresponding markup examples and information about NIMAS (National Instructional Materials Accessibility Standard) markup. NIMAS is the required electronic format in the U.S. for publishers who produce K-12 textbooks. The NIMAS Standard is a minimum required set of elements adopted from the DAISY Standard and is conformant to the DTBook DTD, upon which the DAISY Standard is based.


DAISY 2008 Structure Guidelines
www.daisy.org/z3986/structure/
Digital Talking Book standard
www.niso.org/workrooms/daisy/
NISO/ALPSP Committee Releases Best Practices for Journal Article Versions

NISO, in partnership with the Association of Learned and Professional Society Publishers (ALPSP), has published the Recommended Practice Journal Article Versions (JAV): Recommendations of the NISO/ALPSP JAV Technical Working Group (NISO-RP-8-2008). The publication is designed to provide a simple, practical way of describing the versions of scholarly journal articles that typically appear online before, during, and after formal journal publication.

"Static, single copies of research papers that are essentially facsimiles of a single, unambiguously identified printed document are a thing of the past," stated Bernard Rous, Deputy Director of Publications at the Association for Computing Machinery (ACM) and Co-Chair of the JAV Working Group. "Changes in the way we create, produce, and store articles lead to multiple versions that are often all discovered together through web searches. Our working group addressed the consequent problem: how to identify the versions retrieved and clarify the relationships among them."

In September 2005, NISO launched the partnership with ALPSP to bring together experts from the publishing, library, library systems, and user communities to examine the problems associated with the proliferation of different article versions. Led initially by Cliff Morgan, Vice President, Planning & Development Director at John Wiley & Sons Ltd., the group focused its attention on describing the important stages in the production of scientific articles.

The JAV Working Group also created use cases to explore the lifecycle of journal articles, starting from a base case that describes a typical interaction among author, institutional repository, and publisher. Rather than addressing all possible iterations of an article from origination to publication, the group focused on key stages in recording a document’s development.

“This project has made a significant advancement in the identification and description of common lifecycle stages of articles,” noted Ian Russell, Chief Executive of ALPSP. “The evolution of articles—and possible attributes of each instance of an article version—is a critical component of understanding, managing, and sharing information in the rapidly changing publication landscape.”

Several variables were considered as possible dimensions to identify a particular article version:

- Time: from first draft to latest version
- Added Value: from rough draft to polished publication
- Manifestation/Rendition: different document formats and layouts
- Siblings: multiple mappings between technical reports, conference papers, lectures, journal articles, review articles, etc.
- Stakeholders: author, editor, referee, publisher, librarian, reader, funding organization

Components of the JAV Recommended Practice include a narrative that explains the project background and rationale for recommended terms and definitions, and appendices that cover “Graphical Representation of Journal Article Versions and Relationships with Formal and Grey Literature; Assumptions, Primary Challenges, and Best Practices,” use cases, and comments from JAV Review Group on recommendations received on an earlier draft document.

“The breadth of participation in this process helps to ensure that the group has captured the essence of production lifecycles across a broad range of publishers,” said Todd Carpenter, Managing Director of NISO. “In addition, having the partnership with ALPSP behind the report should encourage the wide adoption of this terminology and descriptive information in our community.”

NISO plans to aggressively promote use of the JAV recommendations in the information dissemination community over the coming months.

Microsoft ‘Save as DAISY’ Plug-In Integrates NISO Digital Talking Book Standard

Microsoft, the DAISY Consortium, and Sonata Software have released the production version of the “Save to DAISY,” translator for Microsoft Word. This Word plug-in converts documents from Word’s Open XML format to the DAISY format that is defined in the NISO standard, Specifications for the Digital Talking Book (ANSI/NISO Z39.86). The DAISY Consortium is the Maintenance Agency for the NISO standard.

This free and open-source software provides feature-rich, structured information to persons who are blind or print disabled, and can assist in the creation of e-books. DAISY Digital Talking Books (DTBs) go far beyond the limits imposed by analog audio books or commercial digital audio books. In a DAISY book, the audio is synchronized with the textual content and images, providing an accessible and enriched multimedia reading and learning experience. A DAISY book also supports multiple outputs, such as Braille and large print.

“Microsoft’s initiative to put Save as DAISY XML in Microsoft Word is the first step to bring fully accessible content to people who are blind or who have a print disability. We know that much of the information in documents today is created with Microsoft Word; this new add-in provides an unprecedented leap forward in the worldwide effort to make information available to all,” said George Kercher, secretary general of the DAISY Consortium.

Groups such as the World Health Organization and the World Blind Union estimate that more than 160 million people throughout the world are either blind or have a significant impairment to their

OAI-ORE Specification and User Guides
www.openarchives.org/ore/0.9/toc.html

OAI-ORE Discussion Forum
groups.google.com/group/oai-ore

Microsoft ‘Save as DAISY’ Translator for Word
www.microsoft.com/enable/ voices/daisy

Sonata Software DAISY Plug-In for Word
www.sonata-software.com/products/wordplug-in

DAISY Digital Talking Book Standard
www.daisy.org

Microsoft ‘Save as DAISY’ Plug-In Integrates NISO Digital Talking Book Standard

Microsoft, the DAISY Consortium, and Sonata Software have released the production version of the “Save to DAISY,” translator for Microsoft Word. This Word plug-in converts documents from Word’s Open XML format to the DAISY format that is defined in the NISO standard, Specifications for the Digital Talking Book (ANSI/NISO Z39.86). The DAISY Consortium is the Maintenance Agency for the NISO standard.

This free and open-source software provides feature-rich, structured information to persons who are blind or print disabled, and can assist in the creation of e-books. DAISY Digital Talking Books (DTBs) go far beyond the limits imposed by analog audio books or commercial digital audio books. In a DAISY book, the audio is synchronized with the textual content and images, providing an accessible and enriched multimedia reading and learning experience. A DAISY book also supports multiple outputs, such as Braille and large print.

“Microsoft’s initiative to put Save as DAISY XML in Microsoft Word is the first step to bring fully accessible content to people who are blind or who have a print disability. We know that much of the information in documents today is created with Microsoft Word; this new add-in provides an unprecedented leap forward in the worldwide effort to make information available to all,” said George Kercher, secretary general of the DAISY Consortium.

Groups such as the World Health Organization and the World Blind Union estimate that more than 160 million people throughout the world are either blind or have a significant impairment to their

OAI-ORE Specification and User Guides
www.openarchives.org/ore/0.9/toc.html

OAI-ORE Discussion Forum
groups.google.com/group/oai-ore

Microsoft ‘Save as DAISY’ Translator for Word
www.microsoft.com/enable/ voices/daisy

Sonata Software DAISY Plug-In for Word
www.sonata-software.com/products/wordplug-in

DAISY Digital Talking Book Standard
www.daisy.org

Microsoft ‘Save as DAISY’ Plug-In Integrates NISO Digital Talking Book Standard

Microsoft, the DAISY Consortium, and Sonata Software have released the production version of the “Save to DAISY,” translator for Microsoft Word. This Word plug-in converts documents from Word’s Open XML format to the DAISY format that is defined in the NISO standard, Specifications for the Digital Talking Book (ANSI/NISO Z39.86). The DAISY Consortium is the Maintenance Agency for the NISO standard.

This free and open-source software provides feature-rich, structured information to persons who are blind or print disabled, and can assist in the creation of e-books. DAISY Digital Talking Books (DTBs) go far beyond the limits imposed by analog audio books or commercial digital audio books. In a DAISY book, the audio is synchronized with the textual content and images, providing an accessible and enriched multimedia reading and learning experience. A DAISY book also supports multiple outputs, such as Braille and large print.

“Microsoft’s initiative to put Save as DAISY XML in Microsoft Word is the first step to bring fully accessible content to people who are blind or who have a print disability. We know that much of the information in documents today is created with Microsoft Word; this new add-in provides an unprecedented leap forward in the worldwide effort to make information available to all,” said George Kercher, secretary general of the DAISY Consortium.

Groups such as the World Health Organization and the World Blind Union estimate that more than 160 million people throughout the world are either blind or have a significant impairment to their

OAI-ORE Specification and User Guides
www.openarchives.org/ore/0.9/toc.html

OAI-ORE Discussion Forum
groups.google.com/group/oai-ore

Microsoft ‘Save as DAISY’ Translator for Word
www.microsoft.com/enable/ voices/daisy

Sonata Software DAISY Plug-In for Word
www.sonata-software.com/products/wordplug-in

DAISY Digital Talking Book Standard
www.daisy.org

Microsoft ‘Save as DAISY’ Plug-In Integrates NISO Digital Talking Book Standard

Microsoft, the DAISY Consortium, and Sonata Software have released the production version of the “Save to DAISY,” translator for Microsoft Word. This Word plug-in converts documents from Word’s Open XML format to the DAISY format that is defined in the NISO standard, Specifications for the Digital Talking Book (ANSI/NISO Z39.86). The DAISY Consortium is the Maintenance Agency for the NISO standard.

This free and open-source software provides feature-rich, structured information to persons who are blind or print disabled, and can assist in the creation of e-books. DAISY Digital Talking Books (DTBs) go far beyond the limits imposed by analog audio books or commercial digital audio books. In a DAISY book, the audio is synchronized with the textual content and images, providing an accessible and enriched multimedia reading and learning experience. A DAISY book also supports multiple outputs, such as Braille and large print.

“Microsoft’s initiative to put Save as DAISY XML in Microsoft Word is the first step to bring fully accessible content to people who are blind or who have a print disability. We know that much of the information in documents today is created with Microsoft Word; this new add-in provides an unprecedented leap forward in the worldwide effort to make information available to all,” said George Kercher, secretary general of the DAISY Consortium.

Groups such as the World Health Organization and the World Blind Union estimate that more than 160 million people throughout the world are either blind or have a significant impairment to their
vision. This number does not even begin to address the additional hundreds of millions of people with physical, developmental, or learning disabilities who can benefit from the rich applications of DAISY.

“Our new Save as DAISY XML functionality for Microsoft Word has the potential to break down barriers for millions of visually impaired individuals around the world and enhance the experience for virtually anyone who loves to read,” said Chris Capossela, senior vice president of the Information Worker Product Management Group at Microsoft. “This tool will make it easier for anyone from a child writing to his or her grandparent, to a government agency providing vital information to its citizens to create accessible content.”

The software works with Microsoft Word 2007, 2003, and Word XP. The open source nature of the Open XML to DAISY XML translation project enables technologists to utilize the source code and other resources for their own applications.

Comments Invited on Interactive Data Dictionary for Information Services Metrics

NISO has posted a new web version of the American National Standard Z39.7-2004, Information Services and Use: Metrics & statistics for libraries and information providers—Data Dictionary, and invites comments from the information community on both dictionary content and web-based usability.

The purpose of the current Z39.7 Data Dictionary remains the same as was for previous versions: to assist the information community in the identification, definition, collection, and interpretation of statistical data used to describe the current status and condition of libraries in the United States. However, it also differs significantly from its predecessors in its approach. This revision absorbs the de facto definitions established through the national program for collecting data about libraries (i.e., the National Center for Education Statistics (NCES) survey program), and absorbs relevant US association definitions, methods and practices of data collection.

In addition, the Data Dictionary has recently been approved to go on to continuous maintenance under ANSI. This will allow revisions to be made to the standard on an ongoing basis, rather than the typical five-year schedule, to allow for the most current and useful definitions to be included. A Standing Committee, chaired by Dianne L. Carty (Massachusetts Board of Library Commissioners), has been established to manage the maintenance process. The new website includes the procedures for the new continuous maintenance process as well as forms for submitting suggested changes.

Carty said, “This new launch of the Z39.7 Data Dictionary takes the standard from a static document to an interactive web-based utility that allows users to contribute to the usefulness of the dictionary by helping to identify standard definitions, methods, and practices that are relevant to library statistics activities.” Carty is Head of State Aid & Data Coordination, Massachusetts Board of Library Commissioners.

“The value of the approach taken by the Standing Committee is that it recognizes the guidelines and best practices in the area of library statistics across the community, not only at the national survey level,” said Todd Carpenter, Managing Director of NISO.

Members of the Standing Committee represent the American Library Association (ALA), Association for Library and Information Science Education (ALISE), Association for Research Libraries (ARL), Association of College and Research Libraries, Chief Officers of State Library Agencies (COSLA), Committee on Research and Statistics (CORS), Government Agency Library, Institute of Museum and Library Services (IMLS), International Organization for Standardization (ISO), Massachusetts Board of Library Commissioners, National Center for Education Statistics (NCES), National Commission on Libraries and Information Science (NCLIS), Project Counter, and Special Libraries Association (SLA).

Proposed NISO CORE Standard to Form Information Bridge Between ILS and ERMS

NISO has approved work on a new standard project, CORE (Cost OfResource Exchange), to facilitate the exchange of cost, fund, vendor, and invoice information between Integrated Library Systems (ILS), Business Systems, Electronic Resource Management Systems (ERMS), and other interested parties such as Subscription Agents.

The new work item was first proposed by Jeff Aipperspach, Product Manager, Serials Solutions; Todd Koppel, A/Gent Verso (ILS) Product Manager, Auto-Graphics, Inc.; and, Ed Riding, Technical Product Manager, SirsiDynix. Describing the project, Riding said, “We have three main goals. First, we want to develop and refine the list of data elements exchanged between an ERMS, ILS, Business Systems and other interested parties holding acquisitions metadata to support the population of the ERMS with financial and vendor information in the automated system. Second, we intend to create a transport protocol useful in moving these data elements from one system to another. Third, we will write a small number of use cases which will help all parties understand the capabilities of the protocol.”

“With real-time lookups and cost-per-click and other cost-related reports in the ERMS, users will be able to avoid manually entering the same data in two different systems,” said Patricia Brennan, Manager, Evaluative Products, Thomson Reuters, and Chair of the NISO Business Information Topic Committee that oversees the new Working Group. “Using defined XML data schemas, the standard will provide a common method of requesting cost-related information from an ILS for a specific electronic resource. Once implemented, this standard could well be expanded to include other elements.”

Todd Carpenter, Managing Director of NISO, added: “We expect great interest from librarians who have implemented an ERMS, ERMS creators, and ILS creators and hope that they will comment on the process every step of the way toward a final standard.”

The Working Group expects to complete the XML schema by Nov. 1, 2008 and have an approved American National Standard by Spring 2009. An interest group e-mail list is available for those who want to follow the progress of the project. Contact Karen Wetzel (kwetzel@niso.org) for further information.

New Versions of ONIX for Serials SRN and SPS Message Formats Piloted

EDHEUR, the international group coordinating development of the
AIIM Offers PDF/Archive Training Courses

AIIM is now offering a training course on the PDF/Archive (PDF/A) standard, ISO 19005-1, Document management - Electronic document file format for long-term preservation - Part 1: Use of PDF 1.4 (PDF/A-1).

The PDF/Archive standard was developed as a joint effort between AIIM and NPES and further developed by an International Standards Organization (ISO) working group. It provides a standard file format that can be used for long-term preservation of electronic documents. PDF/A is only one component of a comprehensive digital preservation strategy that includes records management policies and procedures as well as defined quality control processes.

Since its issuance in 2005, the PDF/A standard has been adopted at much higher rates in Europe than in the United States, even though it has been incorporated into products such as intelligent copiers. Believing that low implementation is due in part to lack of understand-

If you want to get more precise expression of the enumeration and chronology of issues included in a subscription product.

The entire program is available as either Web-based training, public two-day classes held around the country, or in private classes. Public classes scheduled for the last part of 2008 are in Toronto (August), Phoenix (September), Silver Spring, MD (September/October), and Chicago (November). Check the AIIM PDF/A Training webpage for exact dates and schedule changes and to register. Contact Betsy Fanning, AIIM Director, Standards and Member Services (bfanning@aiim.org or 301-755-2682) to schedule private classes.

Additional support for PDF/A implementation is provided through AIIM’s online PDF Experts Corner, where an international community of experts shares implementation experiences and best practices. Among the resources available on the website is a listing of PDF/A compliant products, both software and hardware.

Techstreet and Intertek Launch New Standards Store

Thomson Reuters, the world’s leading source of intelligent information for business and professionals, has partnered with Intertek, one of the world’s largest independent testing, inspection and certification companies, to launch a new online standards store. This new store is supported by Techstreet (a division of Thomson Scientific, a business unit of the Thomson Corporation) and offers a wide range of industry codes and standards to ensure safety, quality, and compliance in manufactured products. Product testing and certification customers also can receive updates on standards via custom email alerts.

“It is more critical than ever for manufacturers and retailers to have the most current standards for their products,” said David Brown, Executive Vice President, Corporate Markets for the Scientific business of Thomson Reuters. “Using the proven e-commerce methods from Techstreet, this store is speeding the delivery of such standards to the organizations who want to get their products to market faster.”

Visitors to the Intertek Standards Store can search, order and download industry codes, standards and technical information from over 350 of the world’s leading standards developing organizations, including ANSI, AIIM, ASQ, ASTM BSI, DIN, IEEE, INCITS, ISO, NISO, and many others. Customers can purchase single copies in print or PDF format, and receive 10% off Techstreet prices. They
Web Portal Created for Case Studies on University Digital Records

The Society of American Archivists (SAA) has created a web portal to collect and disseminate case studies from college and university archivists working on solutions for born-digital records. Suggested elements for the case studies include institutional context and background; nature of the records; key challenges anticipated; appraisal, processing, and preservation accomplished prior to the case study; resources; analysis; and future plans.

Seven initial case studies were posted, resulting from a workshop attended by more than two dozen university archivists on “The Development of Case Studies for the Effective Management of University Digital Records, held at the Bentley Historical Library, University of Michigan, in September 2007. Both the workshop and the portal launch were supported by funding from The Andrew W. Mellon Foundation.

The existing case studies address such projects as: promotion and tenure archival records, university committee and planning records, data-intensive social science research, and the development of ARMA’s standard on the digital records conversion process.

SAA is encouraging and soliciting new case studies, including those for work-in-progress. An online submission form and instructions on how to submit studies are both available from the Campus Case Studies portal. Copyright of posted case studies remains with the author.

ISBN Standard Undergoes Systematic Review

In 2005, a major revision of the International Standard Book Number (ISBN) standard (ISO 2108) was published that replaced the 10 digit number with a 13 digit number using an EAN prefix. Although the standard was issued in 2005, the mandatory use of the 13-digit ISBN was set for January 1, 2007 to allow everyone in the supply chain to update procedures and information systems to support the changes.

Another significant change in the 2005 edition was explicitly identifying electronic publications, chapters and parts of publications, and journal articles (when available separately) as materials eligible to receive an ISBN.

ISO standards are required to undergo a systematic review at least every five years to ensure that they are kept up-to-date and are still relevant in the community of users. Recently, ISO mandated that a new standard have its first systematic review three years after publication. (A revision is also considered a “new” standard.) So although, the 13-digit ISBNs have only been in mandatory use for a little over a year, the standard is now undergoing a review.

This early review will give publishers, libraries, information system vendors, and other users an opportunity to identify any problems encountered in the standard’s implementation that might require a change to the standard. NISO Voting Members will be sending in their comments on the standard and recommendations for whether to reaffirm it by August 29, 2008.

Many of the implementation details of the ISBN are in the Users’ Manual, rather than the standard. This document can be updated separately without rebalotting the standard, as long as it conforms to the standard’s requirements.

PDF format becomes ISO standard

The Portable Document Format (PDF), undeniably one of the most commonly used formats for electronic documents, is now accessible as an ISO International Standard - ISO 32000-1. This move follows a decision by Adobe Systems Incorporated, original developer and copyright owner of the format, to relinquish control to ISO, who is now in charge of publishing the specifications for the current version (1.7) and for updating and developing future versions.

“By releasing the full PDF specification for ISO standardization, we are reinforcing our commitment to openness”, says Kevin Lynch, Chief Technology Officer at Adobe. “As governments and organizations increasingly request open formats, maintenance of the PDF specification by an external and participatory organization will help continue to drive innovation and expand the rich PDF ecosystem that has evolved over the past 15 years.”

ISO Secretary-General Alan Bryden comments: “As an ISO standard, we can ensure that this useful and widely popular format is easily available to all interested stakeholders. The standard will benefit both software developers and users by encouraging the propagation and dissemination of a common technology that cuts across systems and is designed for long term survival.”

The new standard, ISO 32000-1, Document management – Portable document format – Part 1: PDF 1.7, is based on the PDF version 1.7 developed by Adobe. This International Standard supplies the essential information needed by developers of software that create PDF files (conforming writers), software that reads existing PDF files and interprets their contents for display and interaction (conforming readers), and PDF products that read and/or write PDF files for a variety of other purposes (conforming products).
International Committee on Information and Documentation Meets in Stockholm to Advance Standards Work

KAREN WETZEL, NISO’S STANDARDS PROGRAM Manager, and I were joined by nearly a dozen other technical experts who traveled to Stockholm in May for the annual meeting of ISO’s Technical Committee 46 (Information and Documentation) and its four subcommittees: SC 4, Technical Interoperability; SC 8, Quality – Statistics and Performance Evaluation; SC 9, Identification and Description; and SC 11, Archives and Records Management. In addition to the plenary meetings for these committees, there were separate working group meetings for the numerous projects underway. The meeting was hosted by the Swedish Standards Institute (SIS) and was attended by representatives from some 30 countries.

NISO has long served as the U.S. Technical Advisory Group administrator for TC46, an official designation from the American National Standards Institute (ANSI) that gives NISO the responsibility for submitting the U.S. vote and comments on TC46 draft standard ballots and for identifying U.S. participants on each of TC46’s subcommittees and working groups. This is a critical and not well-known aspect of NISO’s role in standards development. While a great deal of advancement in consensus and standards development takes place in the United States, other regions and countries are also pushing technological and information boundaries and facing challenges of similar size and scope as the U.S. domestic market. Often, because of organizational or governmental structures, other countries are better positioned to advance technological, infrastructure, or procedural standards more quickly than the US community. Long gone are the days when national standards could be applied to information exchange, particularly given the rapid changes in creation and distribution of digital publications.

The Stockholm meeting was the first in which I represented NISO in its new role of Secretariat of SC9, Identification, and Description, with Professor Sama Gyun Oh (Sung Kyun Kwan University) as the new Chair, representing the Korean Agency for Technology and Standards (KATS). This subcommittee is responsible for many critical and heavily used standards in our community, including the International Standard Book Number (ISBN) and the International Standard Serial Number (ISSN).

A number of SC9 working groups held meetings to initiate new projects or to advance work of existing projects. These included:

• a revision and combination of the two standards for monolingual (ISO 2788) and multilingual (ISO 5964) structured vocabularies and thesauri, which will have broad application in the development of ontologies and in semantic web technology;
• the International Standard Name Identifier (ISNI) that will address such entities as parties in the content creation process (beyond author), the identification of fictional characters or pseudonyms, and the relationships of multi-party entities, and will be issuing a public call for a registration authority;
• the International Standard Text Code (ISTC), which had been held up for logistical reasons relating to the appointment of a registration authority, but is now pushing toward publication;
• a revision of the International Standard Music Number (ISMN) that had a final draft approved in April with minimal editorial comment and should be published later this year;
• a revision of the Bibliographic References standard (ISO 690), which is readying a draft for ballot; and
• development of a draft terms of reference for a forthcoming revision of the International Standard Recording Code (ISRC).

The project within SC 9 that generated the most comments and discussion was the new draft standard on the Digital Object Identifier (DOI) System, which has raised questions surrounding how the DOI would interrelate and/or incorporate other standard identifiers. Comments were also submitted relating to restricting the character set available for inclusion in DOI identifiers and removing the description of the “Han-
Long gone are the days when national standards could be applied to information exchange, particularly given the rapid changes in creation and distribution of digital publications.
The ISTC is an identifier that identifies the text of a work and links the different print and digital manifestations of that text. The registration authority details of this standard are being worked out and should be in operation during the third quarter of 2008. The ISNI is a new identifier being developed that will identify the public identity of a party (person, legal entity, group, or even fictitious character) across multiple fields of creative activities. This new identifier will be matched with certain metadata that will unambiguously identify the party and will include: its name, date of birth/incorporation, place of origin, the role the party has played relating to the work, and related identifiers.

Taking these scenarios, the group will explore the potential impact and engagement of the various stakeholders in the community. The group has begun work on developing scenarios and hopes to have a working draft of a standard available for public review by mid-2009, with eventual publication in 2010.

Wrapping up the segment of the program on identification was Clifford Lynch, Executive Director of the Coalition for Networked Information (CNI). Mr. Lynch discussed the broader issues surrounding name and party identification. Earlier this year CNI hosted a meeting to discuss name authority and the various projects that are underway in the community related to this issue. The meeting identified more than 25 projects directly represented at the meeting and a large number of additional projects that are ongoing and known to the group. It is almost certain that there are more projects that are under way elsewhere. Among the key problems noted by Mr. Lynch was the fact that the problem is so large and pervasive that there was no simple way to corral all of the many projects and combine efforts. Several communities have their own needs, own metadata, and own systems that rationalization might not be possible. This is despite the many potential services that an integrated identification scheme could provide, particularly with regard to research.

During the Discovery section of the program, Michael Healy again spoke on the issue of Standards for Discovering Online Book Content, which was followed by Jason Price, who is the E-Journal Package Analyst at SCELC and Life Sciences Librarian at the Claremont Colleges, who discussed the NISO/UKSG Knowledge Bases and Related Tools (KBART) initiative. Mr. Healy discussed the work of the AAP Digital Issues Working Group, which has subsequently become an ongoing working group under BISG’s auspices. The group is working toward defining a set of HTTP transitions between a publisher’s digital archive and those of syndication partners. The groups has developed a set of use cases that include transactions such as: Search for Phrase, Enumerate ONIX info, Enumerate Book Info, Browse by pages, Get Full Page Metadata, Get Thumbnail Media, among others. The group is finalizing its Technical specifications and hopes to prototype the specification among the participating groups during this summer. The goal is to have a published specification available in early fall of 2008.

Jason Price discussed the work of the KBART technical working group, of which he is a member. The KBART group is an outgrowth of the preliminary research work commissioned by the UK Serials Group (UKSG) on the chain of metadata that underpins OpenURL link resolver systems. This metadata supply chain provides information about web resources, the holdings of a library and through the OpenURL system, provides users links to the appropriate copy for their access. Often the system breaks because of inaccurate or out-of-date data. The KBART group is working to improve awareness of the OpenURL structure, defining common terminology in the community, improving the quality and timeliness of publisher data, and is developing educational plans to improve the use and understanding of these systems.

The program ended with Nathan Robertson, speaking on the topic of compliance and use. Mr. Robertson is the Electronic Resources Librarian at the Thurgood Marshall Law Library, University of Maryland School of Law. He also serves as the co-chair for the joint License Expression Working Group (LEWG), which is a partnership between NISO, EDItEUR, and the Publishers License Society (UK). The details of copyright and contract law, as well as the complexity of most subscription licenses, add to the confusion surrounding what users may or may not do with subscribed content. Common misunderstandings of the doctrine of Fair Use contribute to the confusion. While NISO and its partners have made some progress with improving the dissemination about license terms through work on the ONIX-PL messages and the ERMI data dictionary, much work remains. The LEWG has invested considerable time in making ERMI encoding easier, through mappings and community education. Also, EDItEUR has released an ONIX-PL Editor, which allows for easier creation and encoding of license terms without needing to understand the underlying XML code. Mr. Robertson ended his presentation with a discussion of the Shared E-Resources Understanding (SERU), which he postulated as a potential option for simplifying the licensing process.

The meeting was a tremendous success. With so much activities taking place in both the publishing and library communities, providing a discussion forum where the two groups can learn from the overlapping activities of the other is critical. BISG and NISO have committed to continuing the series of programs in the coming year and will likely host another similar free seminar next year before the start of the next ALA Annual conference in 2009.

The slide presentations described in this article are available at the NISO website at: http://www.niso.org/news/events/2008/ala08/nisosig08

Todd Carpenter (tcarpenter@niso.org) is managing director of NISO and publisher of ISQ.
Mining for Meaning

IN DAVID LODGE’S 1984 NOVEL, SMALL WORLD, a character remarks that literary analysis of Shakespeare and T.S. Eliot “would just lend itself nicely to computerization....All you’d have to do would be to put the texts on to tape and you could get the computer to list every word, phrase and syntactical construction that the two writers had in common.”

This brave new world is upon us, but the larger question for Google, OCLC, and the Open Library, among other purveyors of warehoused metadata and petabytes of information, is how to achieve meaning. One of the brilliant insights derived from Terry Winograd’s research and mentoring of two bright Stanford Ph.D students is that popularity in the form of inbound links does matter for web pages, at least. In the case of all the world’s books turned into digitized texts, it’s a harder question to assign meaning without popularity, a canon, or search queries as a guide.

Until recently, text mining wasn’t possible at great scale. And as the great scanning projects continue on their bumpy road, the mysteries of what will come out of them have yet to emerge into meaning for users.

**Nascent standards**

In his May NISO presentation, Bill Kasdorf pointed out several XML models for books including NISO/ISO 12083, TEI, DocBook, NLM Book DTD, and DTBook. These existing models have served publishers well, though they have been employed for particular uses and have not yet found common ground across the breath of book types. The need for a standard has never been clearer, but it will require vision and a clear understanding of solved problems to push forward.

After the professor in Small World gains access to a server with digitized books, he grows giddy with the possibilities of finding “your own special, distinctive, unique way of using the English language....the words that carry a distinctive semantic content.” While we may be delighted about the possibilities that searching books afford, there is the distinct possibility that the world of the text could be changed completely.

Another mechanism for assigning meaning to full text has been opened up by web technology and science. The Open Text Mining Interface is a method championed by Nature Publishing Group as a way to share the contents of their archives in XML for the express purpose of text mining while preserving intellectual property concerns. Now in a second revision, the OTMI is an elegant method of enabling sharing, though it remains to be seen if the initiative will spread to a larger audience.

**Sense making**

As the corpus lurches towards the cloud, one interesting example of semantic meaning comes in the Open Calais project, an open platform by the reconstituted Thomson Reuters. When raw text is fed into the Calais web service, terms are extracted and fed into existing taxonomies. Thus, persons, countries, and categories are first identified and then made available for verification.

This experimental service has proved its value for unstructured text, but it also works for extracting meaning from the most recent weblog posting to historic newspapers newly scanned into text via Optical Character Recognition (OCR). Since human-created metadata and indexing services are among the most expensive things libraries and publishers create, any mechanism to optimize human intelligence by using machines to create meaning is a useful way forward.

Calais shows promise for metadata enhancement, since full text can be mined for its word properties and fed into taxonomic structures. This could be the basis for search engines that understand natural language queries in the future, but could also be a mechanism for accurate and precise concept browsing.

**Glimmers of understanding**

One method of gaining new understanding is to examine solved problems. Melvil Dewey understood vertical integration, as he helped with innovations around 3x5 index cards, cabinets, as well as the classification systems that bears his name. Some even say he was the first standards bearer for libraries, though it’s hard to believe that anyone familiar with standards can imagine that one person could have actually been entirely responsible.

Another solved problem is how to make information about books and journals widely available. This has been done twice in the past century—first with the printed catalog card, distributed by the Library of Congress for the greater good, and the distributed catalog record, at great utility (and cost) by the Online Computer Library Center.

Pointers are no longer entirely sufficient, since the problem is not only how to find information but how to make sense of it once it has been found. Linking from catalog records has been a partial solution, but the era of complete books online is now entering its second decade. The third stage is upon us.

Jay Datema (jdatema@bookism.org) is the editor-in-chief of ISQ.
INSTITUTIONAL IDENTIFICATION is becoming more important in a networked electronic world. NISO is developing a standard for these identifiers in the information space building on the work of several projects including the Journal Supply Chain Efficiency Improvement Pilot, Ringgold Identify database and OCLC’s WorldCat Registry. The standard will be developed using a scenario and use-case approach to accommodate the large number of stakeholders involved in this area.

Institutional Identifiers have been around for a long time for many different purposes. In the information world they were used mainly as location definitions, to enable physical product to be delivered, so this included inter-lending codes such as the MARC Organization Code and OCLC Symbol, and Standard Address Numbers. There are also the financial and business coding such as D-U-N-S numbers, tax identifiers and government registries.

In the library world there are several systems, the ISO ISIL (International Standard Identifier for Libraries) run by national agencies and the WorldCat Registry run by OCLC. All these have very different purposes. Ringgold has developed its own registry of institutions which subscribe to scholarly publications currently used mainly by publishers to uniquely identify their customers.

Despite current projects, standards and processes, there exists a need for a community/standards based approach to managing institutional affiliations and organizational structure. In the world of big science, collaborations, country wide consortia and Web 2.0 this means we need a flexible approach to managing this information and its associated metadata as it supports various transactions in the physical and virtual world.

A group of organizations in the information space decided to develop a pilot to look at journal supply chain efficiencies in 2006. This group comprised The British Library, Swets, HighWire Press, several HighWire publishers and Ringgold and reports have been produced and posted on the website (www.journalsupplychain.org). One of the conclusions was that a standard needs to be developed which will include the many different ways that such an institutional identifier can be used, together with a business model and governance to ensure sustainability.

A proposal was made to NISO for a Working Group on Institutional Identifiers, which was accepted by the membership in January 2008. The Working Group has been put together under the leadership of Grace Agnew of Rutgers University and Tina Feick, Harrassowitz. Expressions of interest had been solicited from NISO members and the information world in general, with a huge response from a very broad group of stakeholders than originally envisaged. In order to accommodate these various groups it was decided to develop a scenario based workplan to ensure the Working Group would be
best equipped to accommodate the wide interests and needs represented by the community.

The primary goal for the I2 project is to develop an institutional identifier that is robust, scalable, simple to implement and to deploy in existing workflows and that can support any information workflow requiring an institutional identifier so that multiple institutional identifiers are not needed.

**Stakeholders**

Based on the call for participation and feedback received by NISO and the participants, the following stakeholders have been identified:

**PUBLISHING COMMUNITY:** Publishers, editors, reviewers, fulfillment systems, publisher manuscript systems

Intermediaries: Subscription agents, distributors, aggregators, hosting services

**ACADEME:** Academics, researchers, students, authors, funding bodies, academic administrators, institutional repositories

**LIBRARIES:** Library systems, consortia, ILS systems, ERM systems

**RIGHTS AGENCIES:** Licensing bodies, author rights societies, copyright centers

**ELEARNING:** eLearning vendors, services, content providers

**AGENCIES:** State-wide agencies, national purchasing bodies

**AUTHENTICATION:** Managed federations, OpenID

**PROFESSIONAL GROUPS:** Doctors

**Scenarios**

In developing the workplan, 2 types of scenarios will be developed: those based on topics or use cases; and those based on needs of discrete communities.

**SCENARIO A: ELECTRONIC RESOURCES SUPPLY CHAIN**

This scenario follows most specifically the work of the Journal Supply Chain Pilot and will deal with the issues raised in that work. These relate to the granularity of the identification of the institution. For example publishers are interested in data at the institutional level, whereas agents are interested at the specific account or library level. Another issue is how the publisher, agents and hosting systems are populated with the identifier and metadata and the changes which need to be made in all the back office systems to exploit the identifier and the metadata. The stakeholders in this scenario who will be invited to participate include libraries; agents; publishers; aggregators; distributors; hosting services and fulfillment services.

**SCENARIO B: ELEARNING**

The definition of entitlements to specific elements included in eLearning systems is complex and publishers and vendors need to understand the content, the licensing terms and also the issues of micro-licensing such items and graphics and pictures. Copyright of the content is also a contentious issue. The stakeholders in this scenario who will be invited to participate include publishers, aggregators; academics; rights holders and eLearning vendors

**SCENARIO C: RESEARCH EVALUATION AND FUNDING**

The issues around research evaluation are increasingly complex and international in scope. The traditional methods of evaluating research impact based on purely publication are changing to other metrics which include profiling, usage metrics, and other research outputs. Research outputs are also changing and include all research products that can be digitized and made available, such as data sets, images, lab notes, etc. Funding agencies frequently hold the institution accountable for compliance with reporting requirements, open access for research results, etc. Issues of attribution and allocation of credit are critical and impact university administrators, funding bodies, academics, researchers, and publishers, and the tools and systems they interact with.

**SCENARIO D: AUTHOR REGISTRIES**

Closely associated with research funding is the issue of correct attribution and disambiguation for the creators of works, who generally remain the copyright holder for their original works. Researchers may have multiple affiliations, which may differ for different research papers, for example a paper published as a member of a multi-institution center vs. a paper published as a faculty member of a university. The author’s affiliation is often critical for authenticating the author, providing avenues for follow up with an author and can be a critical credential for providing confidence in the resource. (For example, an overview of the research undertaken by a center may be deemed more authoritative if the author is affiliated with that center. Institutes within academic institutions may have their own separate affiliations (for example CNRS UMRs, or Howard Hughes Medical Institutes). Many publishers have their own author registries, but there is currently no way of identifying common authors and their affiliated institutions across publishers. Academic authors may deposit a work in multiple registries, depending on affiliations. Institutional identification is a critical data element for identification of the author in a registry and also for providing critical provenance information about a publication. The stakeholders in this scenario include publishers, editors, reviewers, authors, and rights agencies,

**SCENARIO E: INSTITUTIONAL REPOSITORIES**

Readership and usage of journal articles is increasingly provided by avenues beyond the traditional journal publication, including repositories sponsored by the author’s institution, the granting agency, the federal government, and by publisher repositories. Each of these dissemination avenues may receive versions of an article that may vary in content, from substantial changes as an article moves from preprint to publication stage through minor errata changes in the postprint stage. As the concept of the journal article moves from an immutable work to an evolving work with a documented lifecycle, the institutional identifier will play a critical role in tracking the location of work versions, providing critical provenance for each deposited version of the work, and tracking the rights ownership of the work.
SCENARIO F: LICENSING/MICRO-LICENSING
With content becoming increasingly granular, and electronic access ubiquitous, the ability to license parts of articles, chapters or images is emerging as an important new model for the distribution and reuse of information. Microlicensing provides opportunities for the reuse of parts of works or aggregations in other works of scholarship, such as the licensing of an image from one work for reuse in another. Microlicensing can provide greater impact for a creator’s work and provide avenues for enriching new works. However, the microlicensing of one work for reuse in another adds exponential complexity to issues of rights and ownership for these hybrid resources. The institutional identifier is a critical provenance element for documenting and tracking the lifecycle and ownership of microlicensed resources. An interesting example here is PLUS (www.useplus.com) which is standardizing licensing terms, definitions and uses for photography and illustration.

SCENARIO G: USAGE METRICS
Although this area has its own set of standards work, one of the issues with usage metrics is the unique identification of the entities involved. The work on definitions, compatibility and systems implementation would be linked to the work within this scenario. All the participants in the information space are stakeholders in this area of work.

SCENARIO H: INSTITUTIONAL COLLABORATION
There is increasing collaboration between participants in the information space as costs increase and the technical barriers to collaboration decrease. In addition to research collaboration, there are such collaborations as the Knowledge Exchange where consortia in the UK, Germany and the Netherlands are collaborating on collection development. Collaboration ranges from federated access to digital resources to collaborative storage and use of shared print collections. Cross licensing between organizations is becoming more commonplace and geographical and territorial definition are needing to be more explicit for licensing and rights. This means that the stakeholders start to include consortia, State-wide sand country-wide systems, national funding bodies, and international funding bodies.

SCENARIO I: AUTHENTICATION
Management of IP addresses has always been a time-consuming and problem area for content providers and consumers. The definition of the entity and the entitlements of that entity are frequently difficult to define and agree. Proxy servers and managed federations such as Athens are a popular choice for managing authentication this but new authentication systems such as Shibboleth and OpenID also need the same definitions. For organizations, the ability to provide a single point of access to commonly held but separately licensed resources, such as digital video collections, provides economies of scale technically while providing access that is both granular and secure. With the increase in cross-licensing and micro-licensing, entity definition is increasingly important. Vendors, institutions, consortia and consumers are all stakeholders.

Community Scenarios

SCENARIO M: MEDICAL COMMUNITY
The medical community is becoming more virtual with examples such as SERMO showing how collaboration and resource-sharing can benefit the whole community. In addition the medical research community has always been complex, with most researchers having multiple affiliations for research, teaching and clinical practice. The national medical associations want to be able to provide more information and resources to their members, and publishers, local and national government bodies, medical vendors and medical research companies and institutes need to be involved in this community effort.

SCENARIO P: PUBLIC LIBRARY COMMUNITY
Public libraries worldwide are characterized by the diversity of their organizations. There is no standard structure, some are publicly funded, others privately. Users may range from documented subscribers to anyone living within a specific geographic range. State–wide, national and consortial licensing are providing access to a increased base of resources but ensuring licensing that provides seamless access for users but accountability for the licensing library is very difficult. Resource sharing has always been important to this community and the ability to define exactly who they are and to what they are entitled is not always easy.

Work Plan
The F working group has been established in July 2008. This will be followed by scenario development, with leaders for each scenario selected from the development team. The first task of each of these sub groups will be to review current identifier standards and usage in their field. This will involve identifying and contacting stakeholders in the field, particularly professional organizations and standards bodies, such as the Internet2 community in higher education. There will then be a need to define identifier standards already in use for important entities (e.g., creators, the work, the institution) the metadata required to support the identifier needs for each scenario, each scenario for institutional identification and to determine the metadata fields which are common to all scenarios and those which are specific to their scenario. Issues such as granularity (e.g., the need to identify subunits of an institution) and compatibility with existing identifier practices must be developed. It is expected that a core set of metadata will be developed (similar to the concept of Dublin Core or the DOI core metadata set) that represents the minimal metadata needed for institutional identification, with the ability to add community specific metadata as needed. A workplan is proposed that supports independent work by all working groups with monthly meetings of the whole to ensure coordination and coherence. The workplan also commits to frequent updates for interested stakeholders. Wikis, blogs and other web-based collaboration tools will be used.

A working draft of the standard will be prepared together with a trial use plan to involve testing against all scenarios and review both internally and by all stakeholders and advisory groups. In parallel the Working
Group will identify one or more potential maintenance organizations and develop a sustainability model that ensures durability and usability for the institutional identifier. Of equal importance will be the determination and development of tools for adoption and implementation of this standard, particularly given the proliferation of institutional identifiers in current use for specific scenarios. After several reviews it will be released for ballot by the members following standard NISO procedure.

Existing standards and guidelines

In the electronic resources space there are already standards and guidelines which will benefit from an institutional identifier. Some of these include:

- ICEDIS (International Committee on EDI in Serials) has long been used to exchange EDI data between publishers and agents but has become less specific and standardized over the year. There are plans to complete a new XML based standard for the 2010 subscription year.
- ONIX has three main subsets – products and subscriptions; online holdings and release notification. Online holdings in particular are dependent on the accurate description of the institution holding the subscription.
- The Linking ISSN will now bring together all manifestations of a serial work and working with an institutional identifier will be able to be more specific about institutional entitlements.
- COUNTER is including a place in the COUNTER usage report specification for the institutional identifier so it is ready when the NISO standard is agreed. This will in turn complement the SUSHI standard.

New standards & guidelines

KBART is a NISO/UKSG Working Group on Knowledge Bases and Related Tools which is intended to streamline the exchange of metadata about entitlements between the content providers and libraries. This Group is building on the work of the UKSG Research Report on entitlements and link resolvers. Project TRANSFER is a UKSG Working Group preparing recommendations for exchange of data between publishers transferring titles with a focus on accurate subscription lists, entitlements, consortia entitlements and data on hosting platforms and changes to deals and bundles.

CORE is a NISO Working Group on Cost of Resource Exchange to create a standard to facilitate the exchange of cost, fund, vendor, and invoice information between Integrated Library Systems (ILS) and Electronic Resource Management Systems (ERMS). Author Registries already exist on a publisher by publisher basis, but there are two current cross-industry initiatives. CrossRef is proposing the CrossRef author registry with DOIs for authors, creating a profile including institutional affiliation(s). ResearcherID is a community being developed by Thomson Reuters with authors being invited to register a profile and identify their publications through disambiguation by linking to citation data.

Establishing a core

The novel scenario based approach to the development of this standard is important because of the large numbers of stakeholders who have identified themselves as having an interest in this area. Although this may be more cumbersome, the intention is to identify a core set of metadata and a format for the identifier at an early stage and to let each scenario work on more detailed areas at their own pace. This may allow early adoption and implementation by some communities. To keep all the stakeholders involved, wikis and blogs will be set up to allow increased participation. The time has now come for the information space to take control of institutional identification.

Helen Henderson is the Managing Director of Ringgold Ltd. (ringgold.com) and has worked as a librarian, publisher and subscription agent. Her wide view of the industry has enabled her to advise in areas of strategic development, such as business models for pricing and packaging of electronic publications of all types. Her company Information Power developed a database of customer information for publishers which now forms the basis for potential standards. Ringgold was formed in 2005 from the merger of several companies. Helen works on several standards committees and is currently the Joint Editor of Serials, the journal for the international serials industry, and on the UK Serials Group committee.
DEVELOPING TOOLS THAT CONNECT users to their library’s resources has been a long standing goal of software developers in the library world. Such tools critically depend on the development and consistent use of standards by key players in the information creation and delivery world.

LibX
LibX is an open source, freely available plug-in that enriches the Web with customized access to library content. It is available for both Firefox and Internet Explorer browsers. Since its inception in 2005, it has been tried out and adopted by over 400 libraries worldwide. The LibX Edition Builder is an automated, interactive web interface that a library can use to create its own customized edition of LibX which incorporates the resources this library offers. These resources include the library’s catalog, OpenURL resolver, off-campus proxy, subscription and local databases, and many more. Using the edition builder, edition maintainers create a configuration file that is then included in the plug-in their users install.

Once a user installs their library’s LibX edition, access to their library is integrated into the web browser experience in a variety of ways. Searching the catalog and other resources is available in a toolbar at the top of the browser interface. When selecting text, LibX offers to search for the selected text using a set of resources configurable by the user. This functionality is adaptive to the user’s selection; for instance, if the user selects a standard identifier such as an ISSN, LibX will offer to search for this item in the appropriate databases. LibX can also “sense” if users have access to a given page through the off-campus proxy and can inform the user of that.

LibX can “autolink” standard numbers it finds on a page, allowing users to search for the resource being referred to using that number in local holdings. In addition, a tooltip displays more information about the item associated with the standard number. For example, LibX uses OCLC’s xISSN service to retrieve such information as title, publisher, and whether a journal is peer-reviewed for each ISSN it detects in a page.

In addition to autolinking, LibX knows how to place “cues” carrying an institution’s logo into some pages, such as book seller sites like Amazon or Barnes and Noble. When users click on the cue, they are taken to the library’s catalog to see if their library owns that item. Thus, when users install their library’s LibX edition, they are taking their library’s resources with them as they use the web.

NISO Standards and LibX
Many of LibX’s features depend on or are facilitated by the use of NISO and other standards. For example, LibX exploits NISO standards Z39.88 (OpenURL) and Z39.84 (DOI), as well as international standards ISBN (ISO 2108) and ISSN (ISO 3297).

Put simply, OpenURL is a standard that allows the description of bibliographic resources as context objects. Based on context objects, OpenURL resolvers lead a user to an appropriate copy of an item and can also provide extended services such as bibliographic imports or access to an ILL system. The OpenURL 1.0 linking syntax is defined in the standard ANSI/NISO Z39.88. OpenURLs are constructed using a base URL which points to the library’s OpenURL resolver where the metadata about the specific item contained in the context object is stored.

An OpenURL resolver implementation maintains a knowledge-base of that library’s holdings and additional services it may offer. The use of OpenURLs has become ubiquitous in how libraries provide access to journal articles. However, whereas OpenURLs are typically placed only on publishers’ sites that can identify where a user is coming from and choose the correct resolver, LibX enables libraries to allow users to use their OpenURL resolver on any page.

Because OpenURL provides a standard way in which to express context objects, OpenURL is easy to configure. The LibX Edition Builder offers the option to integrate a library’s OpenURL resolver into that library’s LibX edition. All a user needs to provide is a name for this resolver and the base URL at which it can be found. Optionally, a logo may be included. To make this process even simpler, we use the edition maintainer’s current IP address to query OCLC’s OpenURL Resolver Registry for a library’s OpenURL settings. Based on survey feedback and log analysis, we know that this feature has streamlined the configuration greatly.

One click retrieval
One of the coolest features of an OpenURL-enabled LibX edition is the ability to select a reference string on a web page or in a PDF document and then retrieve, with a single click, a link to the appropriate copy of the article selected. To implement this functionality, LibX uses Google Scholar as a backend to perform a free text search. LibX then analyzes the results from Google Scholar to find the best match on the first page of results and transforms the links it finds into an
OpenURL and submits it to the library’s OpenURL resolver. We performed a systematic study and found that LibX can lead the user to an appropriate copy in about 81% of cases for a set of randomly selected references in widely read journals.

COinS (Context Objects in Spans) is an important use of OpenURL context objects that is integrated in LibX. Content publishers include “span” tags in pages that embed context objects. By rewriting “span” tags with library specific base URLs and placing library logos into the page, we provide the user with links to appropriate copies of objects. COinS are used extensively by Wikipedia pages, as well as other social reference sites, including CiteULike.

Use of Standardized Numbers and Identifiers
LibX screenscrapes web pages for several types of standardized numbers: ISSN numbers defined in ISO 3297, DOIs (Digital Object Identifiers) as defined in ANSI/NISO Z39.84, and ISBNs as defined by ISO 2108. These numbers are described by fixed formats that make them identifiable; for instance, ISSNs are usually written in the form 0000-0000. DOIs start with a “10.” prefix and contain at least one forward slash. In addition, ISSNs and ISBNs contain a checksum in the last digit. LibX analyzes text on any HTML page by traversing the representation of the page inside the browser, which is based on W3C’s DOM (Document Object Model). It evaluates each text node to see if it contains a potential match. If a string on a web page is found to match one of the standardized number types, LibX inserts a hyperlink to provide the user with the option to access library services. This rewriting of standardized numbers on web pages with a hyperlink to an appropriate electronic resource is also known as “autolinking.” To decrease the likelihood of falsely linking numbers that look like standard numbers, but are not, LibX tests if the checksum matches before inserting these links. Some websites have implemented measures to block autolinking. In this case, LibX inserts icons into the page that serve as visual cues that the user still has access to their library.

LibX also exploits standard numbers when displaying the right-click context menu after a user selects text, such as an ISSN, on a web page. The search options displayed are appropriate for the type of identifier selected (ISSN, ISBN, DOI, or PubMed ID). Tooltips are displayed for autolinked numbers or when the user right-clicks on selected numbers. LibX uses a number of web services to retrieve metadata to display to the user in tooltips. If the user selects an ISSN, LibX queries OCLC’s xISSN service for information such as the journal title and whether the journal is peer-reviewed. For ISBNs, LibX uses the companion xISBN service provided by OCLC. For DOIs, CrossRef’s metadata service is used, and for PubMed IDs, we use NCBI’s metadata service. All of the services are accessible via XML-based web services interfaces that accept the standard identifiers as query input.

For each LibX edition, the edition maintainer can customize how LibX should handle the use of standardized numbers. For instance, if a library does not maintain an OpenURL resolver, the edition maintainer can redirect ISSN queries instead to the library’s catalog.

Z39.50 (and why we’re not using it)
In the future, we would like to offer LibX users holdings and availability information from a library’s OPAC for resources the user may encounter on the web. We investigated using NISO’s Z39.50 standard for computer-to-computer communication for this purpose, since it is widely used in the library community. Unfortunately, Z39.50 is only partially suitable for this purpose. First, Z39.50 is not http-based. Therefore, Z39.50 clients cannot be implemented in JavaScript and would require the creation of distribution of custom client libraries. Although this limitation could be overcome by using a proxy, such as Index Data’s YAZ proxy, there is a second, more important limitation. Z39.50 does not specify a standardized format for holdings and availability information. Instead, practice varies widely between different vendors. Some include availability using the MARC 852 field, some use local MARC 999 fields, and others send information in an ad hoc format as additional information when requesting records of type OPAC. This lack of uniformity, combined with the fact that not all libraries run Z39.50 servers, prevents us from using this protocol at this point.

The DLF (Digital Library Foundation) is currently working to establish standards for the retrieval of this information from ILS. DLF created the ILS Discovery TaskGroup, whose work resulted in the March 2008 Berkeley Accord in which most major ILS vendors agreed to the creation of a basic framework for programmatic access to their ILSs. One key outcome of this effort is the creation of a standard API to retrieve library holdings information that avoids the limitations of Z39.50 in this respect.

Editor’s Note: See the report on the DLF ILS API meeting by Marshall Breeding following this feature.

Annette Bailey (ajbailey@vt.edu) is digital assets librarian at Virginia Tech University Libraries, Blacksburg, VA.
NEXT-GENERATION LIBRARY CATALOGS, or discovery interfaces, represent a major trend in library automation. Commercial and open source development of these products abounds. As more libraries become interested in providing interfaces into their collections using applications other than the online catalog module packaged with their integrated library system, improved interoperability is one of the key issues they face.

The current deployments of discovery interfaces largely depend on interactions with the ILS dependent on proprietary interactions with the local ILS requiring considerable expertise and effort. The Digital Library Federation (DLF) aims to foster a more standardized environment where discovery interfaces can work with any given ILS through a uniform set of application programming interfaces.

DLF forms the ILS-DI Work Group

Toward this goal, the DLF charged a workgroup called the Integrated Library System Discovery Interface Task Group, or (ILS-DI) to explore issues and to develop an applications programming interface (API) to address the interoperability issues between these two kinds of systems.

The committee of representatives from DLF partner institutions, chaired by John Mark Ockerbloom (University of Pennsylvania) also includes Terry Reese (Oregon State University), Patricia Martin (California Digital Library), Emily Lynema (North Carolina State University), Todd Grappone (University of South California), Dave Kennedy (University of Maryland), David Bucknum (Library of Congress), and Dianne McCutcheon (National Library of Medicine).

The ILS Discovery Interface Task Group, initially formed in the summer of 2007, documents its work at www.diglib.org/architectures/ilsdi/. This resource contains the group's official recommendations, example implementations, and a link to a wiki site with supplemental information related to ongoing work.

Assessing Needs and Interest

One phase of the task group’s work involved conducting a poll of libraries to assess their current experiences with discovery interfaces and future intentions. The poll revealed that the majority of libraries surveyed are not satisfied with the traditional OPAC and had plans to implement external products, either commercially provided or open source, to enhance the environment experienced by their users. More detailed information on the questions and responses of the survey are available on the Task Group’s web site.

The task group produced a draft that described the areas of functionality involved in the interactions between discovery interfaces and ILS products and proposing one or more protocols that could serve as bindings. Whenever possible, the committee relied on existing standards or protocols for the bindings. This draft, now superseded, was made available to the broader library and development community for reaction and comment.

The Berkeley Accord

On March 6, 2008, the DLF convened a meeting on the UC Berkeley campus that brought the Task Group together with representatives from organizations involved in developing integrated library systems or discovery interface. The purpose of this group involved discussion of the proposed API and an attempt to reach some level of consensus on what parts of the recommendations could be adopted and implemented.

The development community represented at that meeting included SirsiDynix, Innovative Interfaces, LibLime, Polaris Library Systems, VTLS, Ex Libris, and Talis, BiblioCommons, OCLC, and Medialab Solutions. Of this group, some develop ILS products, others discovery interfaces, and some offer both types of systems.

The discussions of this meeting identified three areas essential for a basic programmatic interface between an ILS and discovery systems. These functions included harvesting, availability, and linking. Harvesting involves the wholesale and incremental export of records from the ILS. The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) addresses this functionality and was proposed as the preferred binding for this function. Availability involves providing a mechanism for a discovery interface to determine the real-time circulation status of an item known to the discovery interface through a harvested record. The proposed binding for availability would be a simple REST interface that returns item status data. Linking provides a stable mechanism for invoking a page for a given item that provides detailed information and any relevant services, such as placing a hold or recall.

This led to the focus on these three functions (harvesting, availability, and linking), with many of the details to be worked out in the future. Each of the developer representatives was polled about their commitment toward supporting these recommendations. Most voted to support the recommendations, though many stipulated qualifications depending on how the details were resolved in the future. Innovative Interfaces abstained with the same basic concern. The general consensus surrounding these functions as the basis for the ILS Basic Discovery Interface, has also been called the Berkeley Accord.

Revised Draft Specification

Following the March 2008 Berkeley meeting, the Task Group prepared a new document representing the recommendations for the interoperability profile. “DLF ILS Discovery Interface Task Group (ILS-DI) Technical Recommendation: An API for effective interoperability between integrated library systems and external discovery applications,” dated June 4, 2008, summarizes the activities of the process to date, presents a new draft of the APIs, and is available through the Task Group’s website.

The document takes the approach of multiple levels of the API, starting at the Basic Discovery Interfaces (BDI), or Level 1, that pro-
provides the minimal level of functionality for a discovery interface to work with an ILS, and included to levels of record harvesting, real-time availability, and linking. At this level, the discovery interface would rely on the native OPAC of the ILS for many aspects of functionality. Levels 2 through 4 provide protocol support that allows the external discovery interface to take on increasing levels of OPAC functionality. At Level 4, the discovery interface can completely replace the OPAC, including such features for specific library types as managing fines, e-commerce payments, and course reserves.

The revised document included specific functions and bindings that would constitute API at each of the four levels. Though the number and complexity of the API elements increases for each level, the task group aimed to use existing standards and protocols when possible.

In addition to creating the document, at least one task group member, John Mark Ockerbloom (University of Pennsylvania) has created a prototype reference implementation of the Level 1 API against a database of online books.

Berkeley II: Engaging the Developer Community

Following the release of the revised draft specification, next steps included a second round of discussion with potential implementers to discuss the recommendations in some detail to identify problems, set priorities, gauge interest, and work toward initial implementations.

The DLF called a second meeting, again on the UC Berkeley campus, on August 7, 2008, including Task Group members and the developer community. Developers represented at this meeting included SirsiDynix, Ex Libris, VUFind (Villanova University), the eXtensible Catalog (University of Rochester), Innovative Interfaces, Talis, OCLC, LibLime, Polaris Library Systems, and BiblioCommons.

The primary agenda for the meeting involved focusing on the Level 1 functions, or the ILS Basic Discovery Interface. Success established at this basic level will provide a foundation for extending the API through the additional recommended functions.

Through the course of the day, participants worked through the four functions that comprise the BDI.

HarvestBibliographicRecords

The function of harvesting data from the ILS to be ingested into a discovery interface relies on OAI-PMH. Some of the issues of OAI-PMH relative to the BDI involve the format of the records, how to deal with deleted records, and details related to harvest intervals. Participants agreed that Dublin Core would stand as the minimal level record, but MARC-XML would be highly recommended for systems that support MARC. Deleted records present complications because ILS’s have a variety of approaches in dealing with items that have been withdrawn from the collection. The group agreed that if the item was no longer available for discovery then it should be presented on an OAI-PMH response as a deleted record. A full harvest involves an OAI-PMH request with no data parameter. Incremental harvest involve providing the date option, indicating the interval desired. Deleted records must be provided for both full and incremental harvests.

HarvestExpandedRecords

This function addresses the need to provide additional information to a discovery interface beyond the basic bibliographic record. These expanded records would provide information such as holdings associated with each record. The components of this response would include the bibliographic identifier, the bibliographic record, and the associated MARC holdings records. Other item-specific elements that should be returned include the location, call number, format, availability, barcode identifier, and other housekeeping data maintained by the ILS. Discovery Systems in need of this item-level information might choose to use HarvestExpandedRecords instead of HarvestBibliographicRecords.

GetAvailability

This function allows a discovery system to query the ILS regarding the real-time status of a title or item. In addition to the basic availability information already specified, participants agreed it would be useful to add location as text strings as an optional response element. Options apply for either title or item level availability. For a title-level request, the response will include the availability messages of each of the associated items.

GoToBibliographicRequestPage

At Level 1, expected functionality involves a hand-off back to the ILS OPAC for user functions such as placing a hold or recall. This function requires documenting and supporting a template for a persistent URL that will invoke a page that provides a set of services for the item specified.

Current status

The discussions of the August meeting resulted in the need for some, though fairly minor, revisions in the recommendations for the Level 1 Basic Discovery Interface API. The Task Group will be reconvened to produce these revisions. Developers were highly encouraged to proceed with implementations. Developers identified tools needed for implementations such as a reference implementation and validation tools. Ideally, a set of service descriptions in WSDL (Web Services Description Language) would be produced to fully document the API. Talis volunteered to work on a validation tool for the XML schemas involved in the API. Developers were encouraged to proceed with at least experimental implementations, with the DLF Fall Forum on November 12, 2008 as the next opportunity for developers to demonstrate components and projects that support some aspect of the ILS-Discovery Interface APIs.

Marshall Breeding is the Director for Innovative Technology and Research for the Vanderbilt University Library and the creator of Library Technology Guides (www.librarytechnology.org). Breeding participated in both meetings described in this report.
Standards in Development: July 2008

Listed below are the NISO working groups that are currently developing new or revised standards, recommended practices, or reports. Refer to the NISO website, Newsline, and Information Standards Quarterly for updates on the working group activities.

DSFTU stands for Draft Standard for Trial Use; RP = Recommended Practice; TC = Topic Committee.

<table>
<thead>
<tr>
<th>WORKING GROUP</th>
<th>STATUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of Resource Exchange (CORE)</td>
<td>Working group formation finalized.</td>
</tr>
<tr>
<td>Co-chairs: Ed Riding, Ted Koppel</td>
<td></td>
</tr>
<tr>
<td>DAISY/NISO Standard Advisory Committee</td>
<td>Z39.86, Specifications for the Digital Talking Book</td>
</tr>
<tr>
<td>Chair: George Kerscher</td>
<td>Proposal to revise at ballot by NISO members.</td>
</tr>
<tr>
<td></td>
<td>Ballot closes Friday, August 29, 2008</td>
</tr>
<tr>
<td>Exchange of Serial Subscription Information</td>
<td>Field testing:</td>
</tr>
<tr>
<td>Joint Working Project with EDItEUR</td>
<td>Serial Release Notification (SRN), v. 0.92 (May 2008)</td>
</tr>
<tr>
<td>Co-Chairs: Priscilla Caplan, Richard Gedye</td>
<td>Serials Products and Subscriptions (SPS), v. 0.92 (June 2008)</td>
</tr>
<tr>
<td>Chair: Priscilla Caplan</td>
<td></td>
</tr>
<tr>
<td>Institutional Identifiers</td>
<td>Standard in development</td>
</tr>
<tr>
<td>Co-Chairs: Grace Agnew, Tina Feick</td>
<td></td>
</tr>
<tr>
<td>Knowledge Base and Related Tools (KBART)</td>
<td>RP in development</td>
</tr>
<tr>
<td>Joint project with UKSG</td>
<td></td>
</tr>
<tr>
<td>Co-Chairs: Peter McCracken, Charlie Rappel</td>
<td></td>
</tr>
<tr>
<td>License Expression</td>
<td>Mapping ONIX-PL to ERMI and ONIX-PL ERMI</td>
</tr>
<tr>
<td>Joint project with DLF, EDItEUR, and PLS</td>
<td>Encoding Format issued.</td>
</tr>
<tr>
<td>Co-Chairs: Nathan Robertson, Alicia Wise</td>
<td>Reformation into two working groups to separately address ongoing issues of ONIX-PL and ERMI underway.</td>
</tr>
<tr>
<td>Metasearch Initiative TG2, Collection and Service Descriptions</td>
<td>Z39.91-200X, Collection Description Specification</td>
</tr>
<tr>
<td>Chair: Juha Hakala</td>
<td>Z39.92-200X, Information Retrieval Service Description Specification</td>
</tr>
<tr>
<td>NCIP Implementers Group</td>
<td>IZ39.83, NISO Circulation Interchange Protocol (NCIP)</td>
</tr>
<tr>
<td>Chair: Candy Zemon</td>
<td>Two-part revision at ballot by NISO members.</td>
</tr>
<tr>
<td></td>
<td>Ballot closes Sunday, August 3, 2008</td>
</tr>
</tbody>
</table>
From start to finish, EBSCONET® simplifies the acquisition, administration, access and evaluation of your e-resource collection.

With extensive access and registration support as well as a plethora of up-to-the-moment e-journal data at your fingertips, EBSCONET delivers on what many competing systems only promise to provide – true information resource management.

With more than 3 million subscription orders handled, 20 million records of rich historical data, 21 languages available and three clicks or less to take you anywhere you need to be, EBSCONET is the most useful and intuitive information resource management system in the world.

Find out what the 16,000 customers using EBSCONET know already.

From acquisition to access, from management to evaluation – we can help.
What links some of the world-changing institutions of our time?