

ARTICLE  
EXCERPTED  
FROM:

# ISQ

## INFORMATION STANDARDS QUARTERLY

SPRING 2010 | VOL 22 | ISSUE 2 | ISSN 1041-0031

SPECIAL ISSUE: DIGITAL PRESERVATION

DIGITAL PRESERVATION  
METADATA STANDARDS

TRUSTWORTHY  
DIGITAL REPOSITORIES

UNIFIED DIGITAL  
FORMATS REGISTRY

AUDIO-VISUAL  
DIGITIZATION GUIDELINES

DIGITAL PRESERVATION  
EDUCATION



Jim LeBlanc



Adam Chandler  
E-Resources & Database  
Management Research  
Librarian, Cornell University

JIM LEBLANC

## Measuring the Quality of OpenURLs: An Interview with Adam Chandler

NISO's Business Information Topic Committee approved in December 2009 the establishment of a new working group called IOTA—Improving OpenURL Through Analytics. Chaired by Adam Chandler, E-Resources & Database Management Research Librarian in Central Library Operations at Cornell University, the working group will build on work previously conducted by Adam at Cornell. Jim LeBlanc, Director of Delivery & Metadata Management Services and Adam's colleague at Cornell, talked to him about the work he had already done and the follow-up project at NISO.

**Q** Let's start with something simple, Adam. What are OpenURLs?

Back in the 1990s, the only way to link from an article citation to a full text document was through something called bilateral linking. Each vendor needed to pre-compute and maintain all the links between their site's content and every other vendor site they linked out to. Then Herbert Van de Sompel and his colleagues at Ghent University came along and figured out a way to pass metadata to software that knows something about a library's collection, a method to exchange information to help a patron answer the question: does the library have access to this resource—print or electronic—and if so where is it? They essentially moved the job of maintaining the links to a brand new node in the supply chain, one optimized for the task: the "link resolver." Then they proposed a standard for the syntax of this "OpenURL" that would allow for predictable transfer of the resource's metadata.

The development of OpenURLs was hugely successful, because it addressed what was known as the "appropriate copy problem," a term that refers to the inadequacy of standard URLs to lead a user from the citation of an article to the most suitable full-text copy of that article. Commercial link resolver software was developed in the early 2000s to take an incoming OpenURL and: (1) determine if the library has a subscription to the journal in question, and (2) if so, present a new URL to the library patron that will connect him or her to full text—or to the library catalog or an interlibrary loan request form, if full text is not available. In 2004, the original OpenURL specification was generalized into a formal standard, ANSI/NISO Z39-88:2004, *The OpenURL Framework for Context-Sensitive Services*.

**Q** What's your specific interest in OpenURLs and quality metrics?

OpenURL was a genuine breakthrough and innovation for libraries. In 2009, Cornell patrons alone clicked on about half a million OpenURL citation links. In a talk last year, Herbert mentioned that a conservative estimate is that over a billion OpenURL requests are made by library patrons every year. The access these links provide can be very satisfying for library patrons, but bad links can be extraordinarily frustrating. Many vendors offer OpenURL links on their sites, but after the links go out to library link resolvers, the vendors have no idea what happens. They get no systematic feedback and don't know if library patrons are able to successfully access resources from their links. The aim of my project is to devise a method to provide feedback to vendors regarding the quality of the metadata content they're sending out, because the reality is OpenURLs don't work 100% of the time. Some OpenURL providers are better at supplying complete and accurate data than others. Nobody knows how often patrons are successful when they click on an OpenURL.

**Q** Where are you now in your research?

I've been gathering up usage log files from different link resolvers from three different institutions and three vendors. I have complete data for 2009 from Cornell, the Georgia Institute of Technology, and Kansas State University, plus sample data from EBSCO, Serials Solutions, and Thomson Reuters—a total of over 4,475,000 OpenURLs. I've written a program that parses each OpenURL, counts the elements that

CONTINUED »

are most likely to be needed for successful linkage (title, ISSN, author, date, and so forth), and indicates whether these elements are present or absent in the OpenURLs. Within each field of the OpenURL, I look for other things, such as whether dates have been entered in the correct form. The results are loaded into a database from which anyone can request reports.

**Q Can content providers request reports on the quality of their own OpenURL data?**

The web reporting system is currently organized by the institution or vendor who supplied the link resolver log file and date, but it is possible to generate an offline report for a vendor. For example, a year and a half ago Eric Rebillard, Professor of Classics and History at Cornell and editor of the bibliographic database *L'Année philologique*, was getting a number of complaints about failed OpenURL links. David Ruddy, Cornell University Library's Director of Electronic Publishing, and I worked with Eric to obtain a planning grant from Mellon to improve links from *L'Année*. The primary focus of the grant was experimental work on something called canonical citation linking. A secondary focus of the proposal was to develop an automated method for evaluating OpenURL quality. Eric is currently working with his programmers to fix the problems we identified when we ran the 900,000 plus OpenURLs through the parser. I recently ran a sample of OpenURLs for another vendor, the American Institute of Physics. I look forward to working with more vendors, as more of them find out about the NISO initiative.

**Q So it's necessary to keep the vendor-supplied data separate from other data in the database?**

I believe it is. The data from the 900,000 citations in *L'Année*, for example, would distort the results from other queries on the database. The point of the current system is to be able to pull data from library link resolvers for a specified time period (quarterly), because we want to monitor changes in quality over time. As vendors are sensitized to the issues and can see how their own OpenURLs compare in quality to those of their peers, they will, I hope, allocate resources to fix the problems that are uncovered. We will write a report on the efficacy of this model after two years and make a recommendation on its continuation. If vendors fix problems, we'll consider the work a success. If they ignore them, well, I might conclude that there is an inherent flaw in the OpenURL linking model that probably won't be fixed.

**Q What's next?**

The OpenURL standard has been around for ten years now, but this is the first attempt to create a feedback loop to help improve the quality of the data passed along in OpenURLs. A related issue is how to improve the proprietary and nonstandard inbound linking from the link resolver to the full-text content provider sites. I've been working on this problem for a while, thanks to our collaboration with Professor Rebillard, but the NISO initiative is helping me bring in other collaborators and solicit more interest in the issue. We have a great group of experts on board for the NISO project; members include Susan Marcin from Columbia University, Oliver Pesch from EBSCO, Ellen Rotenberg from Thomson Reuters, Elizabeth Winter from Georgia Tech, and Rafal Kasprowski from Rice. The existing OpenURL standard was developed under the aegis of NISO, so it makes sense to develop the quality metrics within the structure of NISO. Working through NISO will also keep the process transparent and impartial. We're also working closely with the joint NISO/UKSG KBART (Knowledge Base And Related Tools) Working Group that is developing recommended practices to improve OpenURL knowledge bases.

**Q What can people do who are interested in the project?**

We'd like to get more log files of OpenURL linking from those managing a link resolver, whether library or vendor. Anyone who has data they would be willing to share can contact me <email: alc28@cornell.edu>. An interest group e-mail list is available for anyone who wants to follow the activities of the working group. | NR | doi:10.3789/isqv22n2.2010.10

---

**JIM LEBLANC** <jdl8@cornell.edu> is Director of Delivery & Metadata Management Services, Cornell University Library.

**Canonical Citation Linking and OpenURL**  
cwkb.org

**Improving OpenURL Quality Through Analytics (IOTA) Working Group**  
www.niso.org/workrooms/openurlquality

**OpenURL Quality Metrics Database**  
openurlquality.niso.org

**OpenURL Quality Metrics Interest Group E-mail List**  
openurlqualityinfo-subscribe@list.niso.org



**RELEVANT  
LINKS**