Data flow from publisher to end user via Ex Libris
Presented by Amelia Rowe for the ANZREG 2013 Technical Seminar

Aim
To present the overall data flow behind library discovery services, such as ExLibris
This is a "what happens to bring the information to users via a library discovery service"

A quick overview of a few terms I use freely:
Content providers [Informit] <publisher/database supplier/indexer>

Library discovery services – [ExLibris] is a discovery tool for librarians and their users

- A one stop shop
- One entry point to all [or at least most] of a libraries resources

When publishers talk about the process of getting their information into these discovery services we often talk about syndication
Syndication is:
"Sharing data to engage and promote a common interest"

The common interest is: Users (students/researchers/librarians/etc.)

An overview of the data flow

Resource published
↓
Indexing
↓
Metadata delivery
↓
Metadata processing
↓
Library Discovery service

What I want to talk about is what happens at each of these points in the data flow to bring the data to users, using the flow from Informit to ExLibris as an example

Step 1: Resource published

- I think that is self-explanatory

Step 2: Indexing
The resource is indexed – for example by RMIT Publishing for inclusion on the Informit databases

These first two steps are common steps for most publishing companies

This is where things get different from everyday

Step 3: Meta data delivery

There are two forms of data that Informit as a producer of the metadata provides:

1. Index Metadata

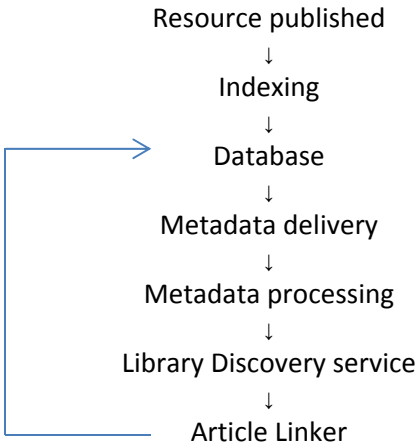
Metadata makes up the ExLibris knowledge base
What: Individual index records for each article indexed
How: XML file supplied on a regular basis (weekly for collections, monthly for PT to reflect their update cycles)

2. Title/holdings data

Contributes to the article linker (SFX)
What: Title and holdings metadata – title, ISSN, and coverage information
How: CSV file accessed by the ExLibris team (monthly?)

This is used to link users to the available articles
Why use a Link service? –allows the discovery service to direct users to holdings they have access to

This enables the flow of information to jump back to the content provider, taking the user direct to the source of the information



Step 4: Metadata processing
The discovery service (ExLibris) needs to process the data provided (by Informit) into their own metadata schema so that it is searchable on their own system.

- This includes mapping metadata tags to match their own fields, this ensures that all records from the many content providers have a uniform record structure, also ensures records display the same to users every time
  <each content provider may use their own tags – such as Subject LC, or MESH, or FAST which may all be mapped to a single Subject field>
- Some fields may require special processing
  <such as subject fields to be facetted out to make it easier for users to search, or Normalisation>
- *De-duping of records*

<If multiple content providers provide data for the same article ExLibis can match these articles together and present the user with a single record then use their article linker SFX to present the user with the options for how to reach the content>

Step 5: Library Discovery service
Once the previous steps are complete the user can access resources from multiple content providers via a single interface.

But this isn't without it's **risks**
A long metadata flow/chain like this one runs the risk of breaking
Like Chinese whispers

Many things can cause a break in the chain which can prevent information from being found, or from being accessible.

A few examples include:

- Incorrect data may get carried to the user (such as an incorrect author entry at the indexing level, or a misprint by the publisher)
- Incorrect title coverage can cause issues when linking
- Metadata from two sources that doesn't match (can cause issues if de-dupped)
- A break in a single link of the system can cause one to fail finding a resource or to fail to access it (such as with ISSNs)
- Out of date or obsolete information may be retained

What is happening to help overcome these limitations?
Standardisation

1. KBART standards are providing us with a better way to present our title and holdings metadata
2. NISO (National Information Standards Organisation) currently have and Open Discovery Initiative (ODI) which is looking at resource discovery to look at a more standardised way to engage in the "the interactions between the creators of these services and the content providers"

Reference:
NISO standards report – ODI survey report: Reflections and Perspectives on Discovery Services
By NISO ODI Working Group (January 2013)