Digital Identifiers Roundtable Reports on Findings and Recommendations for NISO

Introduction

For several years, NISO continued to hear that confusion on how best to create, implement, and support identifiers and identifier systems in digital information systems was increasing the cost of developing and managing systems. In response, NISO convened a Digital Identifiers Roundtable with invited experts who were asked to share their insights, address questions relating to the needs for standards and functional frameworks for digital identifiers, and discuss the role that NISO can play in developing and promoting practices that will further the mission of its community.

The Digital Identifiers Roundtable was held on March 13-14, 2006 at the Lister Hill Center of the U.S. National Library of Medicine in Bethesda, Maryland. Forty experts representing libraries, government information centers, library system vendors, e-learning organizations, and content providers/aggregators listened to panel presentations and held group discussions on the topics of:

- Needs and Requirements – What Makes A Good Identifier?
- Identifiers Embedded In Systems
- The Services & Technical Infrastructure Associated with Digital Identifiers
- NISO's Role in Developing Standards for Identifiers

The final report from the Roundtable is reprinted here in its entirety. A downloadable PDF of the report can be obtained from:


What is an Identifier?

Identifiers for digital objects underpin information systems, and are essential for machine-to-human and machine-to-machine exchange of digital resources. Identifiers enable the packaging of digital information and the creation of services around such information, allowing its robust use in all sectors of human endeavor (including industry, commerce, academia, and government), affecting humans at scales from the individual to large organizations, and embodied in software that is highly visible (text processing, personal bibliographic databases) or nearly invisible (as is the case with much internal business processing). The reach of identifiers extends beyond traditional document management into scientific data sets, business...
processes, and commercial transactions that include a broad diversity of functional requirements and persistence expectations.

Identifiers exist to help manage information within a particular context or environment. This makes it difficult to develop universal or context-free answers. For any general question—“What are the ideal characteristics of an identifier?” “Should identifiers be resolvable?”—the answer is: “It depends.” It depends on the context within which the identifier is designed; it depends upon its purposes; it depends upon its uses within a particular community. Understanding a community’s specific needs is a necessary element for the development of a successful identifier system.

When identifiers proliferate, some items will have more than one identifier and some identifiers may appear to reference more than one item. Although this can be considered an error in some identifier systems, in other circumstances it is not. As an example, most individuals have many different identifiers associated with them: names, social security numbers, medical records numbers, and employee identification numbers. At times, multiple contexts will use the same number, as is the case with the social security number, which may be reused for a variety of identification purposes.

The uses of identifiers also proliferate. As identifiers age, the needs of their user community evolve, leading to “mission creep.” This is how the ISBN, designed to identify books in the publishing and retail supply chain, came to be associated with teddy bears and biscotti. This creates ambiguity around the identifier and its implementation. Mission creep may be preventable by rules and policies, but it may also be mitigated to some degree through identifier systems that are flexible enough to accommodate changing requirements. In any case, organizations that manage identifiers will have a role in directing the inevitable evolution of identifiers over time.

The Need for Identifier Standards

Private identifiers are useful for limited exchange where there are pre-existing agreements. However, the exchange of information across heterogeneous systems requires identifiers (and supporting systems) that are based upon public standards. Standards allow systems to “understand” identifiers that come from other systems, and prevent collisions between identifiers that are created independently. NISO has played a role in promulgating identifiers commonly used within the library and publishing information communities, independently and through its participation in the work of ISO TC 46 / SC 9 (Identification and documentation / Identification and description). With the emergence of the Internet and the World Wide Web, identifiers specified by the IETF (Internet Engineering Task Force) and W3C (World Wide Web Consortium), including URLs, email addresses, and format names, have come to play an important role both within and beyond these technical communities. Unfortunately, a coherent, generalized architecture for digital identifiers is lacking, leading to the emergence of new approaches that sometimes overlap and create confusion. The requirement that legacy (pre-Internet) identifiers must be accommodated in the digital world further complicates the management of globally unique, persistent identifiers.

Although many efforts are being conducted outside of the community of libraries and publishers, they will eventually influence the NISO community as it pursues its traditional role of generating, gathering, archiving, and disseminating information across all domains of human activity. The experience of NISO and its member bodies will help inform a broad interdisciplinary discussion of identifier systems and their requirements.

The Role of Identifier Systems

Some identifiers are tightly bundled with a system that manages the creation, resolution, and maintenance of identifiers. Such systems may supply a brand that a community comes to trust, although in discussions it was widely agreed that the syntactic form of an identifier can not itself convey trustworthiness. The degree to which bundling of identifiers, systems, and resolution services is desirable varies according to the application needs. Bundled identifier systems and resolution services provide complete, integrated solutions that may serve particular business needs. Many organizations, however, requiring autonomy in managing their own information resources, prefer either unbundled systems (e.g., URL-based) or partially bundled systems (such as URN, Handle, or DOI) in which complete responsibility for the final step in resolution rests not in the identifier system, but in the technical systems maintained by the organization.

Work on information systems tends to focus on technical problems, because they are generally better understood and easier to deal with than related socio-organizational problems. In the early days of the World Wide Web, identifier persistence was assumed by some to be a by-product of technology. It is now widely realized that persistence arises from a commitment of a maintenance body, and is not a technical property of an identifier.

The perceived value of an identifier rests in the user’s expectation that it will produce a desired result, and that efficacy arises from the reliability of the systems using that identifier. It is the role of organizations in the community to provide both the policy and technical fabric that supports and maintains identifier systems, and establishes the relationships of trust underlying them. At present, standards coordination is not done through a central authority, but rather through the collaboration of peer bodies (such as IETF, TC 46, NISO,
To have viable community solutions, we need a shared understanding of the variety of roles of identifier systems in the information space and how they can be managed sustainably. Although used every day, identifiers are a mystery to many people, including people responsible for building complex information systems. Business managers may not understand how identifiers can be used to provide services and therefore may not include identifiers in their plans at the service level. Many people repeat the same trial-by-error learning experience as they develop new identifiers. For example, people may create semantically-laden identifiers and then learn how difficult they are to maintain over time. Even technology staff may not understand the various aspects of identifier use or the long term implications of the creation of identifiers in their services.

It can also be difficult to discover that useful identifiers already exist. Opportunities are missed when service creators cannot discover applicable pre-existing identifiers and identifier systems. It can be difficult to find clear and complete information about an identifier and its associated systems. A minimal description should include:

- The intended use(s) of the identifier
- The syntactic rules governing the form of the identifier
- What the identifier is intended to resolve to
- The technical infrastructure that is available to support use of the identifier, and the parties operating it
- Policies governing creation, maintenance, support, and persistence of the identifier
- Information about any metadata related to the identifier is available
- A history of the identifier, including changes in any of the above over time

The overall goal is to facilitate appropriate re-use of identifiers and to permit the creation of services using existing, publicly available identifiers.

**Business Models**

There are costs associated with systems that generate and maintain identifiers. These encompass the creation of standards, the development and maintenance of registries and services, and marketing and education activities. In the library and academic environment, the cost of identifier systems has traditionally been built into the overhead of curatorial organizations. This has hidden the cost from users who have come to see identifier services as free.

An identifier scheme may itself establish a brand, as with a product or an institution. This branding can be part of the business model of the identifier system and therefore is important to its social and business standing. It can be functional branding, designed to help build adoption and to directly inform users of that particular identifier’s context and the way that it can be used.

Different business models can have different goals and values. In the academic community, a common business model is that of leveraging value through the sharing of resources. This business model favors open systems with little attempt to control re-use of resources. In the publishing community, the business model is focused on sales and licensing of resources. This model will favor systems that allow businesses to control access to resources as a way to protect revenue. Both types of systems will use identifiers, but the systems and services supporting those identifiers will have different characteristics, one of which may be that identifier systems developed in the business community may support different kinds of functionality than those in the academic community. There also will be needs in some arenas for the use of private identifiers that are not known outside of their originating system.

**Identifiers and the Web**

The most publicly visible current application of digital identifiers is the World Wide Web. The Web’s primary inventor, Berners-Lee, originally conceived of a family of identifiers, known as Uniform Resource Identifiers (URIs). Only two forms of URI have a notable presence in today’s Web. Far and away the most ubiquitous of these is the Uniform Resource Locator (URL). The Uniform Resource Name (URN) is a related identifier, conceived in the early 1990’s, and deliberated upon within the IETF. The URN is used, for example, to identify XML namespaces when a Microsoft Word™ document is saved as HTML. It has also been deployed within European libraries to retrieve documents through either special-purpose resolution systems, or by embedding a URN within a URL. As originally conceived, URN resolution would have been supported by the shared Internet infrastructure and invoked natively by Web browsers, but this has not come to pass. Roundtable participants expressed frustration concerning the historic difficulty of applying to the IETF to register new URN and URI schemes. Nonetheless, new URN and URI namespaces have appeared (e.g., the Handle URN and the “info” URI) and the URI registration process has recently been revised to make registration more straightforward.

The URL has become the de facto identifier of choice for the Web because of its ease of creation and use, the ubiquity of related tools, its flexibility in disclosing (or concealing) brands, and its ability to point to a location within a document using relative anchors. An identifier system that uses URLs today is immediately functional on a large scale as long as resolution is the expected...
behavior. The success of URLs results in part from the success of ubiquitously distributed Web technology that has become so important to so many communities that we can anticipate the appearance of smooth migration paths when changes are made to the Web infrastructure or its descendants.

In the Web’s nascence, URLs were concerned (as their name accurately reflected) with the location of a resource on the Internet. As such, they were less than perfect identifiers, as the objects to which they pointed often moved or disappeared altogether, leading to the familiar “404 Not Found” error returned to Web clients (“a broken link”, in the vernacular). This led to a widespread perception that URLs were inherently incapable of serving as persistent identifiers. However, it became apparent over time that the flexibility of URL semantics, combined with the Web’s forwarding (redirect) mechanism, allowed suitably committed organizations to provide potentially permanent identifiers in the form of URLs (as has been demonstrated by ARKs, DOIs, and PURLs, for example).

**Definition of Terms**

Discussions of identifiers often do not clearly distinguish between identifiers, identifier systems, and identifier resolution, leading to some confusion if the term “identifier” is used to describe different combinations of these three.

The concept **identifier** has been defined in differing ways. Defined by one Roundtable participant as “a relationship between a string and a resource,” it is more commonly defined as a symbolic stand-in for a digital object, most concretely represented in current practice in the form of a linear series of digital bits organized as discrete characters in accord with a defined protocol, a **string**. The former definition is helpful; however, in emphasizing that an identifier is defined operationally, it is its **association with an object** that makes it an identifier, and that association is most powerfully manifested by the ability to use the identifier to retrieve its associated object. Such retrieval, however, depends upon the existence of a system for performing the retrieval, which in turn assumes the existence of an organization that operates the system.

An identifier can still exist as such in the absence of such a system, as long as a trustworthy entity asserts the relationship between identifier and object. These definitions make it clear at the outset that identifiers can not be fully considered without thought being given to **relationships of trust**, and the organizations associated with them.

An IETF document (RFC 3404) implicitly describes identifier **resolution** to be a process by which an identifier string is employed to access its associated object and/or descriptive information about the object (metadata). This usually involves one or more intermediate mapping operations.

The term **service** is often understood differently depending upon one’s professional affiliation. For computer scientists, it might refer to the technical systems for creating and using identifiers, whereas for librarians and businesspeople, it would more likely refer to the actions they take on behalf of their clients (for example, an interlibrary loan service), which are in turn often built on top of identifier services, and involve human staff who are managing said services and interacting with human clients.

Several attributes of identifiers and their associated objects are worthy of definition.

**Resolvability** refers to the ability of an identifier to be resolved as described above. The terms **actionability** and **dereferenceability** are sometimes used in the same sense.

**Referent** refers to the object which is identified by the identifier, whether or not resolution returns that object.

**Granularity** refers to the extent to which a collection of information has been subdivided for purposes of identification and resolution (for example, at one extreme, an entire body of technical literature could be identified as a collection, and at the other, components of individual reports, such as tables and figures, could be identified for retrieval).

**Persistence** refers to the degree to which the resolvability of an identifier is matched to the business process that supports the association of identifier and referent. As such, it is an attribute of the association between an identifier and its referent rather than of the identifier itself.

**Semantic opacity** relates to the extent to which an identifier may itself carry information about an object (a fully opaque identifier carrying none).

These last terms will be enlarged upon in the following discussion of key identifier attributes.

**Key Attributes of Identifiers**

**Granularity:** When a book was offered in print form, a single identifier sufficed. With books available in digital form, individual chapters, illustrations or other components may serve as free-standing information units. Different levels of granularity within the same resource may be required to serve a variety of needs. When an identifier system does not provide such flexibility, parallel identifier systems may arise for the resource. The granularity “problem” is not one that has a single solution.

**Semantic opacity:** When parts of an identifier string may be inferred to be assertions about aspects of the object identified, such assertions can become misleading, infringing, or offensive due to semantic drift. A fully-opaque identifier does not enable such inferences. Such
an identifier can function over long periods of time during which organizations change names, new trademarks and acronyms arise, and subject hierarchies and language evolves.

Some identifiers (e.g., ISBNs) are probably opaque enough to be long-lived, as their semantics are widely recognizable only to information professionals. Semantically-laden (non-opaque) identifiers, while potentially perishable, can provide usability advantages by allowing one to select or verify them (e.g., from within a list) by drawing inferences about the related object. Such strings function as metadata containers, and their structure and semantics range from ad hoc and unpublished to fully standardized (e.g., OpenURL, SICI).

Some providers support or tolerate exploitation of recognizable identifier semantics, with one popular application being a reverse inference that permits a user to start with the content and correctly guess its identifier. One participant, however, described a system in which such user behavior was viewed as unwelcome “identifier hacking” leading to unintended access to collection resources that were not supported, a clear practical illustration of the importance of context in identifier functionality. Semantically-laden identifiers are difficult to maintain unless their semantic qualities are likely to remain unchanged (for example, the year at the beginning of an LCCN provides useful information that will not change with time).

**Persistence:** The degree required varies by application and is established in the context of a business case manifested as an organizational commitment to continued access. For example, identifiers used in tracking shipments require limited persistence. For information resources, it is often important to maintain the relationship between identifier string and object even if the resource is no longer available. In such instances, a resolution service may, appropriately, no longer resolve to the actual resource, but the string is still a *bona-fide* identifier if the service can provide information about the related resource (itself a kind of resolution).

It is often impossible to provide perfect solutions that address all of these identifier attributes, as the environment is somewhat chaotic and beyond the control of any organization or application. Lessons learned from existing systems can help information providers develop “good enough” solutions that largely meet user needs.

**Summary of Key Points**

Key points that emerged were:

1. Information exchange between systems requires identifiers that are based on public standards, both for shared use of the identifiers and to prevent collisions between identifiers that are developed in different contexts.

2. Identifiers are part of an infrastructure that includes support services for creating identifiers, for binding them to information or objects, and for resolving an identifier to obtain the associated object or metadata about the object.

3. Long term sustainability of identifiers requires community and institutional support backed by viable business models.

4. Identifiers must be usable within the standards of the World Wide Web in order to operate appropriately in the current networked environment.

5. There appears to be broad general agreement on the nature and properties of identifiers, with perceptions to the contrary attributable to past disagreements that can now be seen as having arisen from the differing intended uses of specific identifiers.

**Recommendations for NISO**

A goal of the Roundtable was to develop a list of specific issues where NISO could make a helpful difference. Numerous needs were expressed during the day and a half of the Roundtable. Some of them, such as the lack of universal resolution services for URNs, were deemed outside of the scope and capabilities of NISO and its community. Other issues were seen as appropriate NISO activities. The tasks that were identified as general areas for NISO to pursue are:

1. Establish a registry of identifier schemes that includes information about the associated services and policies for each scheme.

2. Explore use of the “info” URI registry as a focal point for community identifier needs.

3. Prepare a white paper on identifiers. This might include a glossary, a discussion of basic principles, services, and attributes for identifiers, and an implementer’s guide.

4. Educate decision-makers and technology developers in the community about identifiers and their uses.

Greater detail on each of these tasks, with specific goals, are described below.

**Education**

Those who create identifier systems, and those who use identifiers, need to learn from their colleagues across time and space. NISO can provide publications and courses to support services and promote interoperability between systems. NISO could prepare simple, clear documents along the lines of its “Understanding Metadata” publication, addressing those who use identifiers to provide services, system developers implementing identifier-based systems, and would-be creators of identifiers. A number of specific documents were suggested:

- A technical report (possibly written in collaboration with others) describing how URNs are intended to be
used, and how they relate to “info” URLs, addressing issues of location, resolution, and identification.

- A URN “profile”: a technical report outlining the URN specification as it pertains to identifiers of most interest to the NISO community.
- An “Implementor’s Guide” describing the practical aspects of selecting and using an identifier and its related technical services.

**Service & Policy Registry**

The identifier community needs a way to discover available identifier services and policies. Existing services could be used within new information systems, increasing efficiency and reducing the needless proliferation of competing approaches.

A registry will require a business model, as well as development and maintenance resources. A first step would be to create scenarios demonstrating the value of the registry, which could serve as a focus for community discussions of the functional requirements for such a registry. The scenarios would demonstrate discovery services related to otherwise non-actionable identifiers, and would be

- neutral as to business models;
- demonstrate cross-sector services; and,
- define the context dependence of the service.

The scenarios would demonstrate how services can be recombined to provide new functionality. The following scenarios were discussed at the organizers’ meeting at the end of the Roundtable:

1. A menu of services related to identifiers with the string “doi:”. Proxy URLs are a single redirect and DOIs and Handles are often buried in proxied URLs. In response to the string “doi:”, the registry would return a menu of services in structured XML; a Web browser plug-in would display the response. The identifier could be within a PDF file.

2. Resolution of ISBNs to services, in particular use of the OCLC x-isbn list for an ISBN.

3. Contextual resolution of music product identifiers in support of online services such as Napster.

The group also discussed the virtue of developing a vocabulary for standardized statements about services (policies); one application would be the formation of a controlled vocabulary of persistence promises along the lines of the permanence rating scheme devised by the National Library of Medicine. Entry within the registry might serve as a stepping-stone to elaboration as a formal standard for services and policies that demonstrated strong interest among users.

**“info” URI Registry**

NISO is the support agency for the “info” URI. Developed during the time that NISO Committee AX was working on the standard for the OpenURL, this new URI allows the creation of identifiers that the NISO community needs, and includes, importantly, the ability to turn legacy identifiers into URLs. NISO is in the process of formalizing the policies that will be used in “info” URI creation and maintenance, and in setting up the management procedures for the “info” URI registry. This identifier and its registry could serve as a focal point for NISO’s identifier activity, creating a trusted brand and a starting point for community members doing work that requires identifiers.

The members of the Roundtable planning committee were: Rick Rodgers (National Library of Medicine), Emily Fayen (MuseGlobal), Karen Coyle (consultant), Ted Koppel (Ex Libris), and John Kunze (California Digital Library). For the full list of Roundtable participants, visit http://www.niso.org/news/events_workshops/ID-06-wkshp.html

---

**MEMBER SPOTLIGHT**

**ARMA International: The Voice for Records Management Professionals**

*by Cynthia Hodgson, ISQ Editor*

ARMA International ([www.arma.org](http://www.arma.org)), a professional association established in 1955, is the oldest and largest international association dedicated to records and information management (RIM). “ARMA is the leading voice in the United States, and increasingly in the international arena, for the records and information profession,” states William Millican, Director of Professional Resources and Standards Development.

“We are the ‘go to’ organization for anything dealing with RIM.” ARMA provides a wide variety of resources to its over 11,000 members including an annual conference and expo, regional chapters, RIM promotional materials, self-assessment tools, an online forum for members to communicated, educational programs, and a variety of publications. Among the publications are guidelines and standards, the outcome of ARMA’s standards development program.

ARMA has been an ANSI accredited standards development organization since 1986, with a focus on the practical applications of day-to-day management of records and information in both paper and electronic
formats. Over 250 volunteers are involved in the standards program and there are currently 14 projects underway, with about 25% of them slated to become ANSI standards. “ARMA guidelines and standards are very similar in how they are developed,” Millican explains. “They differ, however, in the final review and approval stages. Guidelines are best practices that generally get a peer review while standards go through a formal public review and comment period. Standards, in addition to approval by ARMA’s Standards Development Committee, must also be approved by ANSI.”

Four programs are in their final stages of development and are expected to have published documents later this year. Two will be standards: Glossary of Records and Information Management Terms, a revision of ANSI/ARMA 10-1999; and The Records Conversion Process: Program Planning, Requirements, and Procedures, a new standard. Two will be published as guidelines: Records Management Responsibility in Litigation Support and Guideline for Physical Records Storage Facilities.

Recent corporate scandals involving records and financial statements, followed by legislation such as the Sarbanes-Oxley Act, have brought increased visibility as well as credibility to the records and information management profession. “RIM now gets a higher level of interest by much higher levels of management in many companies,” Millican affirms. “With new records-related regulations being proposed at all levels of government, organizations are seeking to implement best practices ahead of any potential legislation, especially where electronic records are involved.” ARMA is in the forefront of providing those best practices. Two of the most popular ARMA standards in the wake of the new legislation are: ANSI/ARMA 9-2004, Requirements for Managing Electronic Messages as Records, and ANSI/ARMA 8-2005, Retention Management for Records and Information.

As corporations go global, records and information management must also be addressed on a global basis. ARMA is an active participant in the International Organization for Standardization (ISO) committee on Archives and Records Management, a subcommittee (SC11) of ISO TC46 on Information and Documentation. In 2001, this committee issued the two-part standard, ISO 15489, Information and documentation – Records management, which has quickly established the baseline for excellence in records management throughout the world.

ARMA is also active in several other projects of the ISO Archives and Records Management SC including the recently issued Records management processes – Metadata for records (ISO 23081-1:2006), and the new project on Requirements for long-term preservation of electronic records. Kevin Joerling, Manager, RIM Content and Standards at ARMA and a member of the working group developing the preservation standard points out that “the required retention period for many electronic records exceeds the life of the technology and media used to create and store them. The standard we are developing will define the policies and procedures that organizations need to ensure their electronic records are still available and accessible throughout their required life.”

While ARMA continues its support for international standards, “we are not backing off from our national development activities,” Millican asserts. You can follow ARMA’s standards development activities on their website (http://www.arma.org/standards/development/). Individuals interested in participating in ARMA’s standards activities can contact Bill by email at wmillican@arma.org.

NISO NEWS AND NOTES

Three Information Industry Leaders Elected to NISO Board of Directors

The National Information Standards Organization (NISO) Board of Directors elections concluded with the addition of two new members, Michael Jon Jensen and Chuck Koscher, and the re-election of Oliver Pesch.

Michael Jon Jensen has been at the interface between digital technologies and scholarly/academic publishing since the late 1980s. In 2002, Jensen was appointed Director of Web Communications for the National Academies. He remains Director of Publishing Technologies at the National Academies Press, which makes more than 3600 books (over 600,000 pages) from the National Academy of Sciences, the National Academy of Engineering, the Institute of Medicine, and the National Research Council—fully browsable and searchable online for free (www.nap.edu). This site receives more than a million visitors per month, and boasts of some of the most advanced search and discovery tools available on any publisher’s site, most of which were initially developed by Jensen. In 2001, he received the National Academies’ “President’s Award,” its highest staff honor. Previously, Jensen was Electronic Publisher at the Johns Hopkins University Press, and Electronic Media Manager at the University of Nebraska Press. “The information ecosystem is still young and finding its way,” Jensen states. “It’s a fascinating transitional period of explosive evolutionary experimentation, and I’m delighted to be joining an organization committed to helping that ecosystem’s development into a robust, stable environment.”

As Director of Technology for CrossRef, Chuck Koscher has been actively involved for over four years in...
improving the linking infrastructure for their members’ scholarly publications. As a technologist, his focus is on implementation concerns driven by data management, data quality, and the development of new services and standards initiatives. Chuck has more than 17 years of experience in software services relating to content management and structured data delivery systems. On joining the Board, Koscher states, “I believe that NISO plays an important role in providing a forum for the airing and development of new ideas and for the codification of basic blueprints which ultimately lead to their realization. In my joining the board I hope to contribute to the bi-directional communications between publishers and other industry participants focused on standardization development and enhanced awareness of their benefits.”

Oliver Pesch is chief strategist for E-Resources at EBSCO Information Services, overseeing the technical direction for EBSCO’s Electronic Journals Service (EJS), A-to-Z, and LinkSource products. His standards committee memberships include the NISO committee on Library Statistics, Project COUNTER’s Executive Committee, and co-chair of the SUSHI working group. Pesch has been developing information products for EBSCO for nearly 20 years. In his current role he is very much involved with e-journals and the challenges they bring for activation, access and management. “The explosive growth of online information,” Pesch explains, “and the sheer number of products and services being developed to support this online environment represent both tremendous challenges and opportunities, particularly when it comes to information exchange. Collaboration and standards are key ingredients for maximizing our collective efforts and accelerating our progress in creating a better online environment for knowledge seekers. NISO is uniquely positioned to be at the center of these collaborative activities, and I am pleased and honored to be able to continue to be a part of this important work.”

Terms for the newly elected members run for three years.

**SUSHI the Topic of NISO’s First Webinars**

NISO held its first webinar, a two-part series, on the topic of the Standardized Usage Statistics Harvesting Initiative (SUSHI). The webinars, which drew hundreds of online attendees, were provided with the generous support of EBSCO, Ex Libris, Serials Solution, Swets, and Thomson Scientific.

The May 17 session, *Introduction to SUSHI for Librarians and Content Providers*, featured Adam Chandler of Cornell University, co-chair of NISO’s SUSHI committee; Phil Davis, also of Cornell; and Tim Jewell of the University of Washington. Chandler gave background on the issues that resulted in the formation of the initiative, focusing on the need for an automated method to move COUNTER reports into an Electronic Resource Management (ERM) system. Davis followed with a review of Project COUNTER and the COUNTER Code of Practice usage reports. Jewell wrapped up the session by describing the SUSHI project—its participants, goals, timeline, and current status.

**SUSHI: The Technology Unveiled** was presented on May 24 by Oliver Pesch of EBSCO, and Ted Fons from Innovative Interfaces. The second webinar focused on the technical aspects of the SUSHI protocol that is in development. Pesch described web services, the chosen architecture for SUSHI, and explained the components of the protocol: request message, response message, and COUNTER payload. Fons reviewed the roles of client and server in the use of the protocol, handling of security, and how the protocol could be used with an ERM system.

Additional tutorials on SUSHI will take place during the NISO September workshop on Managing Electronic Collections in Denver, and a third webinar aimed at implementers will be held later this year.

**NISO and Project COUNTER Cooperate on XML Schema Maintenance**

NISO and Project COUNTER (Counting Online Usage of Networked Electronic Resources) have signed a Memorandum of Understanding to promote harvesting of COUNTER reports via the web service protocol developed by NISO’s Standardized Usage Statistics Harvesting Initiative (SUSHI).

“SUSHI represents a major step forward in making it easier for librarians to assimilate, manage, and analyze the COUNTER usage data from vendors,” explains Peter Shepherd of COUNTER. “We are delighted to cooperate with NISO on this very important initiative.”

Under the agreement, NISO’s SUSHI program will assume responsibility for the maintenance of the XML Schema versions of the COUNTER Codes of Practice and host the schemas on the NISO website. Currently XML schemas exist for Release 1 and Release 2 of the Journals and Database Code of Practice. The Release 2 schema was developed by the SUSHI committee. The agreement ensures that the COUNTER XML schemas and the SUSHI protocol versions will be in synch.

Further information about SUSHI and the schemas can be found at: [http://www.niso.org/committees/SUSHI/SUSHI_comm.html](http://www.niso.org/committees/SUSHI/SUSHI_comm.html)

Recorded versions of both webinars are available from the SUSHI committee webpage: [http://www.niso.org/committees/SUSHI/SUSHI_comm.html](http://www.niso.org/committees/SUSHI/SUSHI_comm.html)
OCLC Appointed as OpenURL Registry Agency

OCLC (Online Computer Library Center) and NISO announced that OCLC will assume responsibilities as Maintenance Agency for The OpenURL Framework for Context-Sensitive Services (ANSI/NISO Z39.88-2004) for a period of five years. The standard defines the architecture for implementing a form of reference linking that includes extended services, is context sensitive, and open to access from multiple sources. A critical adjunct to the standard is the OpenURL Registry, which contains all the identification information for the components of an OpenURL ContextObject. The primary responsibility of the maintenance agency is to manage this OpenURL Registry including offering online web access, establishing procedures to update the registry, and providing marketing, advocacy, and education about the standard.

“OpenURL has significantly improved the world’s access to electronic journal content,” said Mike Teets, Vice President, OCLC Global Product Architecture. “It is now progressing to bringing the same access to a much broader set of services for electronic resources. As the use of OpenURL expands and more services are automated using this critical infrastructure, there is a growing need for a registry supporting the communication and extension of the current standard as well as the development of community profiles. OCLC has committed our reliable architectures to supporting the OpenURL community and its continued success.”

The OpenURL Registry is available at: http://openurl.info/registry/

NISO Registrations Renewed for ARK and SRU

The NISO Standards Development Committee has approved renewals of the registrations for the ARK Persistent Identifier Scheme and SRU Search/Retrieve via URL for an additional 18 months.

The ARK (Archival Resource Key) is a scheme intended to facilitate the persistent naming and retrieval of information objects. Developed at the US National Library of Medicine, it is currently maintained by the California Digital Library (CDL), which uses ARKs for all of the digital objects that it owns or controls.

SRU defines a protocol for search and retrieval, utilizing CQL (Contextual Query Language), a standard query syntax for representing queries. It includes a companion protocol: SRW (Search Retrieve Web Service). SRU, which builds on Z39.50, is used to search databases containing metadata and objects, both text and non-text.

The NISO Registration Process complements and extends the formal standard consensus process. By offering a lighter-weight review and accreditation, NISO supports the work of organizations and communities that are developing new practices in evolving information services and introduces the NISO community to emerging specifications of potential interest.

Information on all NISO Registrations is available from: http://www.niso.org/registration/registration_approved.html

NISO STANDARDS IN PRACTICE

LBI Seeks Comments on Updated Guide to the Library Binding Standard

The Library Binding Institute is making available an updated draft of the Guide to the Library Binding Standard for public review and comment. Originally published in 1990 as a companion to the American National Standard for Library Binding (ANSI/NISO/LBI Z39.78), the guide is being updated to bring it into harmony with the current version of the standard. Numerous technical challenges identified in the older edition have now been resolved or have industry best practices that were reflected in the 2000 edition of the standard.

The standard was written based on the assumption that readers understand the binding industry—its processes and machinery. The Guide to the Library Binding Standard was written to provide historical context, further explanation, and evaluative information about options. It parallels the numbering scheme of the standard to facilitate side-by-side reading. The guide recognizes that binding choices may vary depending on the needs for durability and openability.

The authors of the Guide are Jan Merrill-Oldham (Harvard University) and Paul Parisi (Acme Bookbinding). The initiators of the project were the members of the Physical Quality and Treatment Committee of the Preservation of Library and Materials Section (PLMS), Association for Library Collections and Technical Services, a division of the American Library Association. The Library Binding Institute provided support for the project.

Proposed Legislation Aims for Electronic Public Access to Research

In May 2006, the Senate introduced the Federal Research Public Access Act of 2006 (S.2695), a bill that would require federal agencies with research expenditures over $100 million to develop public access policies for their research manuscripts. Specifically, the bill states that the agencies must provide free public electronic access to papers resulting from federally funded research that are accepted for publication in peer-reviewed journals, no later than 6 months after publication in the journal. Published agency research findings are to maintained in a stable digital repository for long-term preservation. (The timeframe for long-term was not defined). The public archive would include the author’s final manuscript, not the publisher’s formatted, paginated version that is preferred for citation purposes.

A coalition of national library associations made up of the American Association of Law Libraries, the American Library Association, the Association of College and Research Libraries, the Association of Research Libraries, the Medical Library Association, and the Special Libraries Association, supported the Act.

The bill has been referred to the Subcommittee on Federal Financial Management, Government Information, and International Security, part of the Committee on Homeland Security and Governmental Affairs.

IDF Launches Website for MPEG-21 Rights Data Dictionary

The International DOI Foundation has launched a website for the MPEG-21 Rights Data Dictionary (ISO/IEC 21000-6). The MPEG-21 Multimedia Framework is a family of standards developed to define “the mechanisms and elements needed to support the multimedia delivery chain.” The Rights Data Dictionary (RDD) is the component of MPEG-21 developed to support the MPEG Rights Expression Language (REL) by providing “a structured collection of terms for describing usage rights that might be part of the business rules governing content consumption in the digital environment.” The International DOI Foundation was appointed as the Maintenance Agency for the RDD. Technical operations of the RDD are subcontracted to Rightscom Ltd.

The online RDD contains two parts:

- Terms that are defined in the base standard, ISO/IEC 21000-6
- Additional terms that are registered in accordance with the provisions of the standard

Although freely available, users must register to browse the dictionary. For each term, available information includes a unique identifier of the term in the form of a DOI, the definition, synonyms, relationships to other dictionary terms, special constraints, the element type, the meaning type, comments, and the authority for the term (the standard or the entity registering an additional term). Registration of new terms or term sets can be done using an online form.

The International DOI Foundation encourages anyone building or deploying Digital Rights Management (DRM) systems to use the RDD.

AIIM Spearheads Family of PDF Standards

The development of the international PDF Archival standard, ISO 19005-1:2005, Document management– Electronic document file format for long-term preservation – Part 1: Use of PDF 1.4 (PDF/A-1) was much anticipated in many user communities, including NISO’s. The project was initiated by AIIM, the Enterprise Content Management Association and NPES, the Association for Suppliers of Printing, Publishing and Converting Technologies and developed under the auspices of ISO Technical Committee 171.

Building on the success of PDF/A and the continuing need for standards for electronic document preservation and interchange, AIIM has launched a family of PDF-related standards projects. A second part of the ISO 19005 archival standard is underway. A new ISO standard on Engineering document format using PDF (referred to as PDF/E) will shortly be issued for review as ISO 24517; it is intended to improve the exchangability of engineering workflow documentation. An AIIM PDF/Healthcare committee is developing a standard for specifying PDF tags for healthcare records. A PDF/UA (Universal Accessibility) project is defining a file format to ensure that PDF documents are accessible to those with disabilities.

For information on joining any of the AIIM PDF standards committees, contact the AIIM Standards Program Director, Betsy Fanning (bfanning@aiim.org).
STANDARDS STATUS: JULY 1, 2006

In Development

Listed below are the NISO working groups that are currently developing new or revised standards, recommended practices, or reports. Refer to the NISO website (www.niso.org/committees/) for links to each group’s webpage, which contain member lists, minutes, and working documents.

DSFTU stands for Draft Standard for Trial Use.

<table>
<thead>
<tr>
<th>WORKING GROUP</th>
<th>STATUS</th>
</tr>
</thead>
</table>
  Z39.87-200X (AIIM 20-200X) standard approved by NISO members; awaiting AIIM approval |
| Digital Identifiers                                  | Chair: R. P. Channing Rodgers  
  Pre-standards research |
| Digital Rights Expression                            | Chair: Denise Troll Covey  
  Pre-standards research |
| Exchange of Serial Subscription Information          | Joint Working Project with EDItEUR  
  Co-Chairs: Priscilla Caplan, Richard Gedye  
  Field testing  
  Serial Release Notification (SRN), v 0.91  
  Serials Products and Subscriptions (SPS), v 0.91 |
| License Expression                                   | Joint project with DLF, EDItEUR, and PLS  
  Co-Chairs: Nathan Robertson, Alicia Wise  
  In development |
| Metasearch Initiative TG1, Access Management         | Chair: Mike Teets  
  Working with Shibboleth project to incorporate metasearch requirements into the next version of the Shibboleth specification |
| Metasearch Initiative TG2, Collection and Service Descriptions | Chair: Juha Hakala  
  Z39.91-200X, Collection Description Specification  
  Z39.92-200X, Information Retrieval Service Description Specification  
  Issued as DSFTU through October 31, 2006 |
| Metasearch Initiative TG3, Search / Retrieve         | Co-Chairs: Katherine Kott, Sara Randall  
  Field testing the NISO Metasearch XML Gateway (MXG) protocol |
| Networked Reference Services                         | Chair: Sally H. McCullum  
  Z39.90 – 200X, Question/Answer Transaction Protocol  
  DSFTU extended through December 31, 2006. |
| RFID                                                | Chair: Vinod Chachra  
  In development |
| Standardized Usage Statistics Harvesting Initiative  | Co-Chairs: Adam Chandler, Oliver Peach  
  In development |
| Versions of Journal Articles                         | Joint project with ALPSP  
  Chair: Cliff Morgan  
  In development |
| Web Services and Practices                           | Co-Chairs: Candy Zemon, Ian Davis  
  RP-2006-01, Best Practices for Designing Web Services in the Library Context being prepared for publication |

At Ballot

The following standards are currently being balloted for approval by NISO voting members. The draft standards are available for public review and comment on the NISO website (www.niso.org/standards/balloting.html).

<table>
<thead>
<tr>
<th>STANDARD</th>
<th>BALLOT PERIOD</th>
</tr>
</thead>
</table>
| ANSI/NISO Z39.43-1993(R2001), Standard Address Number (SAN) for the Publishing Industry | Five year review:  
  July 3, 2006 to August 31, 2006                                            |
Libraries, Data Providers, and SUSHI: the Standardized Usage Statistics Harvesting Initiative
by Adam Chandler and Tim Jewell. Against the Grain, 18(2), April 2006.
Provides background on the NISO Standardized Usage Statistics Harvesting Initiative (SUSHI) that is developing a protocol to automate the retrieval of COUNTER statistics. http://www.niso.org/Committees/SUSHI/ATGPre-printSUSHIv18-2.pdf

Metasearch Authentication and Access Management
Summarizes the work and final recommendations of the NISO Metasearch Initiative task group on Access Management. Evaluates 12 authentication methods for their suitability in a metasearch environment. http://www.dlib.org/dlib/june06/teets/06teets.html

Automated Rights Management Systems and Copyright Limitations and Exceptions

The Essential Role of Libraries Serving Persons who are Blind or Print Disabled
Discusses the role that many libraries for the blind are playing in creating accessible, specially formatted materials for their vision-impaired patrons. Reviews the NISO/DAISY Digital Talking Book standard and how it is being implemented. http://www.daisy.org/publications/docs/20060613001101/role_of_libraries.html

Institutional Strategies and Policies for Electronic Theses and Dissertations

Discusses the impact of copyright law and publisher licensing models on the ability of libraries to create and provide access to digitized content and particularly on the preservation of digital information. http://www.ippr.org.uk/ecomm/files/preservation_access_ip.pdf

Special Issue: Doctoral Papers from the European Conference on Digital Libraries (ECDL 2005)

Why OpenURL?
Provides a brief history of linking technologies and then focuses on the NISO OpenURL standard – how it works, its stakeholders, novel uses, and possible future developments. http://www.dlib.org/dlib/may06/apps/05apps.html

Calculating and Reporting Usage Statistics
Summarizes the work on COUNTER v4 to establish new, more accurate metrics for usage statistics reporting. http://www.dlib.org/dlib/may06/murray/05murray.html

The Impact of the Decline in Copyright Royalty Payments
Examines the decline in the amount of money paid by libraries to authors in exchange for the rights of use of their publications. http://www.library.pitt.edu/journal/journals/lnc/40/2-3/397-553.html

September 2006

September 12-13 NISO Board of Directors Meeting
Washington, DC

September 27-29 Managing Electronic Collections
a NISO Workshop
Denver, CO

November 2006

November 2-3, 2006 Discovery to Delivery
a NISO Workshop
Washington, DC

November 3, 2006 NISO Annual Meeting
(following the workshop)
Washington, DC