Against the Grain

Volume 16 | Issue 6 Article 33

2004

CHAOS -- The SICI Emerges, Cicada-Like, After Eight Years of Dormancy

Ted Koppel
The Library Corporation, tedk@tlcdelivers.com

Recommended Citation
Koppel, Ted (2004) "CHAOS -- The SICI Emerges, Cicada-Like, After Eight Years of Dormancy," *Against the Grain*: Vol. 16: Iss. 6, Article 33.
DOI: http://dx.doi.org/10.7771/2380-176X.4460

SHARED MEMORIES
by **Sandra Beehler** (Lewis & Clark College)

Anyone who's tried to identify family photos discovered in an old shoebox recognizes that data attached to such objects is important. Burgeoning numbers of digital photographs pose even bigger problems of identification — and some unique ways of providing metadata about the photos are emerging. Commercial firms can afford to hire a cataloger to provide that data; the rest of us may be able to take advantage of new products being developed to help with this task. Digital cameras already tag photos with date and time; global positioning systems may soon be built in. Data-mining, scene & facial recognition, and usage tracking are other tasks that a computer may soon be able to handle. But the key to photo management may lie in making the pictures public — thus taking advantage of the shared memories of all those who view them — one of whom just might be able to say "that's Uncle Seamus."

See — "Point. Shoot, Kiss it Good-Bye," ***Wired*** 12.10 (Oct 2004), p. 148-152.

GENETICS FOR ALL
by **Pamela M. Rose** (University at Buffalo)

Could terrorists use gene sequence data to engineer new bioweapons? A **National Research Council** report concludes that the possibility of bioterrorism should not stop scientists from freely sharing genome data, noting that the genomes of many dangerous pathogens are already in the public domain, and that it is unlikely that such data would help bioterrorists.

See — "Report Upholds Public Access to Genetic Codes" by **David Malakoff**, ***Science***, Sept. 17, 2004, p.1693.

ACCESS, AGAIN!
by **Pamela M. Rose** (University at Buffalo)

Lauded by librarians and opposed by commercial publishers, **NIH** has drafted a policy requiring grantees to deposit copies of their papers in **PubMed Central** once accepted by a journal. Manuscripts would be posted online six months after publication. The move is intended to increase public access to the results of **NIH**-funded research, but some scientific societies believe it could pose significant risk, and one society calls the plan "an unnecessary expenditure of federal funds for a redundant repository."

See — "NIH Proposes 6-Month Public Access to Papers" by **Jocelyn Kaiser**, ***Science***, Sept. 20, 2004, p.1548.

GDP GYRATIONS
by **Sandra Beehler** (Lewis & Clark College)

Economists are becoming alarmed about the effect of the falling dollar on the world economy. The enormous US trade deficit (currently 6% of GDP (gross domestic product)) must be brought down and a cheaper dollar is one way to accomplish this. Predictions are that the dollar will fall by 20% - 40% in the near future — the question is how fast it will go down and what effect it will have on the world economy. Unfortunately, economic conditions today parallel those in the mid-1970's when the dollar's collapse had severe consequences. Most US debt is held by foreigners, and a drop in the dollar's value reduces the value of their assets. Economist **Paul Volcker** thinks there is a 75% chance of a US currency crisis within five years.

See — "Economics Focus: Checking the Depth Gauge," ***The Economist***, Nov. 13, 2004, p. 84.

PECUNIARY PAIN
by **Sandra Beehler** (Lewis & Clark College)

Most economists agree that the US dollar must fall further in value in order to correct the enormous trade deficit. That the dollar has not fallen further and faster is due to Asian countries propping it up in order to avoid inflation in their own economies. Many factors: the inadequacy of personal savings in the US, the rising cost of oil, the trade deficit, and current US economic policies, not to mention the open-ended cost of waging the Iraqi war, indicate that a further fall in the dollar's value will be painful both to the US and to the world economy.

See — "The Wolf at the Door," ***The Economist***, Oct. 30, 2004. p. 75-76.

MUSIC TO MY POCKET
by **Sandra Beehler** (Lewis & Clark College)

Most agree that downloading of digitized music files is the wave of the future, but the largest music production companies have been holding back. In order to take advantage of downloading revenue they must rethink business models that have served them profitably for decades. Internet distribution also levels the playing field between large companies and independents. The drop in music sales is not all attributable to Internet file-sharing; other factors include a perceived lack of quality music, CD piracy, competition from other media (video) and less retail space devoted to music products. Major producers have pursued short-term profits at the expense of quality — digital music capability should re-open the market to independent and small-scale production. Meanwhile, no one is quite sure yet how to maximize profits from online sales and major producers are proceeding cautiously in that direction.

See — "Music's Brighter Future," ***The Economist***, Oct. 30, 2004, p. 71-73.

CHAOS — The SICI Emerges, Cicada-Like, After Eight Years of Dormancy

by **Ted Koppel** (Senior Product Manager, Standards Implementation, The Library Corporation) <tedk@tlcdelivers.com>

Column Editor: **Pat Harris** (Executive Director, National Information Standards Organization (NISO), 4733 Bethesda Avenue, Suite 300, Bethesda, MD 20814 USA; Phone: 301-654-2512; Fax: 301-654-1721) <pharris@niso.org> *http://www.niso.org*

Column Editor's Note: *Sometimes standards take time to come into their own. As* ***Ted Koppel*** *explains the* ***SICI*** *— the* ***Serial Item and Contribution Identifier*** *— is enjoying a renaissance, a reflection of the forward thinking that went into this standard's development. —* ***PH***

Eight years ago, a revised version of the **Z39.56** standard (***Serial Item and Contribution Identifier***, colloquially known as the **SICI**), was approved by **NISO's** voting members. First released in 1991 the **SICI** proved to be versatile, extensible, functional, and designed for interoperability and was immediately adopted by the then nascent document delivery industry (**CARL Corporation's** *UnCover* and **Faxon's**

continued on page 89

CHAOS
from page 88

Finder). A few publishers—primarily European with mostly low circulation scientific titles—began assigning a **SICI** to each issue, and an article level **SICI** to each article. Some publishers began placing the **SISAC (Serials Industry Systems Advisory Committee) SICI** barcode on printed issues (see the lower left corner of the front page of this periodical). Aside from these few low-profile applications, the **SICI** remained in the backwater of library standards. **Z39.56** became one on a list of many standard protocols to be "required" on most new library automation system RFPs, but very seldom did the requesting library know what technology they were requesting, or how they would use it.

What Has Changed? Why Now?

The **SICI** standards development committee did most of its work between 1993 and 1995, just at the beginning of the "Internet age." In the years since **SICI** was designed, new models of information manipulation, transfer, and use have proliferated. Digital libraries and repositories have been created and bibliographic (and citation) databases have become ubiquitous in research, academic, and library environments. As scholarly publication increased, so did the need of bibliographic and citation databases to uniquely identify each article for retrieval, digital rights management, and full text document retrieval. Current resource sharing and document delivery services are far less efficient when identifiers are less exact. The intervening decade has proven the need for the **SICI**, especially its article level unique identifier.

The **SICI** standard fits neatly into a number of IETF and W3C schemes dealing with **Uniform Resource Identifiers (URIs), Uniform Resource Names (URNs)**, and other Internet naming and classification hierarchies. The only deviation from current standard practice is that **SICIs** make use of special (punctuation) characters that in the intervening ten years have come to have a conflicting functional meaning in the HTTP protocol.

SICIs in Active Use

Database repository access is only as good as the ability to identify a specific, unique element within its collection. The following examples illustrate how databases and article repositories depend on the **SICI** to uniquely identify articles within the collection and provide linkages to the full text material.

- **JSTOR**, a not-for-profit organization whose mission is to archive important scholarly journals, held more than 2.5 million articles from 400 periodicals as of June 2004 and each of those 2.5 million articles needs to be directly accessible and digitally deliverable to **JSTOR's** subscribers. **JSTOR** uses the **SICI** contribution level identifier to create "direct URLs" that point to a specific digital document held in the JSTOR archive.
- **OCLC's FirstSearch** is an online service that gives access to several dozen databases. **SICIs** are included in full article citations in several databases, and can be used as pointers to the full text of the item. **FirstSearch** also makes available its "Direct Article Access Service," in which a library can create a **SICI**-embedded URL link from its own Web presence, directly to a specific full text article housed at **OCLC**. This capability is useful in electronic reserves or library-maintained reading lists and bibliographies.

Several protocols and standards utilize the **SICI**:

- **OpenURL** reference linking (**NISO Z39.88** draft) is a process by which a database source takes citation data and prepares a metadata message for transmission to a link resolver. The link resolver processes the incoming message and points to appropriate targets where the full text of the item can be found.

  The OpenURL standard defines several different roles for data in a metadata string. The referent (abbreviated in the OpenURL string as "rft.") is the desired item to which the OpenURL query refers. Among the referent identifiers allowed in an incoming OpenURL message are both **SICIs** and DOIs (which can contain meaningful **SICIs** within their prefixes). Upon receipt of an OpenURL metadata message containing a **SICI**, the OpenURL resolver decomposes (or

continued on page 90

parses) it into its meaningful parts and acts on each of them (ISSN, volume number, issue number, page, etc.) in order to resolve the citation and deliver the appropriate list of targets to the end user.

In addition to being able to process **SICIs** in incoming OpenURL strings, most Link Resolvers have the ability to create outgoing **SICIs** pointing to text delivery targets. An example of this would be a link resolver processing a citation from an article database which points to a full text article that is housed in JSTOR's archival repository of digital items.

- **The Dublin Core** metadata standard (**ANSI/NISO Z39.85**) is a set of fifteen data elements that can be used to describe a broad array of network resources. It is analogous to, but different from, the descriptive cataloging information contained in a cataloging record. Among the fifteen data elements is the "identifier" element, structurally labeled as "dc.identifier" and defined as "an unambiguous reference to the object within a given context." Identifiers can be ISBNs, ISSNs, DOIs, **SICIs**, etc.
- A very current potential role for **SICIs** that is being explored is the **Serials Release Notification (SRN)** format of the still-developing ONIX for Serials project. SRN, as envisioned, will support distributing data relating to the release of serial resources (issues, or articles within issues) by the publisher at about the same time the physical issues are dispatched. Potential uses for SRN are current awareness services, advance notice of publication for abstracting and indexing services, pre-assignment of DOIs by publishers, and more complete and accurate expected date and check-in information for automated library systems. For SRN to be effective, a unique, unambiguous identifier that can accurately describe a serials contribution (article) or issue is required. Although determination of the SRN unique identifier is very much still in the discussion and consideration stage, the **SICI** is an attractive candidate to fill this role because of its self-referential characteristics and its ease of resolution and parsing.

Conclusion

Eight years after its release as a national standard, the ***Serials Item and Contribution Identifier*** is finally coming into its own as a unique identifier for serials contributions. Advances in technology, Internet based products, and resource sharing prove the need for a mechanism that allows interoperability with digital databases and archives. With the opportunity for re-evaluation and improvement in 2005, the **SICI** can fill the role of a self-derivable, versatile, efficient unique identifier for another generation of serials.

Greater use and acceptance of the **SICI** implies a maturing understanding of the very important role that unique identifiers play in all areas of information transmission and delivery. The need by libraries, database suppliers, and similar industry entities for unique and unambiguous identifiers is already seen in the development of other standards, such as **NCIP**, the **Interlibrary Loan Protocol**, and in **NISO's Metasearching Initiative**. The information industry can build on the lessons learned from the renaissance of the **SICI** standard.

CHAOS SIDEBAR

The Guts of a SICI

SICIs are composed of three different segments. All three of these segments must appear in every SICI, although values within a segment may be omitted if not applicable.

Item Segment, which includes the data elements needed to describe the serial item. These elements include the ISSN, chronology, and enumeration of the piece being described.

Contribution Segment, which includes data elements that describe contribution or article information within the serial item. These include the title code (determined algorithmically) and page number. Depending on the code structure (see below), the contribution segment may additionally include another assigned number.

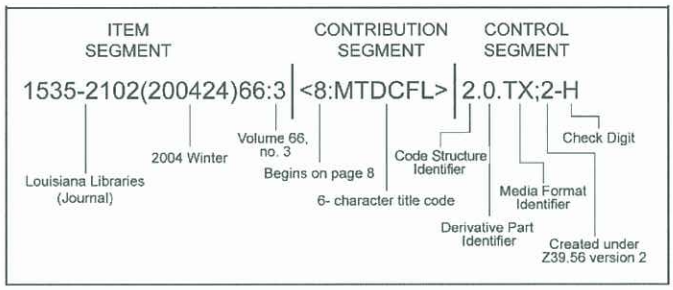
Control segment, which is administrative in function. It contains a version number, to identify the version of the standard being used to create the **SICI**, a **Code Structure Identifier (CSI)**, which designates whether the **SICI** refers to an item, a contribution, or some other alternate type of serial publication. Further, the control segment includes a **Derivative Part Identifier (DPI)**, which further defines the contribution as a table of contents, index, or abstract. The **Medium/Format Identifier (MFI)** distinguishes among the different possible formats of a serial—printed text, online, microfilm, etc. Finally, a calculated check digit, derived and calculated from the values of all the elements and segments that precede it in the **SICI** string, concludes the string.

SICIs can be used to identify either items (issues) or contributions within items (articles). Structurally, an item-level **SICI** and a contribution-level **SICI** are identical; the only difference is that data in the Contribution Segment of an item-level **SICI** is empty. The following examples illustrate different types of **SICIs** and their composition.

Contribution (article) level **SICI** example for:

Amsberryaugier, Lora and Hankel, Marilyn. Mining the Decennial Census for Louisiana Data, 1940-2000. Louisiana Libraries (volume 66, no. 3), Winter 2004

Note: The Code structure identifier is 2 and the Contribution segment is complete. It refers to a single article that begins on page 8.



Item Level **SICI** example for:

Information Week, July 26 2004 issue.

Note: This is an Item level SICI, because it has a code structure identifier of 1 and because the contribution segment is empty.