Metadata for Network Information Discovery and Retrieval: Workshop Report

Fifty experts representing the related fields of librarianship, computer science, online information services, publishing, and networking participated in a three-day workshop on the topic of Metadata on March 1-3, 1995.

WHAT ARE METADATA?

Simply, metadata are data about data. In the context of the networked environment, metadata refer to a set of data elements that one can use to describe and represent information objects. The workshop's primary goal was to develop a simple set of metadata elements that would be broadly applicable to document-like objects for network information discovery and retrieval. The result of the workshop was the Dublin Core Metadata Element Set, a preliminary set of data elements to describe and represent networked information resources or network objects.

Sponsored by OCLC and the National Center for Supercomputing Applications (NCSA), the workshop emphasized the critical need for solutions to assist users struggling with the increasing numbers of servers, documents, and a wide range of other network objects such as images, numeric databases, audio clips, and video archives accessible over the Internet. A recent report in Business Week (February 27, 1995, p. 78) indicates the power and impact of one Internet tool, the World Wide Web, on the extent of networked information: “There are now 27,000 WWW sites, and the population is doubling every 53 days, according to Sun Microsystems. And a research computer scientist at Carnegie Mellon estimates there are some five million documents stored on Web servers, a figure that is doubling every six months to a year.”

(continued on page 2)
NISO PROFILE:
Madeleine “Lennie” Stovel —
NISO Board Member Representing Libraries

Madeleine “Lennie” Stovel, NISO’s newest Board Member, representing libraries, has nearly 30 years experience in library science, 25 years experience in library automation, and a lot of travel that has taken her to Alaska, Antarctica, the Falkland Islands, and the South Georgia Islands.

Stovel first became interested in library science during a one-month summer job in college, when she was given responsibility for organizing a collection of foreign language books for the Ludington Public Library in Bryn Mawr, Pennsylvania. In addition to getting an overview of what goes on behind the scenes in a library, Stovel found that the job appealed to her natural inclinations to “put things in order and bring them under control.”

After receiving her undergraduate degree in biblical literature, history, and interpretation from Wellesley College in Massachusetts, Stovel moved to California and worked for the Palo Alto public library. Later, she obtained a master’s degree in library science at the University of California, Berkeley. She moved on to Information General Corporation, a company that used computer-based systems to do research on topics such as the cost of making machine readable records, and comparative studies of indexing and abstract services. There, she was able to apply her interest in data processing to projects that benefited libraries.

In 1970, Stovel took a job with Stanford University’s library, where she was responsible for designing Stanford’s first library automation system, the Bibliographic Automation of Large Library Operations Using a Time Sharing System (BALLOTS). In 1980, Stovel moved to Alaska, where she served as Librarian for the Alaska State Division of Energy and Power Development, and later as Assistant Librarian for Automation with the Anchorage Municipal Library.

In 1985, Stovel returned to the Bay Area to help develop the Research Libraries Information Network (RLIN) for the Research Libraries Group (RLG) in Mountain View, California. Stovel’s major professional accomplishments include managing the development of Eureka, an easy-to-use interface that allows anyone, regardless of technical ability or expertise, to easily search the RLIN database – an online catalog of more than 56 million items held by over 200 of the world’s leading research institutions.

Today, Stovel is RLG’s Manager of Applications Development. In that capacity, she manages projects pertaining to ANSI/NISO Z39.50 and other inter-system applications. She serves as RLG’s representative to the American Library Association’s Machine Readable Bibliographic Information (MARBI) committee in 1994, and was MARBI’s recorder from 1990-1991. She also reviews and comments on all U.S. MARC Advisory Committee proposals before they are forwarded to MARBI.

“Information standards play a critical role in the work I do, and NISO membership enables me to keep abreast of new developments,” Stovel told ISQ. Her goals as a NISO Board Member include assisting NISO’s efforts to expand membership.

“There are too many organizations that benefit from NISO standards but do not support their development,” Stovel observed.

When Stovel is not working, she enjoys traveling to exotic locations. This past year, she spent a month touring Antarctica, the Falkland Islands, and the South Georgia Islands, taking time to indulge in a favorite hobby — birding.

Metadata
(continued from page 1)

Discovering and retrieving information in the vast information space of the distributed, networked environment becomes increasingly difficult given the incredible growth in numbers of network-accessible resources. Although a variety of tools exist for browsing and accessing networked information, effective and efficient information discovery and retrieval remains relatively primitive. Part of the problem is the lack of bibliographic descriptions and standardized representations of the network objects.

While workshop participants may not have agreed
entirely on what constitutes metadata, they were
guided by a commonsense understanding that
metadata include a basic set of data elements that
can be used to represent or describe information
objects for the purposes of networked information
retrieval and discovery.

Metadata should provide enough information about
an information object to: 1) allow the object to be
discoverable, and 2) allow the object to be retrieved.
The metadata should allow a user to discover and
identify an object, and then to make some judgment
about the utility of particular information objects for
a user’s particular information needs.

The workshop organizers understood the inevita-
'bility of failure if they attempted to agree on metadata
for all the types of network-accessible information
objects, so they confined their attention to metadata
to describe and represent document-like objects
(DLOs). DLOs are analogs to currently existing forms
of information such as books, journals, articles, and
the like. And similar to their analogs, DLOs will often
include pictures, images, and other non-textual infor-
mation.

Metadata for these network objects would pro-
vide a minimal base of descriptive information. Work-
shop participants assumed that the process of creat-
ing the metadata would be decentralized and
distributed, parallel to the distributed nature of the
networked environment and the decentralized char-
acter of information production in this environment.
The metadata would be provided by the producers of
the information objects. Yet that metadata could
serve as a “chief source of information” for catalog-
ers or others to use, who could then enhance the
metadata by doing authority work, adding additional
controlled subject headings, and other professional
cataloging tasks that would add value to the descrip-
tion and representation.

13 CORE D.E.'s DEFINED

A result of the workshop was a preliminary set of
data elements comprising metadata for network
objects. Participants reached general consensus on
thirteen basic, core elements and recognized that
additional core data elements may be needed. Fur-
ther, the core data elements were assumed to be
textensible; subelements could be added for each of
the core elements.

SUBJECT: words or phrases indicative of the infor-
mation content. If the value comes from a controlled
vocabulary, the SCHEME subelement is used to
indicate which vocabulary.

TITLE: the title, name, or short description of the
object.

AUTHOR: the name or creator of the content.

PUBLISHER: the name of the entity responsible for
making the object available.

OTHERAGENT: the name of any other entity respon-
sible for the content of the object; the ROLE

subelement describes the type responsibility.

DATE: the date of publication. Specifically not of the
content but of the actual object described.

IDENTIFIER: a character string or number used to
distinguish this object from other objects; a SCHEME
subelement identifies the authority.

OBJECT-TYPE: conceptual description of the object.

FORM: physical, logical, or encoding characteristics.

RELATION: important known relationship to other
objects; the TYPE subelement describes the nature
of the relationship; the SCHEME subelement identi-
ﬁes the notation used to identify the related object(s).

LANGUAGE: natural language of the object content;
the SCHEME element identiﬁes the controlled vo-
cabulary.

EDITION: a free-text identiﬁcation of the version. See Relation.

SOURCE: object from which this object was derived;
contains a nested object description.

MODELS AND PROBLEMS

The diversity of workshop participants meant that
no one “model” of metadata existed, and the repre-
sentatives from the various communities brought
their perspectives on some of the problems that
metadata would have to address.

Priscilla Caplan, chair of ALA’s MARBI group, pro-
vided an overview of MARC. She emphasized that
creating MARC records implies not only using a spe-
cific record structure (i.e., ANSI/NISO Z39.2-1994) but
also the content designation specified in USMARC
documentation and the conventions (e.g., AACR2) for
determining the content of the records. The bibli-
ographic records in MARC provide metadata for the
bibliographic entities they describe and represent.

C.M. Sperberg-McQueen and Susan Hockey pre-
sented information about the Text Encoding Initia-
tive (TEI), a set of guidelines to assist creators of
electronic texts using Standard Generalized Markup
Language (SGML). The TEI Header is designed to
serve as the chief source of information for a cata-
logger, but it is also designed to elicit relevant infor-
mation from the creator of the electronic text. Filling
in the basic TEI Header information does not require
cataloging skills such as doing authority work.

Elizabeth Mangan from the Geography and Map
Division, Library of Congress, discussed the de-
velopment of the metadata standard for geospatial
data. In determining the data elements that would
comprise the metadata, Mangan listed four basic
criteria:

- Availability — data needed to determine what
  spatial data resources exist
- Fitness for Use — data needed to determine if a
  spatial data resource meets a specific need
- Access — data needed to acquire an identified
  data set
- Transfer — data needed to process and use a data
  set.
In June 1994, the Federal Geographic Data Committee approved the geospatial metadata standard. A major challenge to the library community has been dealing with multiple versions of bibliographic entities. Barbara Tillett, chief of the Cataloging Policy and Support Office at the Library of Congress, outlined the problem when there are different versions or manifestations of similar intellectual content. The issue is how the metadata can represent the various manifestations of a work and how to tie the physical manifestations to a single work. The problem of versions in the networked environment is more problematic given the dynamic nature of the digital objects. Electronic texts can be easily modified and trying to track such dynamic objects raises the question: To what extent is the current cataloging model's response to multiple versions useful in the networked environment? Tillett suggested that using concepts of equivalence and derivative relationships among and between the works and their physical manifestations would move the discussion from the "morass of terminology" of editions, versions, printings, and so on.

Lennie Stovel, manager of Applications Development at the Research Libraries Group, discussed the problem of collection hierarchy relationships. This problem concerns describing and representing component parts of a larger information entity. For example, an issue of a journal is a collection of the articles, editorials, book reviews, etc., that are contained in the journal. An article is a collection of pages, chapters/sections, paragraphs, sentences, and words. In the world of the Web, hyperlinks enable a reader to move easily between parts of the document. The question then becomes whether or not metadata are needed for all addressable parts of the document. The issue is to define or identify the level of granularity or detail at which metadata may be provided as well as showing the relationship among and between the components of the collection. These may be vertical, horizontal, and chronological relationships.

These and other presentations during plenary sessions gave the workshop participants an introduction to the needs, problems, and potential solutions of various stakeholder communities. Participants worked in several breakout sessions to determine a set of principles and assumptions that would guide the development of the metadata. More importantly, the workshop organizers charged the working groups to focus on the purposes the metadata would serve. The focus on purposes alleviated the need to agree on a common definition of metadata and instead pushed the development of an operational definition by identifying the uses of metadata — discovery, searching, retrieval, selection, description, interoperability, discrimination, etc.

**SUMMARY AND NEXT STEPS**
The preliminary Dublin Core Metadata Element Set is an initial, and possibly major, step forward. The workshop began the process to identify and build consensus on a standardized representation of network objects using standard data elements with well-defined semantics. Participants, however, recognized the need to refine the definitions, establish a mechanism for extensibility of each of the data elements (i.e., to accommodate new subelements as they are needed), define a controlled list of terms for some of the data elements (e.g., Object-Type), and develop a set of instructions for using the data elements.

A subgroup of the participants is now taking the next steps to prepare the Dublin Core Metadata Element Set for presentation in various forums such as the Internet Engineering Task Force working groups. Several of the participants are testing the usability of the metadata with samples of network objects. Others are preparing prototypes to search and retrieve metadata records. The metadata will evolve as concrete experience uncovers problems with the preliminary set of data elements.

We need only think about the evolution in modern cataloging techniques during the past 150 years — Panizzi's 91 rules (1841), Cutter's rules for a dictionary catalog (1876), various rules by ALA and LC from the turn of the century on, Paris Principles (1961), AACR2 (1967, 1988). It took over 150 years to expand and refine the library community's understanding of bibliographic description for the relatively stable bibliographic entities such as books. Even those bibliographic items still pose challenges to catalogers. The metadata workshop will not get everything right the first time, but it lays out an approach that has the potential to grow into solutions for networked information discovery and retrieval.

Those interested in more details on the metadata workshop can check out its web page: <http://www.oclc.org:5046/conferences/metadata/metadata.html>

Bill Moen
Syracuse University

---

**Implementers want to know**

**An ISO 12083 "Q and A"**

As more and more companies and organizations are implementing ISO 12083, an SGML implementation for books, serials, and journal articles, questions have come up about particular aspects of using the standard. Eric Van Herwijnen, the principal editor of ISO 12083 and the author of *Practical SGML* (Kluwer, 1994), offers a response to some of the most frequently asked questions. If you have a question, please submit it to NISO and we will forward it to Eric for reply.
Formatting tables continues to be problematic. Can you give me some pointers?

Concern about formatting tables is legitimate. The problem with tables is that coding them in SGML offers few advantages. Tables are similar to mathematics in that they both fall into that gray area somewhere between content and format. Compared to simple text, the informational value of a table is lost if the formatting is poor. The fact that an object with a 2 dimensional structure can be described by rows and cells with certain spannings between them does not give any information about the content of the data inside the cells, and the relationship between the data in neighboring cells. However, an advantage of using the SGML syntax for tables, even without meaningful semantics, is that it is a vendor-independent format. Provided that the DTD describes all visual properties of tables you can use it as a hub for converting into other (text processing) formats. The ISO 12083 table fragment is for simple tables only:

```xml
<!ENTITY % p.tbl "table" — table matter — >
<!ELEMENT table · · (no?, title?, tbody) · (%i.float;) >
<!ELEMENT tbody · O (head*, tsubhead*, row*) >
<!ELEMENT row · O (tstub?, cell*) >
<!ELEMENT tsubhead · O %m.ph; >
<!ELEMENT (tstublcell) · O %m.pseq; >
```

You can include another table fragment into ISO 12083 as follows:

```xml
<!DOCTYPE book [ 
<!ENTITY % p.tbl "mytable" — table matter element defined in xyz table fragment — >
<!ENTITY % xyz SYSTEM "path-name-of-xyz-table-dtd-fragment" > %xyz;
<!ENTITY bookdtd SYSTEM "path-name-of-iso12083-bookdtd" > %bookdtd; ]>
<book> etc.
```

This method of including or redefining elements is compatible with ISO 12083. Technically, you can add any table DTD fragment that you like and use it with ISO 12083.

At this time, the most practical solution is to support the three most common formats: ArborText, CALS and SoftQuad. You can probably make filters or have filters made between these DTDs. The editing environment for creating a table, i.e., having a WYSIWYG table editor, can also be a significant productivity issue. The ArborText and SoftQuad products were both designed to work with a proprietary WYSIWYG table editor.

The AAP Tables update committee has set out to define a vendor independent, lasting standard DTD for tables. I believe this work should be expedited — please contact me if you are interested in working on the standardization of table DTDs.

My final word of advice: an important criterion for deciding to describe data with SGML is re-usability. Use SGML whenever your data can be re-used.

---

WWW Opens ISO Meeting to World Audience

In the first known use of the Internet's World Wide Web (WWW) at an ISO standards meeting, daily minutes and technical issues of ISO TC184/SC4 were placed 'online' to allow standards experts from around the world to actively participate in the technical dialog. The ISO TC184/SC4 develops standards for representing and exchanging computer information describing physical products; it has a membership of over 400 experts from 22 countries. Standards Australia hosted the meeting from 19-24 March in Sydney and provided computing support and full Internet services under an arrangement with Dialix Services.

Materials from each daily session were placed on a Web server which was created for the meeting by AusSTEP Consultants in Perth, Western Australia. The Web pages allowed users to instantly link into a number of other Web servers in Germany, Japan, Slovenia, Switzerland and the United States. This enabled attendees to have immediate access to detailed SC4 Committee project information and related standards material.

Typical of the information available through the Web was:
- TC184/SC4 Subcommittee Home Page
- STEP Project Online Information Service
- PLIB Project Information Service
- Project Qualification Activities
- Product Data Implementation Activities and Project Schedules

For those who would like to browse the meeting materials, try the URLs:

Any questions or suggestions for improvement can be addressed to the Chairman of the SCR committee, Bradford Smith at: smithb@cme.nist.gov
NISO News and Notes

RECENT BALLOTING RESULTS

**Z39.14-199x Writing Abstracts**
(A revision of Z39.14-1979)
Ballot period: December 15, 1994 - March 15, 1995
Interim results:
- 14 Yes
- 8 Yes with comments
- 1 No
- 0 Abstain (OhioNet)

The draft has been referred back to the standards committee for revision. A second ballot will be conducted in 1996.

**Z39.18-199x Scientific and Technical Reports: Elements, Organization, and Design**
(A revision of Z39.18-1987)
Ballot period: October 6, 1994 - January 6, 1995
Interim results:
- 21 Yes
- 6 Yes with Comments (ALA, ASIS, LC, NIST, NLM, DoE)
- 0 No
- 1 Abstain (OhioNet)
- 5 Comments from interested parties

The standard is approved and will be published by NISO Press in 1995.

**Z39.22-1989 Proof Corrections - Reaffirmation Ballot**
Ballot period: September 1, 1994 - November 30, 1994
- 28 Reaffirm
- 1 Revise (AJL)
- 0 Abstain

If the negative vote is not withdrawn or resolved, a default ballot will be conducted in 1995.

**Z39.41-1990 Printed Information on Spines - Reaffirmation Ballot**
Ballot period: October 6, 1994 - January 6, 1995
- 25 Reaffirm
- 1 Revise (LC)
- 2 Abstain (NARA, OCLC)

The standard will be revised and released for a second ballot in 1995.

**Z39.49 Computerized Book Ordering - Withdrawal Ballot**
Ballot period: October 6, 1994 - January 6, 1995
- 28 Withdraw
- 1 Abstain (NARA)


**Z39.50-199x Information Retrieval: Application Service Definition and Protocol Specification.**

---

NCLIS Studies Internet Costs for Libraries

The U.S. National Commission on Libraries and Information Science (NCLIS) has funded a study to investigate cost models related to libraries and the Internet/National Information Infrastructure (NII). The study will explore the development of cost elements, categories, and models associated with public libraries connecting to and using the Internet/NII.

From February through April 1994, researchers for NCLIS will collect data on local public library and statewide initiatives to connect libraries to the Internet/NII. The cost information will be used to develop a preliminary list of cost elements, cost categories, and draft models based on the type of connection and levels of Internet services provided.

The Commission's study attempts to provide cost information based on institutional networking experience. Information resulting from this study is intended to assist libraries planning Internet involvement by identifying, describing, and categorizing the costs associated with Internet connectivity and services. The results of the study will address current policy issues related to the participation and roles of libraries in the provision of networked information.

This NCLIS study builds on a 1994 survey sponsored by the National Commission which resulted in publication of Public Libraries and the Internet: Study Results, Policy Issues and Recommendations. This report, providing baseline data on public libraries’ access to and use of the Internet, was done by Dr. Charles R. McClure, John Carlo Bertot, and Dr. Douglas L. Zweizig.

Dr. McClure, professor at the School of Information Studies at Syracuse University, and Syracuse doctoral students John Carlo Bertot and John Beachboard are the investigators for the new cost study.

The library Internet costing will investigate the following seven costing criteria: program planning and management; connectivity (including computer, telecommunication line, staff and maintenance costs); facilities upgrade and maintenance; navigation and access (including software and site license costs); training and education; content, and research, development and evaluation. Those public or state libraries that have collected cost data related to the Internet are encouraged to contact John Bertot (jcbertot@mailbox.syr.edu) or John Beachboard (jcbbeachb@mailbox.syr.edu) phone: 315-443-2911.

The National Commission plans to issue a report of the findings from this preliminary study in June 1995.

The U.S. National Commission on Libraries and Information Science is a fifteen-member, independent agency of the federal government charged by Public Law 91-345 to advise the President and Congress on national and international library and information policies and plans.
## Standards Status: April 1, 1995*

<table>
<thead>
<tr>
<th>Page</th>
<th>Status</th>
<th>Standard or Committee</th>
</tr>
</thead>
<tbody>
<tr>
<td>8,9</td>
<td>Ballot</td>
<td>Z39.4-199x Guidelines for Indexes and Related Information</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Retrieval Devices (A revision of Z39.4-1984)</td>
</tr>
<tr>
<td></td>
<td>Resolution</td>
<td>Z39.5-199x Abbreviations of Titles of Publications</td>
</tr>
<tr>
<td></td>
<td>Resolution</td>
<td>Z39.7-199x Library Statistics</td>
</tr>
<tr>
<td>6</td>
<td>Publication</td>
<td>Z39.18-1995 Scientific and Technical Reports</td>
</tr>
<tr>
<td></td>
<td>Revision</td>
<td>Z39.20-1995 Criteria for Price Indexes for Library Materials</td>
</tr>
<tr>
<td></td>
<td>Reconsideration</td>
<td>Z39.29-199x Bibliographic References</td>
</tr>
<tr>
<td>8</td>
<td>Ballot</td>
<td>Z39.32-1981 Information on Microfiche Headings</td>
</tr>
<tr>
<td></td>
<td>Revision</td>
<td>Z39.63-1969 Interlibrary Loan Data Elements</td>
</tr>
<tr>
<td></td>
<td>Reconsideration</td>
<td>Z39.69-199x Patron Record Data Elements</td>
</tr>
<tr>
<td></td>
<td>Resolution</td>
<td>Z39.70-199x Format for Circulation Transactions</td>
</tr>
<tr>
<td></td>
<td>Development</td>
<td>Z39.71-199x Holdings Statements for Bibliographic Items (SC AL)</td>
</tr>
<tr>
<td></td>
<td>Revision</td>
<td>Z39.72-199x Format for Submission of Data for Multimedia CD-ROM Mastering (SCAE)</td>
</tr>
<tr>
<td></td>
<td>Publication</td>
<td>Z39.73-1994 Single-Tier Steel Bracket Library Shelving</td>
</tr>
<tr>
<td>8</td>
<td>Ballot</td>
<td>Z39.74-199x Guides to Microform Sets (SC AC)</td>
</tr>
<tr>
<td>8</td>
<td>Ballot</td>
<td>Z39.76-199x Data Elements for Binding of Library Materials</td>
</tr>
<tr>
<td></td>
<td>Revision</td>
<td>Z65.1-1980 Permanent and Durable Library Catalog Cards</td>
</tr>
<tr>
<td>8</td>
<td>Resolution</td>
<td>NISO/ANSI/ISO 3166 Country Codes</td>
</tr>
<tr>
<td></td>
<td>Development</td>
<td>SCMM Env. Conditions for the Exhibition of Library... Materials</td>
</tr>
<tr>
<td></td>
<td>Formation</td>
<td>SCQQ Physical Preparation of Theses and Dissertations...</td>
</tr>
<tr>
<td></td>
<td>Development</td>
<td>SCSS Information to be Included in Ads [etc.] for Products Used for the Storage, Binding or Repair of Library Materials</td>
</tr>
<tr>
<td></td>
<td>Development</td>
<td>SCZZ Library Binding and Library Prebound Books</td>
</tr>
<tr>
<td></td>
<td>Development</td>
<td>SCAJ Format for Downloading Records from Bibliographic and Abstracting Indexing Databases</td>
</tr>
<tr>
<td></td>
<td>Development</td>
<td>SCAF Sorting of Alphanumeric Characters and Other Symbols</td>
</tr>
</tbody>
</table>

* Note: this list does not include current, approved standards not being revised.
Ballot period: September 1, 1994 - November 30, 1994
Final vote:
25 Yes
2 Abstain (AJL, NLM)
1 Comment from an interested party
The standard is approved; Version 3 of Z39.50 will be published by NISO Press in 1995.

NISO/ANSI/ISO 12083 Electronic Manuscript Preparation and Markup
Ballot period: March 1, 1994 - May 30, 1994
25 Yes
1 Yes with comments
2 Abstain
0 Comments from interested parties
The standard is approved and will be published by NISO Press in 1995.

NISO/ANSI/ISO 3166 Codes for the Representation of Names of Countries
U.S. Adoption of ISO 3166
Ballot period: June 1, 1994 - September 1, 1994
24 Yes
2 Yes with comments (ALA, ASIS, LC, NIST, DoEnergy)
1 No (AJL)
2 Abstain (RLG, Ohionet)
0 Comments from interested parties
It is anticipated that the negative vote can be resolved, if not a default ballot will be conducted in 1995.

1995 BALLOTING CALENDAR:

February 1995
Z39.74-199x New standard: Guides to Microform Sets
Z39.76-199x New standard: Data Elements for Binding of Library Materials
NWI Proposed New Work Item: Title Pages of Conference Proceedings

March 1995
Z39.57-1989 Holdings Statements for Non-Serials: 5 year Review/Reaffirmation ballot
Z39.64-1989 East Asian Character Code: 5 year Review/Reaffirmation ballot

April 1995
Z39.32-199x Revision of Z39.32-1981, Information on Microfiche Headers

May 1995
Z39.4-199x Revision of Z39.4-1984, Basic Criteria for Indexes
The draft is titled: Guidelines for Indexes and Related Information Retrieval Devices

NEW STANDARDS OUT FOR REVIEW


This standard is particularly useful to micro publishers and other microfiche producers, as well as users and collectors of fiche publications.

This standard promotes uniformity in the eye-legible identifying information to be included on the headers of microfiche intended for library use. These fiche publications may include books, journals, series, pamphlets, theses, report literature, public documents, and newspapers.

This standard specifies the location of elements in headers, the order of elements in each location, the minimum type size, and the degree of contrast between characters and background. The standard applies to computer-output microfiche of monographic or serial materials that meet the dimensional requirements of ANSI/AIIM MS5-1992.

Major changes in the draft revision include:
- New text that addresses the placement of information concerning restrictions on the use of the document, such as copyright or security classification.
- A new section that addresses the inclusion of both the collection title of the fiche and the title of work on the individual fiche. When a collection title is not provided, the user cannot easily identify the collection to which a fiche belongs. When the title of the work on the fiche is not provided, the user cannot identify the individual work or browse through the collection.
- A new section that specifies where publishers' company or product logo or identification should be placed.
- New text that elaborates on the need for the blank area in the identification area. Many publishers currently place their identification numbers, company or product logos or other identification in this area. This precludes libraries and information centers, from adding local identification information in that area when needed.

This draft was prepared by Harriet Rebuldela of the University of Colorado-Boulder. Prior to drafting the revised text Ms. Rebuldela and her colleague Tim Byrne studied the current production practices of twelve major producers of microfiche. Their study revealed that, overall, producers comply with the NISO standard.

The ballot period for Z39.32-199x is April 1, 1995 - July 7, 1995. Copies of the draft revised standard were distributed to all NISO Voting Members. The draft is available from NISO Press for $30.00 per copy plus shipping. Orders should be sent to: NISO Press, P.O. Box 338, Oxon Hill, MD 20750-0338. Telephone: 301-567-9522, Fax: 301-567-9553.
Indexes and Related Information Retrieval Devices (ANSI/NISO Z39.4-199x), will be distributed for a second ballot and review from May 1 - August 15, 1995. This standard is designed to keep pace with recent changes in publishing and information retrieval methods and technologies by identifying those generic criteria that apply to all types of indexes. It is a useful reference for anyone concerned with indexes used for information retrieval.

The standard emphasizes three processes essential for all indexes: comprehensive design, vocabulary management, and the provision of syntax. It deals with the principles of indexing, regardless of the type of material indexed, the indexing method used (intellectual analysis, machine algorithm, or both), the medium of the index, or the method of presentation for searching.

Many of the changes made to this draft were done in response to comments from Voting Members and others during the earlier review of the standard.

The draft revision includes a new section that explicitly addresses such issues as relations among indexers, authors and publishers; briefing of indexers; provision of documents to be indexed; time frames for compilation and delivery of indexes; changes after the delivery of an index; copy-editing and proofreading; remuneration; acknowledgment of indexers; and copyright. Many definitions have been modified, and a number of new terms have been added to the glossary.

The ballot period for Z39.4-199x is May 1, 1995 - July 7, 1995. Copies of Z39.4-199x are available from NISO Press for $40.00 per copy plus shipping. Orders should be sent to: NISO Press, P.O. Box 338, Oxon Hill, MD 20750-0338. Telephone: 301-567-9522, Fax: 301-567-9553.

STANDARDS MEETINGS AND EVENTS

Thesaurus Design Workshop to be Held May 12
A one-day workshop on thesaurus design sponsored by St. John's University, Division of Library and Information Science will be held Friday, May 12 in New York City. The Workshop will introduce the design of controlled vocabularies, covering concepts of formulation of descriptors, term relationships, thesaurus format, and screen displays. Computer-assisted techniques of thesaurus development will also be discussed.

The Workshop will be held at the Sixty East Club, 60 East 42nd Street, New York, NY (the Lincoln Building opposite Grand Central Station.)

The registration fee is $100. The registration fee includes lunch, coffee breaks and all handouts.

The workshop will be taught by Dr. Bella Hass Weinberg. Dr. Weinberg chaired the NISO standards committee which developed the most recent revision of the NISO American National Standard on thesaurus construction (ANSI/NISO Z39.19-1993). Dr. Weinberg is also a NISO voting member representative. She is a leading expert on thesaurus systems, indexing, and information management.

To register for the Workshop or for further information contact: James Benson, Division of Library and Information Science, St. John's University, Jamaica NY 11439. Telephone: 718-990-6200, Fax: 718-380-0353.

SLA Standards Update
Marjorie Hlava, Chair of the Special Libraries Association Standards Committee, will keynote a standards update to be featured during the SLA's 1995 annual meeting June 11-14 in Montreal. The standards session will be held Tuesday June 13, 1:30 - 3:30 p.m. Speakers include Pat Harris, NISO Executive Director, and Nolan Pope, Chair of the NISO Standards Development Committee. SLA attendees: this is your opportunity to stay on top of the standards scene!
ALA IRRT to Bring International Standards into Focus

The annual program meeting of the ALA International Relations Round Table will use the theme of international standards to focus on issues that affect libraries globally. Featured speakers will represent the ISO, standards implementors, and standards developers. Speakers include: Catherine Lupovici representing Jouve, a leading bibliographic systems vendor headquartered in Paris; Clifford Lynch, University of California, a leading proponent for the internationalization of Z39.50; Peggy Morrison, UNC-Chapel Hill Medical Library, the principal drafter of the proposed international standard for bibliographic electronic citations; Frances Schrotter, executive in charge of ANSI's international programs; and Pat Harris, NISO Executive Director and administrator of the U.S. TAG to ISO TC46. The meeting will be held Monday, June 26, 1995, 2:00-4:00 p.m.; check the conference program for the meeting room.

NISO WELCOMES THREE NEW VOTING MEMBERS

The College Center for Library Automation (Tallahassee, FL) joined NISO in March. Representing the CCLA are Dr. J. Richard Madaus (Voting Representative) and Ann Armbrister (Alternate VR). EBSCO Information Services has also joined NISO. EBSCO is represented by Sandra Hurd (Voting Representative) and Mary Beth Vanderpoorten (Alternate VR). Innovative Interfaces, Inc. joined as a Voting Member in January. The III is represented by Gerald M. Kline and Sandra Westall. By supporting NISO these companies and individuals are demonstrating their commitment to standards. Call the NISO office (301-654-2512) for details on how you can become a Voting Member and be a part of the standards solution.

SICI Revision Update

The NISO Standards committee to revise ANSI/NISO Z39.56-1991 has met three times within the last year, most recently on February 28. Between meetings much mail has flowed among the committee members, who have made solid progress on revising a complex standard that affects many operations across the globe and in diverse constituencies. These constituencies include librarians, publishers, agents, systems vendors, A&I services, the ISSN Network, a working group in the Internet community on URIs, the EDI community, providers of tables of contents and document delivery services, and users of bar code technology. Accordingly, clarification of the issues and negotiation of solutions has not been fast or easy. When the standard was published in 1991 it had been used only experimentally and in one context, for bar code serials checking. Now that use, while still significant, has been eclipsed by some other applications.

The progress on this standard falls into five categories. (1) Confirmation of the ISSN as a key data element for identifying the basic serial title, which was in some doubt at one point in the committee's work. (2) Some changes of rules to permit easy use by those who do not have at hand all of the desirable data elements required in Z39.56-1991. (3) The addition of three new data elements: a code structure identifier (CSI), a derivative part identifier (DPI), and a medium/format identifier (MFI). (4) Finally, the accommodation of parties such as Adonis, who wish to use a sequential or random (rather than algorithmic) identifier for the contribution. (5) Finally, it is the sense of the Committee that the algorithm for creating a contribution abbreviation should be expanded to seven characters from the current four as, currently it generates too many non-unique strings.

The history or details for these changes are explained at length in the minutes of the committee meetings. The minutes are available either on the mailserver for this standard (z39_56@faxon.com) or on the Faxon WWW server (http://www.faxon.com) in the standards section. The Faxon WWW server also has additional information about NISO. For instructions of joining the mailserver, please reach Fritz Schwartz at the Faxon Company (schwartz@faxon.com) or read them on the Faxon WWW server.

Although the discussion and negotiation of the changes between the 1991 and 1995 version of this standard took longer than anticipated, work is moving well now and the committee expects to be able to submit the standard to the NISO for distribution to voting members by July 1, 1995. Since great efforts have been made to include in the process those with strong opinions on this standard, the committee hopes to be able to resolve any negative votes reasonably quickly and to submit the final manuscript to NISO for publication by the close of 1995.

Fritz Schwartz
The Faxon Company
# NISO Standards:

<table>
<thead>
<tr>
<th>Standard</th>
<th>Description</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z39.2-1994</td>
<td>Information Interchange Format</td>
<td>$30</td>
</tr>
<tr>
<td>Z39.4-1984</td>
<td>Basic Criteria for Indexes</td>
<td>$30</td>
</tr>
<tr>
<td>Z39.5-1995</td>
<td>Abbreviation of Titles of Publications</td>
<td>$30</td>
</tr>
<tr>
<td>Z39.9-1992</td>
<td>International Standard Serial Numbering (ISSN)</td>
<td>$30</td>
</tr>
<tr>
<td>Z39.14-1979</td>
<td>Writing Abstracts</td>
<td>$35</td>
</tr>
<tr>
<td>Z39.19-1993</td>
<td>Guidelines for the Construction, Format, and Management of Monolingual Thesauri</td>
<td>$49</td>
</tr>
<tr>
<td>Z39.22-1989</td>
<td>Proof Corrections</td>
<td>$34</td>
</tr>
<tr>
<td>Z39.43-1995</td>
<td>Standard Address Number for the Publishing Industry</td>
<td>$30</td>
</tr>
<tr>
<td>Z39.44-1986</td>
<td>Serials Holdings Statements</td>
<td>$40</td>
</tr>
<tr>
<td>Z39.47-1993</td>
<td>Extended Latin Alphabet Coded Character Set for Bibliographic Use (ANSEL)</td>
<td>$40</td>
</tr>
<tr>
<td>Z39.48-1992</td>
<td>Elements, Organization and Design</td>
<td>$45</td>
</tr>
<tr>
<td>Z39.53-1994</td>
<td>Codes for the Representation of Languages for Information Interchange</td>
<td>$35</td>
</tr>
<tr>
<td>Z39.56-1991</td>
<td>Serial Item and Contribution Identifier</td>
<td>$49</td>
</tr>
<tr>
<td>Z39.57-1998</td>
<td>Holdings Statements for Non-Serial Items</td>
<td>$45</td>
</tr>
<tr>
<td>Z39.62-1993</td>
<td>Eye-Legible Information on Microfilm Leaders and Trailers and on Containers of Processed Microfilm on Open Reels</td>
<td>$45</td>
</tr>
<tr>
<td>Z39.63-1998</td>
<td>Interlibrary Loan Data Elements</td>
<td>$52</td>
</tr>
<tr>
<td>Z39.64-1989</td>
<td>East Asian Character Code for Bibliographic Use</td>
<td>$44</td>
</tr>
<tr>
<td>Z39.66-1993</td>
<td>Durable Hard-Cover Binding for Books</td>
<td>$35</td>
</tr>
<tr>
<td>Z39.67-1993</td>
<td>Computer Software Description</td>
<td>$45</td>
</tr>
<tr>
<td>Z39.73-1994</td>
<td>Single-tier Steel Bracket Library Shelving</td>
<td>$40</td>
</tr>
</tbody>
</table>

# NISO/ANSI/ISO Standards:

<table>
<thead>
<tr>
<th>Standard</th>
<th>Description</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>3166</td>
<td>Codes for the Representation of Names of Countries</td>
<td>$59</td>
</tr>
<tr>
<td>9660</td>
<td>Volume &amp; Fint Structure of CD-ROM for Information Exchange</td>
<td>$52</td>
</tr>
<tr>
<td>12083</td>
<td>Electronic Manuscript Preparation and Markup</td>
<td>$99</td>
</tr>
</tbody>
</table>

# Draft NISO Standards:

<table>
<thead>
<tr>
<th>Standard</th>
<th>Description</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z39.4-199X</td>
<td>Guidelines for Indexes and Related Information Retrieval Devices</td>
<td>$40</td>
</tr>
<tr>
<td>Z39.5-199X</td>
<td>Abbreviations of Titles of Publications</td>
<td>$30</td>
</tr>
<tr>
<td>Z39.14-199X</td>
<td>Writing Abstracts</td>
<td>$30</td>
</tr>
<tr>
<td>Z39.32-199X</td>
<td>Microfiche Headings</td>
<td>$30</td>
</tr>
<tr>
<td>Z39.63-199X</td>
<td>Interlibrary Loan Data Elements</td>
<td>$40</td>
</tr>
<tr>
<td>Z39.69-199X</td>
<td>Patron Record Data Elements</td>
<td>$30</td>
</tr>
<tr>
<td>Z39.70-199X</td>
<td>Format for Circulation Transactions</td>
<td>$30</td>
</tr>
<tr>
<td>Z39.71-199X</td>
<td>Holdings Statements for Bibliographic Items</td>
<td>$35</td>
</tr>
<tr>
<td>Z39.72-199X</td>
<td>Data Format for CD-ROM Mastering</td>
<td>$40</td>
</tr>
<tr>
<td>Z39.74-199X</td>
<td>Guides to Accompany Microform Sets</td>
<td>$30</td>
</tr>
<tr>
<td>Z39.75-199X</td>
<td>Alphabetical Arrangement of Letters and the Sorting of Numerals and Other Symbols</td>
<td>$35</td>
</tr>
<tr>
<td>Z39.76-199X</td>
<td>Data Elements for Binding of Library Materials</td>
<td>$40</td>
</tr>
</tbody>
</table>

# Standards Reference Tools:

<table>
<thead>
<tr>
<th>Tool</th>
<th>Description</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP2ISO</td>
<td></td>
<td>$150</td>
</tr>
<tr>
<td>Abbey pH Pen (Chlorophenol Red)</td>
<td></td>
<td>$5</td>
</tr>
<tr>
<td>The Annotated ISO 12083; Electronic Manuscript Preparation and Markup</td>
<td></td>
<td>$259</td>
</tr>
<tr>
<td>ASIS Thesaurus of Information Science and Librarianship</td>
<td></td>
<td>$35</td>
</tr>
<tr>
<td>Challenges in Indexing Electronic Text and Images</td>
<td></td>
<td>$30</td>
</tr>
<tr>
<td>Effects of Aging on Printing and Writing Papers: Proceedings of a Workshop held July 6-8, 1994</td>
<td></td>
<td>$99</td>
</tr>
<tr>
<td>Electronic Data Interchange; An overview of EDI standards for libraries</td>
<td></td>
<td>$25</td>
</tr>
<tr>
<td>Environmental Guidelines for the Storage of Paper Records</td>
<td></td>
<td>$35</td>
</tr>
<tr>
<td>From A to Z39.50: A Networking Primer</td>
<td></td>
<td>$35</td>
</tr>
<tr>
<td>The Government Information Locator Service (GILS)</td>
<td></td>
<td>$45</td>
</tr>
<tr>
<td>Implementation Guidelines for Electronic Data Interchange</td>
<td></td>
<td>$99</td>
</tr>
<tr>
<td>Information Standards Quarterly (ISQ)</td>
<td></td>
<td>$65/$85</td>
</tr>
<tr>
<td>Models for Open System Protocol Development</td>
<td></td>
<td>$25</td>
</tr>
<tr>
<td>North American Permanent Papers</td>
<td></td>
<td>$15</td>
</tr>
<tr>
<td>Practical SGML (Second Edition)</td>
<td></td>
<td>$65</td>
</tr>
<tr>
<td>Serial Item Identification: Bar Code Symbol Implementation Guidelines</td>
<td></td>
<td>$15</td>
</tr>
<tr>
<td>Scholarly Information and Standardization</td>
<td></td>
<td>$40</td>
</tr>
<tr>
<td>SGML TagWizard (Version 1.6)</td>
<td></td>
<td>$229</td>
</tr>
<tr>
<td>Standards: A Resource and Guide for Identification, Selection, and Acquisition</td>
<td></td>
<td>$45</td>
</tr>
<tr>
<td>TEI Guidelines (Text Encoding Initiative)</td>
<td></td>
<td>$79</td>
</tr>
<tr>
<td>The Unicode Standard, Version 1.1</td>
<td></td>
<td>Call</td>
</tr>
</tbody>
</table>

Prices effective 4/1/95. All prices subject to change without notice.
The 1995 NISO Press Catalog has been mailed to over 8000 standards followers in the US, Canada, and around the world. Your copy will arrive shortly.

The new catalog includes:
• 29 standards
• 12 draft standards
• 21 reference tools — publications that are designed to help you use and implement standards.
• three new pieces of software to support implementation of NISO/ANSI/ISO 12083
• the latest edition of the Unicode Standard, and
• the print and CD-ROM version of the TEI (Text Encoding Initiative) Guidelines — an SGML application for textual analysis.

The full list of NISO Press offerings is featured on page 11.

To place an order or request a copy of the catalog, contact NISO Press Fulfillment at: 1-800-282-NISO or Fax 301-567-9553. NISO Press now accepts MasterCard and Visa so ordering is easier than ever.

ISO Online Now Available

The ISO has announced the opening of ISO Online, an ISO Home Page accessible on the World Wide Web. The ISO Online address is:

http://www.iso.ch/welcome.html (in English)
http://www.iso.ch/welcomef.html (in French)

ISO Online will give you access to:
• The Catalogue of all ISO International Standards including drafts. A search tool is provided to find information by keywords or by ISO reference number. Information on how to order ISO standards and publications is given.
• Lists of ISO members and technical committees.
• General background on ISO, its structure, the scope of responsibility of each technical committee.
• Facts on the ISO 9000 Forum, the information service on ISO 9000, the series of International Standards for quality management.

Please send suggestions, questions, and comments about this service to webmaster@isocs.iso.ch.

Welcome to ISO Online!