

Selected Abstracts of Standards Under Review, 2006-2007 Notes from the ASIS&T Standards Committee (ASIS&T-SC)

by Marcia Zeng

ANSI/NISO Z39.85 - 2007 The Dublin Core Metadata Element Set

Related international standard: ISO 15836

Abstract: Defines 15 metadata elements for resource description in a cross-disciplinary information environment.

Status: Approved and published standard

ASIS&T-SC Note: The following major changes were made in this standard including:

- Changed element names from uppercase to lowercase in line with emerging Semantic Web practice;
- Revised definitions;
- Changed label from “subject and keywords” to “subject”
- Revised best practice guides and updated with newer versions of encoding schemes. For example, RFC 2396 (URI Syntax) → now RFC 3986; RFC 3066 (Language Tags) → now RFC 4646.

Please note that the current NISO version is different from ISO 15836:2003. Also, the web-based version (2006-December revised) on the DCMI website only reflects changes from the time when the specification was sent to NISO and therefore still includes obsolete content. All of these versions are still numbered as Version 1.1.

For example, the following comments are for “subject” element:

- ISO15836-2003: “Typically, Subject will be expressed as keywords, key phrases or classification codes that describe a topic of the resource.

Marcia Zeng served as chair of the ASIS&T Standards Committee 2007. She is also professor in the School of Library and Information Science at Kent State University and can be reached by email at mzeng@kent.edu

Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.”

- Current version (2006-12-18) on the DCMI website: “Typically, the topic will be represented using keywords, key phrases or classification codes. Recommended best practice is to use a controlled vocabulary.”
- Z39.85-2007, passed May 2007: “Typically, the subject will be represented using keywords, key phrases, or classification codes. Recommended best practice is to use a controlled vocabulary.”

NISO Z39.93 - 2007 The Standardized Usage Statistics Harvesting Initiative (SUSHI) Protocol

Abstract: Defines an automated request and response model for the harvesting of electronic resource usage data utilizing a Web services framework that can replace the user-mediated collection of usage data reports. The standard is built on the SOAP (simple object access protocol) for transferring request and response messages. The standard includes a versioned web services description language (WSDL) to describe the web service namespace and operations and an XML schema constraining the syntax of the SUSHI transaction.

Status: Approval pending

ANSI/NISO Z39.87 - 2006 Data Dictionary - Technical Metadata for Digital Still Images

Abstract: Defines a set of metadata elements for raster digital images to enable users to develop, exchange and interpret digital image files. The dictionary has been designed to facilitate interoperability between systems, services and software as well as to support the long-term management of and continuing access to digital image collections.

Status: Approved and published standard

NISO Z39.91 - 200x Collection Description Specification

Abstract: This draft standard defines a means of describing collections, where a *collection* is defined as an aggregation of terms. The draft standard takes the form of a Dublin Core application profile, a specification of how metadata terms from the Dublin Core metadata vocabularies and from other metadata vocabularies, some constructed for use in association with this Dublin Core association profile, are used to construct a description of a collection, in accordance with the DCMI abstract model. It also specifies an XML binding for serializing such descriptions for interchange between applications.

Status: Draft for trial use

ASIS&T-SC Note: This standard takes the form of a Dublin Core Collection Description Application Profile. The elements are associated with six namespaces: Dublin Core Metadata Element Set, v1.1; Dublin Core Terms; Dublin Core Type Vocabulary; MARC Relator Code Properties; Collection Description Terms (collection-specific terms); General Description Terms (non-collection-specific terms); and Collection Type Vocabulary Terms. A summary can be found at www.ukoln.ac.uk/metadata/dcmi/collection-ap-summary/

NISO Z39.92 - 200x Information Retrieval Service Description Specification

Abstract: This standard describes a method of describing information retrieval oriented electronic services, including but not limited to those services made available via the Z39.50, SRU/SRW and OAI protocols. The standard addresses the need for machine readable descriptions of services in order to enable automatic discovery of and interaction with previously unknown systems. It specifies an abstract model for service description and a binding to XML for interchange.

Status: Draft for trial use

ASIS&T-SC Note: SRU (search/retrieve via URL) protocol is developed based on the relevant aspects of Z39.50. It provides a means for expressing search requests, encoding them in URLs, sending them to servers and getting back results in a well-defined XML format. There is also a binding of the same protocol semantics over SOAP (simple object access protocol), wherein client requests and server responses are returned as SOAP messages. This binding was previously known as SRW (search-retrieve webservice) and now as “SRU over SOAP.”

Information regarding this change can be found at www.loc.gov/standards/sru/. The Open Archives Initiative Protocol for Metadata Harvesting (aka OAI protocol, OAI-PMH) provides an application-independent interoperability framework based on metadata harvesting. Version 2.0 specifications can be found at www.openarchives.org/OAI/openarchivesprotocol.html.

ISO STANDARDS

ASIS&T-SC Note: Only selected standards are included in this list. A great number of additional ISO standards went through systematic reviews (re-affirmation) and are not listed below. Also, some of the ballots were for the initiation of developing or revision of a standard. Voting members can nominate individuals to serve on the to-be-formed committees. For other standards, NISO has one vote that represents all NISO members' comments and suggested votes.

ISO/CD 28500, Information and documentation - The WARC (Web ARChive) File Format

Summary: The WARC (Web ARChive) file format offers a convention for concatenating multiple resource records (data objects), each consisting of a set of simple text headers and an arbitrary data block into one long file. The WARC format is an extension of the ARC file format [ARC] that has traditionally been used to store web crawls as sequences of content blocks harvested from the World Wide Web. The WARC format is expected to be a standard way to structure, manage and store billions of resources collected from the web and elsewhere. It will be used to build applications for harvesting (such as the open source Heritrix web crawler), managing, accessing and exchanging content.

ISO NWI, Internationalized (non-ascii script) code for the representation of names of countries and other significant geopolitical entities

Summary: Currently, applications and protocols use ASCII (A-Z, 0,9) but this creates a barrier for non-ASCII users, e.g., users writing in Japanese. A need has transpired for the internationalization of the codes within ISO 3166-1 (defines 2-3 digit alpha country codes). For example, many countries that do

not use Romanized scripts have no country code information printed on their passports. This standard will create codes in the scripts of the world's writing systems currently in use and harmonize the Romanized codes of ISO 3166-1 with their internationalized counterparts.

ISO/IEC CD 19773-1, Information technology – Metadata Registries (MDR) Modules Part 1: Overview

ASIS&T-SC Note: This standard is being issued by JTC1/SC32. Due to a liaison arrangement with TC46/SC11, we have the opportunity to comment on the draft.

ISO/NWI 25964, Information and documentation – Guidelines for the establishment and development of thesauri

ASIS&T-SC Note: This is a proposal to revise ISO 2788 and ISO 5964 regarding the development of monolingual and multilingual thesauri. The proposal document and attached Outline provided only extracts from BS 8723: *Structured Vocabularies for Information Retrieval – Guide*. The BS standard has five parts, four of which had been reviewed by invited reviewers. The equivalent standard in the United States is the NISO Z39.19-2005 *Guidelines for the Construction, Format and Management of Monolingual Controlled Vocabularies*.

ISO/WD 19005-2, Document management – Electronic document file format for long-term preservation – Part 2: Use of PDF 1.7 (PDF/A-2)

Summary: The new working draft is a second part to the PDF/A standard. Part 1 was based on version 1.4 of the PDF specification while this new part 2 is based on version 1.7 of PDF and addresses such added capabilities as file attachments, PDF packages, transparency and JPEG 2000 compression.

ISO/DIS 20775, Information and documentation – Schema for holdings information

Summary: The schema is designed to cover the holdings of all types of resources, physical and electronic. The schema is primarily designed to be included in responses to queries. This schema is intended as a more comprehensive and better defined alternative to the OPAC schema defined within ISO 23950 and the associated Z39.50 holdings schema that has not been widely implemented. This schema has been derived from the Z39.50 holdings schema, but updated to reduce the complexity. The schema is designed to be used as a fragment in other schemas or XML bibliographic resource descriptions such as MODS. It is being included as the structure for the “possible suppliers” element of the ISO OpenURL Request Transfer Message.

ISO/PDTS 23081-2 Information and documentation – Records Management Processes – Metadata for Records - Part 2: Conceptual and implementation issues

Summary: The purpose of this framework is to enable standardized description of records and critical contextual entities for records; provide common understanding of fixed points of aggregation to enable interoperability of records and information relevant to records, between organizational systems; and enable re-use and standardization of metadata for managing records over time, space and applications. It further identifies some of the critical decision points that need to be addressed and documented to enable implementation of metadata for managing records. It aims to identify the issues that need to be addressed in implementing metadata for managing records; identify and explain the various options for addressing the issues; and identify various paths for making decisions and choosing options in implementing metadata for managing records. ■