SAGE White Paper on Discoverability in the Twenty-First Century: Collaborative Opportunities for Publishers, Vendors, and Librarians

Fiesole Collection Development Retreat Pre-Conference
Fiesole, Italy, April 12, 2012

Mary M. Somerville, MLS, MA, PhD
University of Colorado Denver, USA

Disclaimer: This white paper was supported by SAGE in an effort to contribute further to the conversation and debate around discoverability. It does not necessarily reflect the views or policies of SAGE. Additionally, the views of this presenter, while supported by SAGE, do not necessarily reflect the views or policies of SAGE.
Discoverability White Paper  Scope:

- Best practices for access and discovery of content in libraries
- Big problems that publishers, vendors, and libraries need to solve
- Real solutions that librarians and publishers can implement
- Further observations for improving discoverability and visibility


Symbiotic (but disrupted and fragmented) Scholarly Ecosystem:

- Librarians manage systems for institutional collection, dissemination, and retrieval of scholarly corpus
- Publishers produce and promote authors’ work through formats findable on the open Web and in library catalogs
- Publishers’ technology vendors supply e-pub platforms and strategic SEO advice
- Libraries’ technology vendors connect publishers’ digital content to OPACs through ERMs and web scale discovery services
Cross Sector Interviewees:

- Scholarly publishers and published authors
- Search engine developers, database providers, A&I services, and academic publishers
- Electronic Resource Management (ERM) and Integrated Library System (ILS) vendors
- Librarians who advance institutional discoverability

Some Discoverability Starting Points and Search Catalysts:

- Open URL link resolver for ‘known item’ /citation search
- Specialized bibliographic databases—w/out or with full text
- A-Z database or journal list in OPAC listing library holdings
- Library website research pages listing authoritative resources
- Publisher websites, journal homepages, and alert services
- Web search engines, like Google Scholar, to direct traffic to publishers and libraries
- Discovery layer services which conduct ‘Google like’ searches of formerly ‘silo-ed’ library content through a single index
- Referrals from colleagues, blogs, and other ‘invisible college’ (trusted experts) network, including Web 2.0 media contacts
Discoverability (working definition):

Scholars’ capacity to locate relevant content in the scholarly corpus as needed to advance research and other creative activity at appropriate points in workflow.

Analysis of Industry Experts Found:

• Improved discoverability depends on cross sector strategies
• Discovery acceleration tools in familiar web environments
• Detailed indexing for highly relevant and precise results
• Seamless identification and fulfillment user experiences
Some ‘conversation starters’ to improve discoverability and visibility, access and discovery, and usage and creation of the scholarly corpus:

- Establish common standards for structured metadata, information organization, resource presentation, and usage statistics
- Explore and implement cross-platform and cross-publisher industry best practices and shared standards
- Create online product interfaces and publisher website designs that conform to (yet to be determined) standards and functionalities
- Monitor (changing) researcher behaviors and apply findings to publisher and library educational tools and system and interface redesigns
Collaboration Through Combination of Technologies, Standards, and Practices:

- Open URL (link resolver) navigation technology that shows options for obtaining target content and shows ‘best’ version of scholarly content for which users have ‘rights’ through academic affiliation validated by institutional authentication (National Information Standards Organization/NISO and Knowledge Bases and Related Tools/KBART)

- Open Researcher and Contributor ID/ORCID assigns unique identifiers to associate researches and entities with research outputs, identifies version of record and most recent or authoritative version of given work through its life publication cycle (NISO has also recommended standard version terms and CROSSRef has released a new feature for version validation, CrossMark)
Collaboration Through Combination of Technologies, Standards, and Practices (continued):

- ScholarlyArticle offers structured data schema to enable improved discovery of appropriate content through consideration of a variety of unique properties, including publisher, editor, reviewer, genre, reviews, ratings, institution, location, creation date, and modification date, as well as author, title, and source – all value added signifiers of provenance and authority.
Recent Developments in Cross Sector Communication and Collaboration:

New NISO Open Discovery Initiative Committee will convene open discovery libraries, information content providers, and discovery service providers to develop formal standards and recommended best practices for ‘next generation’ library discovery services using an aggregated index search of a wide range of resources, licensed and free, from multiple providers. This will necessarily require, for instance, creation of consistent vocabulary and business practices to clarify exactly which resources are available in uniquely licensed and purchased electronic content and which are indexed in full text or by citations only, or both, and whether the metadata derives from aggregated databases or directly though the full text. (released October 25, 2011)
Recent Developments in Cross Sector Communication and Collaboration (continued):

National Federation of Advanced Information Service (NFAIS) Discovery Service Code of Practice Draft:

“discovery services have the potential to provide ease of information discovery, access, and use, benefitting not only its member organizations, but also the global community of information seekers. However, the relative newness of these services has generated questions and concerns among information providers and librarians as to how these services meet expectations with regard to issues related to traditional search and retrieval services; e.g., usage reports, ranking algorithms, content coverage, updates, product identification, etc. Accordingly, this document has been developed to assist those who choose to use this new distribution channel through the provision of guidelines that will help avoid the disruption of the delicate balance of interests involved.” (released to NISO, Feb 1, 2012)
Recent Developments in Cross Sector Communication and Collaboration (continued):

COUNTER Code of Practice for Usage Factors Draft:

“provides specifications for the recording and reporting of Usage Factors that are consistent with the COUNTER Code of Practice for E-resources. …including specifications for the metadata to be recorded, the content types and article versions whose usage may be counted, as well as the Publication Period and Usage Period to be used.” (released Mar 2012)
Recent Study on *Collaboration Opportunities Between Publishers, Librarians, and Vendors:*

*Libraries at Webscale* (OCLC) January 2012 reports on the question: “what next?” and concludes “big collaboration in the information ecosystem will come not only from broader collaboration across libraries, library groups, consortia, and partnerships across the broader knowledge community – across researchers, publishers, commercial vendors, and Webscale providers such as Google, Amazon, and Facebook.” (page 31)
OCLC Libraries at Webscale ‘Thought Leaders’ Recommendations:

• “Connect users with content regardless of format or where it is stored by creating new models of partnership with all types of content providers
• Develop new forms of knowledge through dialogue and discourse that are easily distributed, reviewed, and added to the collective collection
• Build creative ‘spaces’ that encourage collaborations of pure exploration and invention among any ecosystem members, organizations, or groups
• Build bridges, links, and tunnels to wells of information that make it easy to find, connect, compare, mix, or mash up all content into any format.” (p. 32-33)
Championing discoverability

- **White paper, released January 2012**
- **The development of more sophisticated discovery and visibility strategies very much depends on heightened cross-sector collaborations**