KNOWLEDGE BASES AND RELATED TOOLS:
A NISO/UKSG RECOMMENDED PRACTICE

Jason Price, PhD Claremont Colleges/SCELC
KBART Working Group Member
NASIG June 6 2010
Palm Springs, CA
Why you should you back this horse?

---

Jarrett Campbell from Cary, NC, USA
Vitals

– What – a universally acceptable holdings data format
– Where – throughout the supply chain & at the UKSG info hub http://www.uksg.org/kbart
– When – Now
  • Endorsement Phase -- June 2010
– Why – Better access for users through accurate holdings data
– How can you help?
  • For librarians
  • For publishers
Who is behind KBART?

• Standards organizations
  – UKSG and NISO

• The working group members, representing:
  – Knowledge Base vendors
    • ExLibris, Serials Solutions, Ebsco
  – Content aggregators
  – Publishers
  – Subscription Agents
  – Libraries & Consortia

• Full list -- [http://www.uksg.org/kbart/members](http://www.uksg.org/kbart/members)
What is KBART?

• “A set of practical recommendations for the timely exchange of accurate metadata between content providers and knowledge base developers”

• A universally acceptable holdings list format
  – Expresses title level coverage by date & volume/issue

• A single solution for sharing holdings data across the scholarly content supply chain

• A NISO recommended practice
A simple metadata exchange format...

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>publication_title</td>
<td>Publication title</td>
</tr>
<tr>
<td>print_identifier</td>
<td>Print-format identifier (i.e., ISSN, ISBN, etc.)</td>
</tr>
<tr>
<td>online_identifier</td>
<td>Online-format identifier (i.e., eISSN, eISBN, etc.)</td>
</tr>
<tr>
<td>date_first_issue_online</td>
<td>Date of first issue available online</td>
</tr>
<tr>
<td>num_first_vol_online</td>
<td>Number of first volume available online</td>
</tr>
<tr>
<td>num_first_issue_online</td>
<td>Number of first issue available online</td>
</tr>
<tr>
<td>date_last_issue_online</td>
<td>Date of last issue available online (or blank, if coverage is to present)</td>
</tr>
<tr>
<td>num_last_vol_online</td>
<td>Number of last volume available online (or blank, if coverage is to present)</td>
</tr>
<tr>
<td>num_last_issue_online</td>
<td>Number of last issue available online (or blank, if coverage is to present)</td>
</tr>
<tr>
<td>title_url</td>
<td>Title-level URL</td>
</tr>
<tr>
<td>first_author</td>
<td>First author (for monographs)</td>
</tr>
<tr>
<td>title_id</td>
<td>Title ID</td>
</tr>
<tr>
<td>embargo_info</td>
<td>Embargo information</td>
</tr>
<tr>
<td>coverage_depth</td>
<td>Coverage depth (e.g., abstracts or full text)</td>
</tr>
<tr>
<td>coverage_notes</td>
<td>Coverage notes</td>
</tr>
<tr>
<td>publisher_name</td>
<td>Publisher name (if not given in the file’s title)</td>
</tr>
</tbody>
</table>
Representing years of thinking...
Where does KBART apply?
Where...how it affects users

What’s in your knowledgebase?
What’s in your knowledgebase... the hard way

**Problem Type**

- Over-reported access dates: 31 titles
- Titles not listed: 85 titles
- Uncertain accuracy: 59 titles
- Title change not reported: 7 titles
- Under-reported access dates: 44 titles
- Match w/in a quarter: 1226 titles
Why KBART?

• Maintenance of accurate package content coverage data
  – Supports openURL Link Resolvers
  – Supports ejournal MARC record delivery services
  – Enables automated updating by KB providers

• Addresses common holding list inadequacies
  – Re-use of ISSNs
  – Embargo period ambiguities
  – Inconsistent date/enumeration formats
Desired impact on our work

- An end to our role as translators
  - No more badgering publishers to send complete access lists
    (List of necessary elements is standardized)
  - No more teasing out title changes to make the #’s match
    (Best practice is to include former titles & ISSNs)
  - No more waiting for the KB data team to translate data
    (Standardized format leads to automated ingest)
  - No more out-of-date access lists
    (Regular updates direct from publisher to knowledge base)
How Librarians can help

• Lobby publishers to adopt the KBART practices
  – LEARN about what KBART is and what it does
  – INSIST on ‘knowing’ what you are buying!
    • Require delivery of a usable holdings list *before* you pay
      – Initially & annually going forward
    • When the list is inadequate, point them to KBART
    • Only your insistence will make it happen
  – ENABLE publisher sales staff to make the case to their company
  – FOLLOW UP with continued requests as necessary
Applying KBART in practice

• Two case scenarios – AIP & ABP Journals
  – American Institute of Physics
    • self-initiated, complete, accurate
    • Driving expansion to other formats (conference proceedings)
  – A Big Publisher
    • Recognizes the problem
    • Need to establish priority of the change
    • Needs to get HOSTING SERVICE to program ability to produce
    • Takes persistence from dozens of customers with a mantra:
      – We need good data. Make it speak KBART!
AIP -- the ideal scenario

- First publisher KBART adopter
Providing KBART formatted data

<table>
<thead>
<tr>
<th>publication_title</th>
<th>print_identifier</th>
<th>online_identifier</th>
<th>date_first_issue_online</th>
<th>num_first_vol_online</th>
<th>date_last_issue_online</th>
<th>num_last_vol_online</th>
<th>title_url</th>
<th>first_author</th>
<th>title_id</th>
<th>embargo_info</th>
<th>coverage_depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal of International Accounting Research</td>
<td>1542-6297</td>
<td>1558-8025</td>
<td>2002</td>
<td>1</td>
<td>1</td>
<td><a href="http://link.aip.org/link/?JIAIAOP">http://link.aip.org/link/?JIAIAOP</a></td>
<td>JIARXX</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Irrigation and Drainage Engineering</td>
<td>0733-9437</td>
<td></td>
<td>1983</td>
<td>109</td>
<td>1</td>
<td><a href="http://link.aip.org/link/?JIDE">http://link.aip.org/link/?JIDE</a></td>
<td>JIDEDH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Legal Affairs and Dispute Resolution in Engineering and Construction</td>
<td>1943-4170</td>
<td></td>
<td>2009-02</td>
<td>1</td>
<td>1</td>
<td><a href="http://link.aip.org/link/?JLADH">http://link.aip.org/link/?JLADH</a></td>
<td>JLADAH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Management Accounting Research</td>
<td>1049-2127</td>
<td>1558-8033</td>
<td>2000</td>
<td>12</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Manufacturing Science and Engineering</td>
<td>1087-1357</td>
<td>1528-8935</td>
<td>1996-05</td>
<td>118</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Mathematical Physics</td>
<td>0022-2488</td>
<td>1089-7566</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Mechanical Design</td>
<td>1050-0472</td>
<td>1980-01</td>
<td>1982-10</td>
<td>104</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Mechanical Design</td>
<td>1050-0472</td>
<td>1050-0472</td>
<td>1990-03</td>
<td>112</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Mechanisms and Robotics</td>
<td>1942-4302</td>
<td>1942-4310</td>
<td>2009-02</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Mechanisms Transmissions and Automation in Design</td>
<td>1050-0472</td>
<td>1983-03</td>
<td>1989-12</td>
<td>111</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Medical Devices</td>
<td>1932-6181</td>
<td>1932-619X</td>
<td>2007-03</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Micro/Nanolithography, MEMS and MOEMS</td>
<td>1537-1646</td>
<td>2007-04</td>
<td>1</td>
<td>1</td>
<td>2004-10</td>
<td>3</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Microlithography, Microfabrication, and Microsystems</td>
<td>1537-1646</td>
<td>2002-04</td>
<td>1</td>
<td>1</td>
<td>2004-10</td>
<td>3</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Nanophotonics</td>
<td>1934-2608</td>
<td>2007</td>
<td>2007</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Nanotechnology in Engineering and Medicine</td>
<td>1949-2944</td>
<td>1949-2952</td>
<td>2009</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Performance of Constructed</td>
<td>0887-3828</td>
<td>1987-02</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
What do publishers need to do to adopt the KBART best practices?

2. Format ejournal and ebook content availability data to meet the requirements.
3. Self check your datasheet(s) on the KBART website to ensure that they conform to the recommended practice and make any necessary corrections.
4. Ensure that you have a process in place for regular data exchange as outlined in section 5.2 of the KBART report.
5. Register your organization on the KBART registry website, providing a link to download the newly KBART formatted dataset(s) [http://bit.ly/kbartregistry](http://bit.ly/kbartregistry).
The Registry -- a contact and metadata content clearinghouse

Knowledgebase management before KBART

Proactive reconciliation of an ejournal package list

☐ General Process – library, consortium or KB vendor
  ☐ Request updated access list from publisher
  ☐ Sample publisher list for accuracy
  ☐ Translate publisher list to match KB list
    ■ Number of titles never matches
    ■ Perform ISSN match with MS Access
    ■ Watch for & integrate title changes, mergers, acquisitions and losses
    ■ Watch for publisher-reuse of ISSNs/title combinations
    ■ Identify date discrepancies manually (inconsistent formats)
  ☐ Decide when its ‘good enough’ and go live/distribute new list
  ☐ Lather, Rinse, Repeat
Our vision...after KBART

Phase 1 – Universally accepted standardized publisher metadata, regularly distributed AND available on demand

Phase 2 – Broad adoption, More content type coverage

Phase 3? – Consortia & Institution level holdings metadata distribution based on what’s actually accessible from a particular IP

(this last is more my dream)