KBART and the OpenURL: Increasing E-Resource Use Through Improved User Access

- **Introduction**
  - Todd Carpenter, Managing Director, NISO

- **KBART: Improving Access to Electronic Resources**
  - Peter McCracken, Co-founder and Director of Research, Serials Solutions, and co-chair, NISO/UKSG KBART Working Group

- **KBART: Benefits to Link Resolver Vendors**
  - Thomas Ventimiglia, East Asian Library Technical Support Specialist, Princeton University Library

- **KBART: A Librarian’s Perspective**
  - Christine Noonan, Hanford Technical Library, Information Services Department, Pacific Northwest National Laboratory

- **KBART: View from a Content Provider**
  - Jenny Walker, Vice President, Marketing, CredoReference

---

NISO 2009 Events

- May 13 (NISO/COUNTER Webinar): Usage Issues
- June 1 (Baltimore, MD) - Assessment & Performance Measurement
- June 10 (NISO Webinar): Interoperability Issues
- July 9-13 (Chicago, IL): NISO @ ALA Annual 2009

Please join us for the NISO/BISG Forum on Friday, July 10th!

- August 12 (NISO Webinar): E-books
- September 9 (NISO Webinar): Licensing Introductory Issues/SERU
- October 12-13 - Library Resource Management Systems (Boston, MA)
- October 14 (NISO Webinar): MARC/RDA/Bibliographic Control
- November 11 (NISO Webinar): Data Systems Population/Data Migration
- December 9 (NISO Webinar): ONIX-PL/ERMI Update

Register for three webinars, get one free!

---

OpenURL Evolution

- The evolution of the OpenURL in use:
OpenURL Evolution

- The evolution of the OpenURL in use:

If links fail, patrons will turn to the tool that always works.

Three main problems with OpenURL today:
- Bad data
- Bad formatting
- Lack of knowledge

The Measure of Success

- Better access for patrons
  - Fewer false positives and false negatives

- Best-case scenario:
  - IF a patron is seeking an article, and her library offers access to it through exactly seven online resources,
  - THEN the OpenURL resolver returns exactly seven accurate links

KBART: An Introduction

- Knowledge Bases And Related Tools
- UKSG and NISO collaborative project
- Get better data for everyone –
  - Those who provide data (publishers, aggregators)
  - Those who process data (link resolvers, ERMs, etc.)
  - Those who present data (libraries, consortia)
- All for THOSE WHO USE DATA – library patrons
- Ensuring timely transfer of accurate data to knowledgebases, ERMs, etc.
KBART: The Membership (Phase I)

- **Core working group** chaired by me (Serials Solutions) and Charlie Rapple (TBI Communications; formerly Ingenta)
  - Link resolver/ERM suppliers – Ex Libris, Openly/OCLC, Serials Solutions
  - Publishers – Taylor & Francis, SAGE, British Medical Journal
  - Subscription agents/aggregators – Swets, EBSCO, Ingenta, Credo
  - Consortia – SCELC, California Digital Library

- **Monitoring group**
  - More of these plus other related groups e.g. NASIG
  - Anyone can join monitoring group
    www.niso.org/lists/kbart_interest/

Solving the Negatives:
Lack of Knowledge

- Some content providers simply aren’t aware of what OpenURL does and why it benefits them
  - Education & advocacy
- Follow recommendations of “Link Resolvers and the Serials Supply Chain” report; provide useful information to those content providers
  - How to implement correctly
  - Offer contacts for those needing assistance

KBART: Focusing on the Negatives

- **“OpenURL’s Negatives”**
  - Inaccurate data leads to bad links
  - Incorrect implementation doesn’t transfer data properly
  - Lack of knowledge means some vendors aren’t using it

Solving the Negatives:
Incorrect Implementations

- Help content providers determine what is working, and what isn’t
  - Cornell project to focus on source OpenURLs
  - Identify correct and incorrect implementations
  - Give opportunity for vendors to grade selves
  - Offer more and better examples of OpenURL
- Standardize transfer of data within and among supply chain participants
Solving the Negatives: Inaccurate Data

- How do we handle incorrect data?
  - Grading? Policing? Shaming?
  - Biggest and most difficult problem to solve
- Highlight to content providers how important completely accurate data is to their end users
  - Consider the ‘false positive’: arrrgh, that’s frustrating...
  - Consider the ‘false negative’: much, much worse: how would you ever know?

KBART Phase I Deliverables

- Create a report that provides general guidance on problematic issues
  - Data problems
  - Incorrect implementation
  - Limited knowledge
- Offer best practices guidelines for how to effectively transfer accurate data among parties
- Provide better understanding of supply chain

Thanks!

- [http://www.uksg.org/kbart](http://www.uksg.org/kbart)
- [http://www.niso.org/workrooms/kbart](http://www.niso.org/workrooms/kbart)

- Peter McCracken (NISO co-chair)
  - peter@serialssolutions.com
  - Co-founder & Director for Research, Serials Solutions

- Charlie Rapple (UKSG co-chair)
  - charlie.rapple@tbicommunications.com
  - Head of Marketing Development, TBI Communications

KBART: Benefits to Link Resolver Vendors

Thomas Ventimiglia
Princeton University Library
NISO KBART Webinar
April 8, 2009
My Background

- 2004-2008: Part of the development team for the OCLC WorldCat Link Manager Knowledgebase
  - (Previously Openly Informatics’ “1cate”)
- Contacted content providers, compiled title lists, wrote in-house software to process metadata.

Challenges of Compiling a KB

- “Order from Chaos”
  - Over 100 providers, 2 million records each month
    - This data cannot realistically be entered by hand
    - Each record cannot be error-checked manually each month
  - Computer programs must process/validate this data and put it in a single, unified DB

- Each provider’s data is in a different format, with a different set of fields
- Data formats can change
- Large amount of work put into maintaining in-house software and writing new software when new providers are added
- KBART identifies a data format that is easy to process
  - Code can be reused
  - Less code needed

KBART reduces these challenges

- If KBART guidelines are followed, KB vendors can spend less time writing code to convert KB data into a unified format, and more time improving the quality and accuracy of their software and data
- Reduced likelihood of processing errors
- Benefits all parties – better search results, better likelihood of connecting patrons to resources
Chaos: Title List Formats

- Ideally, title list should be a table, so it can be imported directly into a database.
- In what format do we receive title lists?
  - Spreadsheet, XML, CSV file (good)
  - Web page, PDF brochure (not so good)
- In a web page, data is buried in HTML
  - If site design changes, processing code must change
  - Each title may have its own webpage, in which case many separate files must be downloaded and merged into a single file
- PDF: hard to extract tabular data

KBART: Title List Formats

- KBART recommends CSV (comma-separated-values) format
  - Common export format from DB and spreadsheet software
- Plain text: no colors, typefaces, etc.
- UTF-8 encoding: covers all languages
- As a result, almost no processing needed to put the data in a tabular format

Chaos: Data Fields

- Some title lists lack important data fields
  - ISSN: More stable identifier than a title
  - URL: Where to access the content
  - Coverage dates: Tells us what's accessible
- Attempts can be made to get this data from other sources (Web, Union Catalogs), but the process is messy, time-consuming, and often inaccurate.

KBART: Data Fields

- KBART recommends a set of metadata fields important to the basic functions of a link resolver
- Includes:
  - Title
  - Author (for books)
  - ISSN/ISBN
  - URL & Title Identifier
  - Coverage dates/volumes/issues
  - Embargo (coverage delay)
  - Content type (abstracts-only, fulltext, etc)
Chaos: Dates

- Content providers often use the date format that appears on the cover of the print edition
- Dizzying array of formats
  - Language variations: April vs. 四月
  - Numerical date variations (Is 1/4/09 Jan 4 or Apr 1?)
  - Many other designators (“Summer”, “1st Quarter”, “Pentecost”)
- These variations may all appear within the same title list

KBART: Dates

- Dates should be in YYYY-MM-DD format
  - YYYY-MM and YYYY fine for lower-frequency publications
- If a journal is covered up to the present, the coverage end fields may be left blank
  - KB will still be current when new issue is released

Other Recommendations

- Data should be updated each month
  - Or as often as it changes, if less frequent
- Data should be posted to an FTP site
  - Preferable to emailing between individuals
  - Can be harvested automatically
- Designated contact person on both provider and vendor ends

Role of Content Providers

- KBART will likely not be a burden to content providers
  - Many of these recommendations are already being practiced to some degree
- By making small adjustments to the way data is presented, providers will increase the chances that their online holdings are accurately conveyed to customers
Thank you!

My Background

- Electronic Resources Management Team Lead, Emerging Technologies Team Lead
- Experience with collection development, acquisitions, licensing, end-user and staff training and education, usability and website management, and research consultation
- NISO KBART Core Working Group Member
- LITA Education Committee Member

Today’s Environment

- 10+ years of increasing acquisition and use of electronic products in libraries
- 98% Collection Development budget
- The 3 Ps
  - Pricing
  - Packaging
  - Platform
- We need to:
  - Link resources in a meaningful way
  - Optimize the accessibility and use of resources
What does the Library do?

- OpenURL link resolver which is registered with variety of providers
- Customize and manage resources in the link resolver’s KB on an ongoing basis
  - Access limitations
  - Current subscriptions
  - Cancellations
  - Archival access
  - Consortia arrangements
  - Etc

What is a knowledge base?

- Database, “the brain”
- Contains information about web resources
  - e.g., what journals are accessible in ProQuest’s Research Library
  - How you would link to articles in them
- Identifies resources a library leases, owns, or activates
  - Not just journals!
  - May include print or other non-electronic resources
  - May include free, trial, or open access online resources

Why do we care about the knowledge base?

- Makes the connections
- Knows where the “stuff” is
- Directs traffic to library-specific resources
- Only place the user can get to the “appropriate copy”

“Hell is the place where nothing connects.” – T.S. Eliot

When linking goes wrong

- The elusive needle
- GIGO – Garbage in, garbage out
  - Wrong metadata
  - Poor linking
  - Darned ISSN and title changes
  - Link does not resolve to the right target
  - Dead end
- Why can’t you find the needle?
  - Outdated data!!!
Why KBART? Why now?

► UKSG research report – *Link Resolvers and the Serials Supply Chain*
► Highlighted many issues including
  ► Lack of awareness of OpenURL capabilities
  ► Impacts quality and timeliness of data provided to link resolver knowledge bases
  ► Undermines the potential of the sophisticated technology
► Unique partnership between US and UK entities
  ► Mission
  ► Guidelines
  ► Education
  ► Information

How will KBART impact libraries?

► Empowerment of library staff
► Education
  ► Tools and resources
  ► FAQs
  ► Presentations
► Easy to understand terminology
► Better understanding of the OpenURL supply chain
► Improved data accuracy and increased reliability of linking from, within, and between OpenURL compliant products
► Normalization of package names, date representation, title list formats, etc
► Resolution of errors
► Maximized usage of electronic products
► Improved user experience

Final Thoughts

► Online collections will continue to increase
► Increased metadata accuracy and transfer rates are crucial
► Standards and best practice guidelines
► Joint efforts further our ability to get everyone at the table to solve some problems!

► Thank you!
What is Credo Reference

- Online, full-text reference service
- Over 390 authoritative titles
- 3,000,000+ entries
- 60+ publishers
- Thoughtful integration

Integration with other library resources

- Metasearch via Z39.50 and XML Gateway
- Customisable links to local and 3rd party resources
- Inbound linking via persistent URLs
- OpenURL linking
  - Credo as Link Resolver Source
  - Credo as Link Resolver target

Credo as Link Resolver Target

- Increases traffic to Credo
- Many of the Credo reference resources are available direct from the publisher or from other aggregators
  - Need to link to the ‘appropriate’ copy
- Credo title lists available on Credo Web site (on request)
  - Link Resolver providers can download in order to update their knowledge bases
Some Issues to address to meet KBART guidelines

- Format of data (+ standard heading for the title list)
- General vs Specialist Reference titles
- Selective titles from Credo
Summary

- Credo welcomes the KBART initiative
  - Standardized procedures and formats
  - Trusted source of information
- Credo will make necessary changes to meet the KBART recommendations

Questions?

All questions will be posted with presenter answers on the NISO website following the webinar:


Thank you!
Jenny Walker