OPENURL KNOWLEDGE BASES AND LINK RESOLVERS

A report on the current landscape

Marshall Breeding
Director for Innovative Technology and Research
Vanderbilt University Library
Founder and Publisher, Library Technology Guides
http://www.librarytechnology.org/
http://twitter.com/mbreeding

27 March 2012
National Library of Sweden Workshop
Agenda

- Summarize findings and analysis of the current state of the OpenURL Link Resolver and Knowledge Base arena
General Observations

- Both commercially produced and community-based projects underway
- Commercial products currently are more comprehensive with more sophisticated support mechanisms
- Commercial:
  - Serials Solutions: 360 Link / KnowledgeWorks
  - Ex Libris: SFX / SFX Global Knowledgebase
  - EBSCO: LinkSource /
  - OCLC: WorldCat knowledge base
Non-commercial Products and projects

- Simon Fraser University:
  - CUFTS / GODOT
- Colorado Alliance of Research Libraries
  - Gold Rush
- Kuali OLE Project
  - Global Open Knowledge Base (GOKb)
Related Initiatives

- **KBART**
  - More standardized approach for information providers to deliver representations of their content products to KB providers

- **Project Transfer**

- **JISC initiative for shared services for electronic resource management**

- **KnowledgeBase+**

- **Open Discovery Initiative**
  - Specifically related to index-based discovery
Future Directions
Transition away from Standalone Products

- Knowledge bases were originally developed for OpenURL link resolution
- Increasing emphasis on knowledgebase as part of the infrastructure of broader strategic platforms
- Ex Libris: SFX Global KnowledgeBase will become part of the Community Catalog of Alma
- Serials Solutions: KnowledgeWorks will expand to include print and other materials in support of Intota
- OCLC: WorldCat Link Manager already discontinued in favor of WorldShare License Manager, WorldShare Management Services
Emerging Automation Models

- New genre of Library Services Platforms replacing Integrated library Systems
- Support for multiple formats:
  - Print and physical media
  - Electronic: subscribed and owned
  - Digital: local and external collections of images, digital recorded sound, video
- Globally shared metadata models: knowledge bases that support appropriate formats
- Deployed through multi-tenant software as a service platforms
- Open API’s for interoperability and extensibility
Evolving Role of Knowledgebases

- Shift from multiple knowledge bases for different product categories to more consolidated approach
  - Support infrastructure for electronic resource management
  - Basis for reference linking in broader end-user access environment for scholarly content
  - Basis for smart or direct linking within discovery interfaces
  - Basis for scoping for index-based discovery services
All the vendors report that they will continue to offer stand-alone linking products and knowledge bases in the short term.
Product leverage

- Initial acquisition of link resolver and KB
- Subsequent investment in discovery
- Subsequent investment in new automation products
In the past, link resolvers and their knowledge bases were mostly independent decisions.

Today, knowledge bases and link resolvers are mostly tied to discovery strategies. Many libraries are shifting to new resolvers to accommodate changes that they have made in discovery products.

- Using a link resolver from another provider may result in the need to operate multiple knowledge bases. Difficult to synchronize.

In the next cycle, knowledge bases will also be associated with new library management products.
Decision Drivers

- Today and in previous years link resolvers were acquired on the basis of their functionality and knowledge base completeness.
- Increasingly the link resolver will come as components of broader systems.
- Differentiation among knowledge bases will narrow as data becomes available more universally and consistently through KBART.
Open Access Knowledge Bases?

- Previous efforts of open access knowledge bases have not been sustained
  - JAKE
- GODOT not competitive with current commercial offerings
- Open access knowledge bases may pose more substantial competition
  - Global Open Knowledge Base (Kuali OLE / JISC)
  - KnowledgeBase+ 
Commercial Products
## Comparison

<table>
<thead>
<tr>
<th>KnowledgeBase</th>
<th>SFX Global KnowledgeBase</th>
<th>KnowledgeWorks</th>
<th>LinkSource</th>
<th>WorldCat knowledge base</th>
</tr>
</thead>
<tbody>
<tr>
<td>Company</td>
<td>Ex Libris Group</td>
<td>Serials Solutions</td>
<td>EBSCO</td>
<td>OCLC</td>
</tr>
<tr>
<td>Link Resolver</td>
<td>SFX</td>
<td>360 Link</td>
<td>LinkSource</td>
<td>None (OCLC Link Manager withdrawn)</td>
</tr>
<tr>
<td>A-Z utility</td>
<td>SFX A-Z list</td>
<td>360 Core</td>
<td>EBSCO A-to-Z</td>
<td>WorldCat Local includes integrated A-Z listing capability</td>
</tr>
<tr>
<td>MARC Delivery Service</td>
<td>Marclt</td>
<td>360 MARC Update Service</td>
<td>MARC Updates service</td>
<td>Pilot project</td>
</tr>
<tr>
<td>Discovery Service</td>
<td>Primo</td>
<td>Summon</td>
<td>EBSCO Discovery Service</td>
<td>WorldCat Local</td>
</tr>
<tr>
<td>Electronic Resource management</td>
<td>Verde</td>
<td>360 Manager</td>
<td>ERM Essentials</td>
<td>WorldShare License Manager</td>
</tr>
<tr>
<td>Library Services Platform</td>
<td>Alma</td>
<td>Intota</td>
<td>None</td>
<td>WorldShare Management Services</td>
</tr>
</tbody>
</table>
## Comparison

<table>
<thead>
<tr>
<th></th>
<th>KnowledgeBase</th>
<th>SFX Global KnowledgeBase</th>
<th>Knowledge Works</th>
<th>LinkSource</th>
<th>WorldCat knowledge base</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Launched</strong></td>
<td>2000</td>
<td>2003</td>
<td>2003</td>
<td>2006</td>
<td></td>
</tr>
<tr>
<td><strong>Support FTE for KB</strong></td>
<td>29</td>
<td>12</td>
<td>29</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td><strong>APIs</strong></td>
<td>Full APIs to support other interfaces</td>
<td>None now, but in 2012-13 roadmap</td>
<td>None</td>
<td>Through WorldShare Platform</td>
<td></td>
</tr>
<tr>
<td><strong>Number of installations</strong></td>
<td>2350 in 62 countries</td>
<td>973</td>
<td>3000+ Recent licensing changes makes LinkSource available to all libraries using A-Z</td>
<td>301</td>
<td></td>
</tr>
</tbody>
</table>
Comparing Comprehensiveness and Quality

- Assessment of the scope of the products cannot be accomplished based on numerical comparisons.
- Numbers provided by each supplier not comparable.
- Process for how libraries localize their implementation for materials not covered differs.
- Quality and consistency as reported by supplier sometimes not entirely consistent with library perceptions.
Library Survey

- Conducted a survey completed by over 500 libraries using link resolvers
- Questions
  - How do you rate the overall performance of the Link Resolver and its Knowledge Base?
  - How do you rate the performance of the link resolver in its effectiveness of connecting users to appropriate copies of electronic resources?
  - How do you rate the comprehensiveness of the knowledge base relative to the number of e-journals and other electronic resources held by your library?
  - How do you rate the accuracy of the knowledge base in representing electronic resources?
  - How do you rate the promptness in which errors reported by libraries are corrected in the knowledge base?
  - How do you rate the tools offered for the library to configure and administer the knowledge base and link resolver?
Ex Libris SFX & Global KnowledgeBase

- First to acquire and commercialize OpenURL link resolvers (Feb 2000)
- Promoted OpenURL 0.1 as a de facto standard and OpenURL 1.0 as a NISO Standard
- SFX Global KnowledgeBase
- Most Installations worldwide
- Many new customers associated with ongoing sales of Primo and Alma
- Some erosion of customers to Serials Solutions
Serials Solutions KnowledgeWorks and 360 Link

- KnowledgeWorks a core product of Serials Solutions
- Emerged from E-Journal lists
- Original ArticleLinker became 360 Link
- Multi-tenant software as a service for all products
- Uses both automated processes and human intervention to maintain KB
- Created Summon as the first index-based discovery service
- Recently announced Intota as its comprehensive library management product
OCLC WorldCat Knowledge Base

- Originated from the 1Cate product acquired from Openly Informatics
- Integrated into the WorldCat platform
- Basis for WorldCat Link Manager
- No longer a separate utility but provides the link resolution functionality of WorldCat Local
- Part of the underlying infrastructure of WorldShare License Manager
EBSCO LinkSource and Knowledge Base

- EBSCO uses SmartLinks to create direct links within the EBSCOhost environment
- LinkSource provides link resolution to content outside EBSCO products
- Very large number of libraries use EBSCO A-to-Z
  - Recent product license changes makes LinkSource available to any library subscribing to A-to-Z
Lower-level product

Smaller knowledge base with fewer automated processes for updates and management

Lower cost
Simon Fraser University: CUFTS GODOT

- Open source link resolver
- Open access knowledge base
  - Maintained primarily by technical services group at Simon Fraser University
- TrueSerials offers commercial subscriptions.
Kuali OLE: Global Open Knowledge Base (GOKb)

- Proposal to the Andrew W. Mellon Foundation
  - Submitted Jan 4, 2012 (Recently Approved)
  - $498,621 requested for 1-year project
  - April 1 2012 – March 31, 2013

- Collaboration between JISC and KualiOLE

- Lead institution: North Carolina State University

- Principal Investigators:
  - Susan Nutter (NCSU)
  - Rachel Bruce (JISC)
Broad Objectives

- Create an community supported knowledge base that can be used by multiple projects for the management of electronic resources
- Initially for JISC and Kuali OLE, but also for other institutional and commercial uses
- Hosted on a platform built on open source software
  - Deployed at multiple host sites (Initially Europe and US)
  - Data and services delivered through open API’s
Addresses current problems

- Insufficient data regarding publishers and providers; simplistic representation of relationships among content and provider roles
- Absence of identifiers for entities in the supply chain
- Inadequate normalization of data across or within systems
- Inability for libraries to openly access and repurpose data within current knowledge bases
- Inability to track changes over time of any given title or other entities over time
Global Open Knowledge Base goals

- A repository of high-quality metadata about the electronic resources and their representation across the supply chain from publisher to licensor to library vendor to library customer, including representation of marketing and packaging constructs as defined by the suppliers;
- the ability to track changes to product entities and their associated organizations over time;
- an open services platform that will enable libraries and others to employ that data in local systems.

Quoted from the proposal to the Andrew W. Mellon Foundation
Data needed for Electronic Resource Management

1. **Publications & their packaging** - E-resource titles and their associated packages and platforms as supplied by a given publisher or provider, associated with brief descriptive metadata, such as subject terms and identifiers.

2. **Actors** - Information about organizations and their role in the supply chain (publisher, provider, licensor, subscriber).

3. **Terms** - Information about licensing, terms of acquisition, use and ongoing access.

4. **History** - The historical record of changes in such as titles and organizational roles, against which license terms and institutional entitlements can be recorded.

5. **Local entitlements (holdings)** - typically date ranges (in various formats) and statements about embargoes if applicable and access information (URL).

Quoted form the proposal to the Andrew W. Mellon Foundation.
Proposed Solution

- Bill of Materials data model
  - used in engineering and manufacturing sectors
  - Data for item can be exploded into constituent parts
- Ensure compatibility with open linked data expectations through RDF data modeling
- Relational RDMS as fallback
Data Elements

- **Organization.** Institutional entity with a role in the supply chain.
  - Elements: Names, roles (publisher, provider, licensor, vendor, subscriber), administrative information about organizations.

- **Standard numbers.** Identifiers for the journal title.
  - Elements: ISSN, e-ISSN, ISSN-L, DOI

- **Collection (combination).** A branded set of titles instances offered by a provider.
  - Elements: Names, title IDs, internal identifier.

- **Title instance (component).** A title as supplied by a given licensor.
  - Elements: Names, identifiers, base URLs, terms under which a title is offered ("status"), entitlement offered

- **Platform.** A software solution for hosting journal title instances.
  - Elements: Names, internal identifier.

- **Model license agreement.** Terms of acquisition and use for a deal or titles instances as offered by the provider and prior to local negotiations.
  - Elements: Document, coded values, deal ID, title instance IDs, internal ID.

Quoted from the proposal to the Andrew W. Mellon Foundation
License for GOKb

- Science Commons Protocol for Implementing Open Access Data

- Open Data Commons Public Domain Dedication and License (ODC-PDDL)

- Creative Commons CC0 waiver
Sustainability and Data Maintenance

- Project will produce initial population
- Develop ongoing methods of updates
KnowledgeBase+

- develop a centralised, shared, above-campus knowledge base of data useful to electronic resources management (ERM) at institutional level
- provide institutions, and the services they use, with timely, accurate, verified and structured ERM information, including e-resources publication, licensing, subscription and entitlements data
- extend and enhance existing data and services rather than creating yet another silo
Society of College, National and University Libraries

Funded a 2011 project to analyze use case process in UK Higher Education institutions