Deep Blue / Michigan Research Experts Integration, Next Century Library Share Summit

Welzenbach, Rebecca

http://hdl.handle.net/2027.42/148275
Deep Blue / Michigan Research Experts Integration
Rebecca Welzenbach, Research Impact Librarian
Next Century Library Share Summit
November 14, 2018
The Michigan Research Experts Information Architecture

UNIVERSITY OF MICHIGAN FEEDS
- Person metadata
- Grants metadata

SHIBBOLETH SSO
User authentication to support single sign on

EXTERNAL DATA FEEDS
- Dimensions
  Scopus, Google
  Web of Science, Crossref
  Europe PubMed Central
- CiNii
  PubMed, RePEc, dblp

DATA SYNC
Harvest
Harvest Deposit

DEEP BLUE
Institutional Repository (Dspace)

Altmetric
Online attention tracking for over 12m outputs

MICHIGAN RESEARCH EXPERTS
https://experts.umich.edu

Dimensions
Dimensions Profiles
Discovery and analysis platform for data fed from UMICH Elements system. Currently includes:
- 4,689 experts
- 2,460,564 publications
- 17,640 grants
- 11,901 patents
- 1,895 clinical trials

Dimensions Plus
Richly interlinked research database used by UMICH to analyse global research and contextualise their work. Currently includes:
- 97m publications
- 4.1m grants
- 426k clinical trials
- 361k policy documents

Elements
Research Information Management solution for data collection, curation and enrichment.
Profile data claimed or added by UMICH researchers is used to automatically populate downstream systems including Dimensions Profiles, UMICH Deep Blue Repository and other reporting tools.

Dimensions
Elements
Data source for additional UMICH Grants, Patents and Clinical Trials not yet captured in Elements

Dimensions
readcube
Enhanced PDF Viewer for OA articles in Dimensions Profiles and Dimensions Plus

Source: Byrne, Kate and Stephen Cawley. Connections, Collaborations, & Impact: Data-Driven Approaches to Understanding institutional research expertise. Digital Science case study. October 2018 (pending publication)
The Michigan Research Experts
Information Architecture

- UNIVERSITY OF MICHIGAN FEEDS
  - Person metadata
  - Grants metadata

- Shibboleth SSO
  User authentication to support single sign on

- External Data Feeds
  - Dimensions
    - Scopus
    - Google Scholar
    - Web of Science
    - arXiv.org
    - CrossRef
    - Europe PubMed Central
    - CiNii
    - PubMed
    - RePEc
    - dblp

- Altmetric
  Online attention tracking for over 12m outputs

- Michigan Research Experts
  https://experts.umich.edu

- Elements
  Research Information Management solution for data collection, curation and enrichment.
  Profile data claimed or added by U-Mich researchers is used to automatically populate downstream systems including Dimensions Profiles, U-Mich Deep Blue Repository and other reporting tools

- Deep Blue
  Institutional Repository (Dspace)

- Dimensions
  Discovery and analysis platform for data fed from U-Mich Elements system.
  Currently includes:
  - 4,859 experts
  - 260,566 publications
  - 17,640 grants
  - 11,901 patents
  - 1,895 clinical trials

- Dimensions Profiles
  - 97m publications
  - 4.1m grants
  - 426k clinical trials
  - 364k policy documents

- Dimensions Plus
  Richly interlinked research database used by U-Mich to analyse global research and contextualise their work.
  Currently includes:
  - 97m publications
  - 4.1m grants
  - 426k clinical trials
  - 364k policy documents

- Data Sync
- Harvest
- Deposit

- ReadCube
  Enhanced PDF Viewer for OA articles in Dimensions Profiles and Dimensions Plus

This is a diagram showing the underlying information architecture and connections behind Michigan Research Experts
RT2 - Repository Integration

Add a to-do
- Please confirm all possible ways of how content is currently ingested into the IR. 2 comments (Completed by Jim Ottaviani on 12 Sep)
- If you have an existing IR, how many items does it currently contain? 3 comments (Completed by Jim Ottaviani on 12 Sep)
- What repository platform is your IR based on? 2 comments (Completed by Igor Kondrashov on 11 Sep)
- Do you require one way or bidirectional sync? 2 comments (Completed by Igor Kondrashov on 11 Sep)
- If you use a third party to host or maintain your IR please confirm provider. 1 comment (Completed by Igor Kondrashov on 11 Sep)
- Do you maintain the IR server in-house or is it a hosted service? 2 comments (Completed by Igor Kondrashov on 11 Sep)

RT2 - Preparation

Add a to-do
- Confirm that Deep Blue folks can log in to their Elements accounts. 1 comment (Completed by Rebecca Welzenbach on 11 Oct)
- Review repository tools integration documentation. 8 comments (Completed by Jim Ottaviani on 2 Oct)
- Besancamp training for the Deep Blue Team (Jim Ottaviani, Martha Stilt, Jose Blanco, Becky Welzenbach). We want to make sure we're following best practices and that we assign appropriate roles (who checks off that tasks are complete, etc.) - Not sure who to assign this to, so starting w/Igor. Please re-assign as appropriate. Or if it's a full session doesn't seem needed, perhaps we can confirm practices and roles on our next weekly call. (Completed by Igor Kondrashov on 25 Sep)
- Create a table of all the metadatas in DeepBlue. 2 comments (Completed by Igor Kondrashov on 25 Sep)
- Configure the repository API. 5 comments (Completed by Igor Kondrashov on 25 Sep)
- Provide details of current repository software. 1 comment (Completed by Igor Kondrashov on 30 Aug)
- Ports and Firewalls. 2 comments (Completed by Igor Kondrashov on 20 Aug)
- Request Repository Licence from Symplectic. 2 comments (Completed by Igor Kondrashov on 10 Aug)

To-do lists

Add a to-do

RT2 - Harvest Crosswalks

- Build the crosswalk map file (at least 0.5 day per object type intended for harvest) 8 comments (Evelina Buicig - Fri, Nov 2)
- Take full backup of Elements, enable verbose logging, limit volume of data for testing phase (1 day) Evelina Buicig - Mon, Nov 5
- Test and review crosswalk (1 week) Jim Ottaviani - Mon, Nov 12
- Adjust and retest crosswalk (1 week) Evelina Buicig - Mon, Nov 19
- Enable harvest (1 day) Jim Ottaviani - Tue, Nov 20

RT2 - Deposit Crosswalks

- Identify functionality required for crosswalk Jim Ottaviani - Tue, Nov 27
- Build the crosswalk map file (at least 0.5 day per object type intended for deposit) Evelina Buicig - Tue, Dec 4
- Take full backup of Elements, enable verbose logging, limit volume of data for testing phase (1 day) Evelina Buicig - Wed, Dec 5
- Test and review crosswalk (1 week) Jim Ottaviani - Mon, Dec 10
- Adjust and retest crosswalk (1 week) Evelina Buicig - Thu, Dec 13
- Enable deposit (1 day) Jim Ottaviani - Fri, Dec 14

Add a to-do
<table>
<thead>
<tr>
<th>Underlying field in Elements</th>
<th>data type</th>
<th>Default mapping field usage</th>
<th>Default source field in Repository</th>
<th>object/item type (internal values) -&gt; artefact</th>
<th>Source field for type ‘artefact’</th>
<th>Field usage for type ‘artefact’</th>
<th>book</th>
<th>Field usage for type ‘book’</th>
</tr>
</thead>
<tbody>
<tr>
<td>title</td>
<td>text</td>
<td>Title</td>
<td>dc.title</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>abstract</td>
<td>text</td>
<td>Abstract</td>
<td>dc.description.abstract</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>authors</td>
<td>person-list</td>
<td>Authors</td>
<td>dc.contributor.author</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>author-url</td>
<td>url</td>
<td>Author URL</td>
<td>[no default mapping]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>editors</td>
<td>person-list</td>
<td>Editors</td>
<td>dc.contributor.editor</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>series</td>
<td>text</td>
<td>Series</td>
<td>dc.relation.ispartofseries</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>edition</td>
<td>text</td>
<td>Edition</td>
<td>[no default mapping]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>volume</td>
<td>text</td>
<td>Volume</td>
<td>[no default mapping]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>pagination</td>
<td>pagination</td>
<td>Pagination</td>
<td>[no default mapping]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>publisher</td>
<td>text</td>
<td>Publisher</td>
<td>dc.publisher</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>publisher-url</td>
<td>url</td>
<td>Publisher URL</td>
<td>[no default mapping]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>place-of-publication</td>
<td>text</td>
<td>Place of publication</td>
<td>[no default mapping]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>publication-date</td>
<td>date</td>
<td>Publication date</td>
<td>dc.date.issued</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>isbn-10</td>
<td>isbn-10</td>
<td>ISBN-10</td>
<td>dc.identifier.isbn</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>isbn-13</td>
<td>isbn-13</td>
<td>ISBN-13</td>
<td>dc.identifier.isbn</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>doi</td>
<td>doi</td>
<td>DOI</td>
<td>[no default mapping]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>medium</td>
<td>text</td>
<td>Medium</td>
<td>dc.format.medium</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
<tr>
<td>publication-status</td>
<td>choice</td>
<td>Status</td>
<td>[no default mapping]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
<td>[as default]</td>
</tr>
</tbody>
</table>
Source: Byrne, Kate and Stephen Cawley. Connections, Collaborations, & Impact: Data-Driven Approaches to Understanding institutional research expertise. Digital Science case study. October 2018 (pending publication)