C N I P *R O J E C T* B *R I E F I N G*

Connecting the Docs:
New models and tools to link bibliographic databases and full text journals

Bette Brunelle and Sean O'Doherty, Ovid Technologies

Abstract:

The full text journal marketplace is complex, confusing, and changing daily. In this presentation, we will describe some common models for full text delivery, and address the advantages and disadvantages of each approach. We will explore the attributes of distributed vs aggregated full text vendors, examine the impact of prepackaged vs. customized groupings of content, and look at the way that static vs. live linking affects administration of full text systems. We will also explore trends for future full text delivery.

With these models as a common basis for discussion, we'll provide an overview of the development of Ovid's OpenLinks Toolkit. The OpenLinks Toolkit provides Ovid sites with the tools to define live links from Ovid bibliographic databases to external full text targets using metadata from the citations. We will discuss this project both in terms of technical and marketplace requirements and challenges, and will also place it in context of available and projected full text delivery models.

F u l l   T e x t   J o u r n a l   M o d e l s
Publisher-supplied Full Text Journals

<u>Definition</u>: Journals published directly by societies and commercial publishers, and made available via standalone websites

<u>The good news</u>

- No longer bound by the constraints of print, publishers can add prepublication data of interest
- Publishers are well-positioned to add value in terms of additional content, including supplementary material (including links to videos, backup data, related websites)
- The electronic form allows continual addition of commentary; letters; corrections and addenda, making the documents alive and constantly changing.
- Society publishers in particular are highly innovative in terms of e-publishing

<u>The challenges</u>

- Keeping track of the myriad of different journal sites and access methods is difficult for both users and librarians charged with making them available to users.
- Sites become silos -- users can't search across journals directly
- Electronic versions often don't contain all of the data available in the print version.
- Electronic publishing continues to be a by-product for many publishers, which means that journals are converted to electronic form after print production, causing quality control issues and additional costs.
- Publisher expertise and investment in online journal operations vary, and many are not prepared for the requirements of a large scale 24x7 operation, nor of the ongoing requirements to improve software.

Aggregator-supplied Full Text Journals

<u>Definition</u>: Journals published through 3rd party aggregators who offer a common interface and search system for accessing many journals simultaneously at a single website.

<u>The good news</u>

- Aggregators can offer more and better integration, both between full text sources by linking references, and also between bibliographic and full text resources
- Aggregators can enable full text searching across many full text sources, allowing users to search the full text directly and naturally.
- Aggregators provide one-stop shopping for lots of content and typically provide a high level of online service to their customers.
- Integrated resources available from one source provide seamless access, including common software tools and search techniques, and consistent system performance.
- Because they have access to a variety of content forms, aggregators are in a position to create products that integrate content from a variety of sources (journals, textbooks, dictionaries) and which present different levels of content as needed

<u>The challenges</u>

- Aggregators usually do not own content and must work to gain and keep rights to distribute the content.

- Some publishers will not license content to aggregators, and users want access to this content as well.
- Aggregated content may not be as current, either because of contractual restrictions or because of processing constraints
- Aggregators usually do not get access to the additional materials and commentary when made available on publisher sites, creating "versioning" and archival issues.
- Users often want access to the aggregator version and the publisher version of a journal, but don't want to pay for it twice.

Distributed, linked Full Text Journals
Definition: Journals are located at a variety of sites, and linked via bibliographic databases. In some cases, this model can be a hybrid of the two above.
The good news

- Full text access is provided via familiar bibliographic databases and is a natural and expected extension of current products
- Users can choose the bibliographic tools and system that they like best to use for linking

The challenges

- There are currently no real standards for linking, so bibliographic database providers and full text publishers alike must register their full text and exchange linking protocols and/or exchange data in order to create links.
- There's no cost model for linking, but links cost money to produce and maintain.
- The linking database products offer either generic or subscription-based links. Generic links mean users will find links that don't work because they don't have access. Subscription-based links avoid unusable links but require administrative overhead to register subscriptions.
- Journal site performance and software tools will vary substantially.
- This model solidifies the bibliographic database searching model (which may need updating), and increases the challenges of integration of different types of content.
- The full text is not searchable.   Since 40-60% of the material in journals does not contain abstracts, searching is limited to titles and indexing, substantially reducing their accessibility.

THE OVID OPENLINKS TOOLKIT PROJECT

Project Goals:

- Provide a method for integrating Ovid bibliographic databases with non-Ovid full text
- Provide as integrated access as possible in the distributed world
- Provide "open" tools to allow sites to create their own links
- Make "live links" based upon citation metadata rather than production-based links
- Include predefined links to popular full text targets

Product Features:

- Links can be system wide or database-specific
- Many target search syntaxes are supported
- Journal identifiers can be mapped as needed
- OpenLinks will only appear for linked journals and available years of coverage
- Supports HTML, PDF, ASCII and TIFF full text formats
- No coordination of links and updates required
- Can link to other resources, not just journals

Project challenges we'll discuss:

- Each link target may have require a different syntax
- Not all targets support direct journal access from citation metadata
- How do we handle the differences between databases in terms of available metadata?
- How do we handle changing journal names and ISSNs?
- What is the best way to manage site-specific information to support customized links?
- What is the best way to communicate with users about what happens when they link and how to get back to their search session?
- How and will users differentiate between external links and internal links?