

February 2008

Standards Column -- The Problems of Institutional Identification: Toward a Universal Institutional ID

Todd Carpenter
NISO, tcarpenter@niso.org

Follow this and additional works at: <https://docs.lib.purdue.edu/atg>



Part of the [Library and Information Science Commons](#)

Recommended Citation

Carpenter, Todd (2008) "Standards Column -- The Problems of Institutional Identification: Toward a Universal Institutional ID,"
Against the Grain: Vol. 20: Iss. 1, Article 40.
DOI: <https://doi.org/10.7771/2380-176X.2722>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Standards Column — The Problems of Institutional Identification: Toward a Universal Institutional ID

by **Todd Carpenter** (Managing Director, NISO, 1 North Charles Street, Suite 1905, Baltimore, MD 21201; Phone: 301-654-2512; Fax: 410-685-5278) <tcarpenter@niso.org> www.niso.org

One of the common jokes relating to standards is that they are such wonderful things because there are so many to choose from. Particularly as they relate to institutions and information exchange, this is true. Unfortunately, for our community no one type of institutional identifier satisfies every need or use; an entity can have many different types of identifiers, each with a different purpose. Institutions, as multifaceted entities that conduct a myriad of interactions, have, as a result, ever-expanding lists of identifiers that relate to each particular institution. So one might ask, “Do we really need another?” That question — and its answer — is not so simple as it may seem, however, and this complexity forms the core rationale for the organization of a new working group within NISO.

Institutional identification is a complicated issue with many overlapping concerns, problems, and use cases. For every identifier, the core questions that need answering are: What are you trying to identify? What makes the entity being identified unique? And how will the identifier be used? By asking these questions, we can begin to evaluate the existing institutional identification standards. We can also discover gaps that tell us where a new identifier would be useful — indeed, necessary.

Organizations have a variety of business functions and needs, and there are a number of unique identifiers related to these. Most organizations in the U.S. have an **Employer Identification Number (EIN)**, also known as a **Federal Tax Identification Number**, which is assigned by the **IRS** and used to identify a business entity. Another type of business identifier is the **D-U-N-S (Data Universal Numbering System) Number**, developed and regulated by **Dun & Bradstreet** — a for-profit financial reporting firm. A **D-U-N-S Number** is a “unique nine-digit sequence recognized as the universal standard for identifying and keeping track of over 100 million businesses worldwide.”⁷¹ Both the **EIN** and **D-U-N-S** identifiers are essential at an organizational level for many business practices and needs. Obviously, paying or reporting taxes are critical business functions, as is organizational identification for auditing and contractual purposes. However, these identifiers all focus on the top-level needs of an organization. They lack the specificity and granularity that publishers and information providers need, particularly when information is provided to smaller divisions within an entity, rather than to the organization as a whole. There are also layers and relationships between entities that are needed in the business of providing information, but that are not captured by either the **EIN** or **D-U-N-S** identifiers.



What about those identifiers that have been created for use by the information community? For example, the **Standard Address Number (SAN)** is an identifier for publishers, standardized by the NISO community to identify the specific delivery addresses of an organization.² Managed by **R.R. Bowker**, the **SAN**, **American National Standard ANSI/NISO Z39.43-1993**, is a seven-digit numeric identifier used to signify a specific address of an organization in (or served by) the publishing industry (including book and serial manufacturers, libraries, publishers, etc.). Originally created to expedite paper-based transactions, such as purchase orders and returns, the **SAN** has been designated as the organizational identifier for use in **EDI** transactions in the publishing industry.³ It is the method used by **PUBNET**, **X*NET**, and **PubEasy.com** systems and is required in all electronic data interchange communications using the **Book Industry Systems Advisory Committee (BISAC) EDI** formats.⁴ However, the **SAN** only represents delivery addresses, and thus lacks the relational information critical to effectively manage digital delivery of content, such as site license subscriptions or consortia relationships.

Further, there are identifiers that either exist or are in development for identification of libraries and related organizations. The **International Standard Identifier for Libraries and Related Organizations (ISIL)** is a standard (ISO 15511-2002)⁵ within the portfolio of **ISO's Technical Committee 46 on Information and Documentation**. In the original development of the standard, the primary focus was to create an identifier to enable access to the holdings of library collections. The **ISIL** also supports granularity, in that library units within an organization may obtain their own identifiers. While the **ISIL** captures the identification of libraries, it is limited in that it does not extend to the diverse world that receives information from the publishing community, which includes not only libraries, but also other entities, such as book stores, jobbers, etc. Similarly, the **MARC** system, maintained

by the **Library of Congress** and the **OCLC** communities, includes unique identifiers that are assigned to libraries within their own structures and systems. However, these proprietary identifiers are unique to these communities. Like **ISIL**, they focus specifically on the library community and do not include the variety and scope of other organizations served by the publishing community.

Because of these gaps and limitations, then, **NISO's Business Information Topic Committee** has charged a new working group — approved by the **NISO Voting Membership**

in early January — to look at the institutional identifier environment and develop a standard that can be implemented in all library and publishing environments, and will meet the needs of the majority of participants. There are a number of potential areas that this proposed new institutional identifier could improve both for libraries and publishers, as well as agents and aggregators. The following potential use cases are drawn from the final draft report of the **Journal Supply Chain Efficiency Improvement Pilot (JSCEIP)**.⁶

Benefits of a potential unique institutional identifier for libraries:

- Precision in order processing and renewal continuity
- Faster access activation to electronic content
- Coherent reporting
- Central IP registration and servicing
- Institutional relationships cataloged
- Documentation of holdings

Benefits of a potential unique institutional identifier for publishers:

- Market penetration knowledge
- Usage analysis
- Customer identification
- End-user identification
- Automatic institutional holdings reports

NISO's goal in creating this identifier is to help remove some of the challenges associated with the information supply chain and to reduce the transaction costs of managing a diverse information flow. The need for such an identifier has become especially acute among the customer service and sales functions within publishers and for digital subscription distribution. Building on the work of the **JSCEIP**, which was limited to journals, the **NISO** working group will conduct an initial analysis of existing identifiers and will explore the possibilities of incorporating these identifiers into a unique institutional identifier that will broadly address the entire library and publishing supply chain for all types of media. The group will then explore the structure and associated metadata for such a new identifier, if their initial review determines that one is indeed needed. In addition, finding ways to incorporate and interact with other existing identifiers is another goal of the group. While a variety of potential business models and services might be developed using such an identifier, the **NISO** working group will focus only on the development of the identifier and its associated data, along with a management structure for correcting and maintaining the metadata.

A slate of nominated participants to join the new working group is being compiled, for

continued on page 83

Technology Left Behind — Let's Go Zotero

Column Editor: **Cris Ferguson** (Electronic Resources/Serials Librarian, James B. Duke Library, Furman University, 3300 Poinsett Highway, Greenville, SC 29613; Phone: 864-294-2713) <cris.ferguson@furman.edu>

The development of bibliographic management services has fundamentally changed the way many researchers keep track of their research resources and create bibliographies. Gone are the days of 3x5 note cards with facts and citation information scribbled on them. Instead, software programs and Websites designed to help users keep track of their research resources have replaced the traditional, and more tedious, processes for citing sources. Products such as **EndNote**, **RefWorks**, **ProCite**, and **Noodletools**, only a few examples of the citation management products available on the market, have become a fixture on the desktops of researchers at academic institutions of all shapes and sizes. While they differ in their approach to the task, these programs, which range from very simple bibliography generators to more high-level research tools, enable users to save citation information and generate bibliographies electronically.

Introducing Zotero

In October 2006, the **Center for History and New Media** at **George Mason University** released the beta version of **Zotero**, an alternative to commercial bibliographic management products like **EndNote** and **RefWorks**. Funded by the **United States Institute of Museum and Library Services**, the **Andrew W. Mellon Foundation**, and the **Alfred P. Sloan Foundation**, **Zotero** is a free, open source research tool designed to help users collect, manage, and cite their research sources. Since its release, **Zotero**, whose name is “loosely based on an Albanian word meaning ‘to acquire, to master’” (<http://gazette.gmu.edu/articles/9088/>), has grown into a robust tool for harnessing and organizing research resources.

Zotero is a plug-in for the **Firefox** browser. Users must first have **Firefox** installed on their computer to be able to download, install, and

use **Zotero**. The program cannot be used with **Internet Explorer** or other browsers. Once **Zotero** is installed, the **Zotero** icon appears in the lower right hand corner of the **Firefox** window. Opening the **Zotero** plug-in is as easy as clicking on the icon. While it is necessary to have **Firefox** open to be able access the **Zotero** window and the citation library, it is not necessary to be online to use the application.

One of the key factors that sets **Zotero** apart from other bibliographic management tools is this integration into the **Firefox** browser. (Similar bibliographic management tools are, by design, independent from the browser. **EndNote** must be installed on a user's computer, saving the citation information on the computer's hard drive, while **RefWorks** is a Web-based solution. Users navigate to the **RefWorks** Website to access their saved bibliographic citations.) When open, **Zotero** appears in the lower half of the **Firefox** window, enabling users to pull item(s) from a Website or from a list of search results directly into their citation library without ever leaving their browser. **Zotero** is able to remain open while users browse the Web, and items can be added to the citation library with the click of a mouse.

The Zotero Interface

The **Zotero** window has an easy-to-use iTunes-like interface, divided into three columns. The left column contains the “My Library” folder, which holds the complete library of items that have been saved. The full library can be sub-divided into individual collections. Each collection is represented by a folder, and the folders can be arranged hierarchically. The middle column displays all of the items in whichever collection (folder) is highlighted in the left column. An item could be a book, article, Webpage, or other information source, and a single item can be included

in multiple collections and folders. The right column is divided into five tabs: Info, Notes, Attachments, Tags, and Related. Each tab contains information about the item that is highlighted in the middle column.

The Info tab contains all of the citation information about an item, as well as the URL for the item, if it is available. On the Notes tab users have the ability to make notes about the selected item. These notes are searchable, and multiple notes can be added to a single item.

The Attachments tab enables users to attach files to an item. For example, a PDF image or the HTML version of an article could be attached to the article's record. In addition to capturing PDFs, **Zotero** can be used to archive Web content by capturing screenshots of Webpages. Users can highlight text and add annotated notes to the screenshot of the Webpage. This functionality means that **Zotero** can serve not only as a citation library, but also as a document library. By attaching images of articles and Webpages to the citation information in the **Zotero** library, users can use **Zotero** to keep track of and archive their actual research resources, not just the accompanying citation information.

The Tags tab enables users to categorize their items by inserting tags or subject headings. Tags can be browsed, searched, and edited globally. In some cases, when importing records from library catalogs or databases, **Zotero** will automatically pull the subject headings in with the citation information and pre-populate the Tags tab. The Related tab is used to relate one item in the **Zotero** library to another. For example, if a user has imported records for items on the same topic or by the same author, the items can be related to each other.



Standards Column from page 82

review and approval by the **Business Information Topic Committee**. Along with this core working group, a larger monitoring group will be formed that will allow interested individuals and organizations to be kept abreast of the working group's activities and to have the opportunity to provide input and feedback to the working group. If you or your organization is interested in participating in the monitoring group, please contact the **NISO** office.

The initial interest in this project has been significant and speaks to the complexity of the problems our community face managing the relationships among and between institutions that purchase content. While many important issues remain to be addressed, we have confidence that the developing working group will satisfactorily resolve them. In the end, we hope

there will be some added clarity surrounding much of the key identification data needed to provide information to institutions. 🌱

Endnotes

1. **Dun & Bradstreet Website** — http://www.dnb.com/US/duns_update/
2. **The SAN** — http://www.niso.org/standards/standard_detail.cfm?std_id=528
3. **SAN History** — <http://www.isbn.org/standards/home/isbn/us/san/san-history.asp>
4. **BISAC** — <http://www.bisg.org/bisac/>
5. **ISO 15511 - ISIL** — http://www.iso.org/iso/catalogue_detail?csnumber=27979
6. **Journal Supply Chain Efficiency Improvement Pilot Report** — **Chvatal, D.**, Editor. http://www.journalsupplychain.com/white_papers/JSCEI_Report_Final_Draft_20080111r.pdf

Importing and Exporting

Zotero is unique in its ability to pull citation information from Websites and import it directly into the **Zotero** citation library. Translators are created for various Websites that allow **Zotero** to “sense” or extract citation information and metadata elements on Websites. (A list of **Zotero**-compatible Websites can be found at: <http://www.zotero.org/index.php/translators/>). When **Zotero** detects items such as books or articles on a Webpage, an icon will appear on the far right side of the **Firefox** location bar, where the Web address or URL is displayed. To save the citation information about the item displayed on the Webpage in **Zotero**, it is as simple as clicking on the icon in the location bar. If more than one item is displayed on a Webpage, such as in search results from a library catalog or on a Website like **Amazon**, a list of items will be

continued on page 84