DATA CITATION INDEX

Nigel Robinson
3 December 2014
Overview

• What is the Data Citation Index
• Progress and current status
• 2015 developments
Data Citation Index

• Discover, track and measure impact of non traditional scientific output in the context of traditional literature

Launched October 2012
4M data records
Data Repositories

Over 1200 repositories identified
## Types of Data by Discipline

<table>
<thead>
<tr>
<th>ART &amp; HUMANITIES</th>
<th>SOCIAL SCIENCES</th>
<th>SCIENCE &amp; TECHNOLOGY</th>
</tr>
</thead>
<tbody>
<tr>
<td>CULTURAL HERITAGE</td>
<td>POLL DATA</td>
<td>MAPS</td>
</tr>
<tr>
<td>LANGUAGE CORPUS</td>
<td>ECONOMIC STATISTICS</td>
<td>ALGORITHMS</td>
</tr>
<tr>
<td>IMAGE COLLECTIONS</td>
<td>LONGITUDINAL DATA</td>
<td>GENOMICS</td>
</tr>
<tr>
<td>RECORDINGS</td>
<td>NATIONAL CENSUS</td>
<td>SKY SURVEYS</td>
</tr>
<tr>
<td></td>
<td>PUBLIC OPINION SURVEYS</td>
<td>ASTROPHYSICS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>REMOTE SENSING</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MUSEUM SPECIMENS</td>
</tr>
</tbody>
</table>
Repository Selection & Evaluation

Considerations:

• Editorial Content
• Persistence and stability
• Adequate descriptive metadata
• Links to research literature
Required Metadata

- Unique ID in repository
- Date provided
- Author
- Repository
- URL/DOI
- Title
- Year Published

• Elements for creation and tracking of a data citation
Metadata Processing

Repository provides metadata feed
  • Deep partnerships and collaboration on metadata handling

Normalisation and enhancement of metadata
  • Controlled vocabularies
  • Indexing

Integrate into Web of Science platform & link citations to literature

Metrics
  • Citation counts
Indexing a Data Repository on Web of Science

- Record types
  - Repository/Source
  - Data Study
  - Data Set
Search Results within the Data Citation Index present the powerful Web of Science options for exploring a body of information. Data becomes discoverable alongside literature.
Data deposition makes it possible to show related data from the repository.
Because data are accessible and able to be cited, they can be linked to publications describing research which uses them.
Link out directly to the original item, in this case a Data Study.
Start to build citation maps associated with data through the association of data and literature.
Provide assistance in how to associate data and literature through citation.
2015 Developments

- Infrastructure in place
- Content expansion
- Citation universe expansion
- Citation tracking tools
Content Expansion

- Beyond data sets
- Non traditional digital scholarly output
  - Software
Citation Universe Expansion

50% of articles mention or cite data

- Capture
- Link to publications

75% informal mentions

25% formally cited in bibliography
Data Citation Tracking Tools

- January 2015

- Web service API to provide citations for specified data items
Data Citation Index

• Discovery of data most important to scholarly research
• Data linked to published research literature
• Measures of data citation, use and reuse with attribution
• New metrics for digital scholarship
Time’s Up!

About your speaker:

Name: Nigel Robinson
Company: Thomson Reuters
Email: nigel.robinson@thomsonreuters.com