Better data-level metrics: Quality and impact

10.30 Welcome and introduction  
  Alex Ball

10.40 Knowledge Exchange, DCC, and CASRAI contributions  
  Alex Ball

11.00 NISO Altmetrics project  
  Mike Taylor

11.20 Making Data Count project  
  Martin Fenner

11.40 RDA/WDS Data Publishing Bibliometrics Working Group  
  Sarah Callaghan

12.00 Open Discussion  
  Chair: Sunje Dallmeier-Tiessen

12.30 Close
Dataset-Level Metrics
Knowledge Exchange, DCC, CASRAI

Alex Ball
University of Bath

‘Better data-level metrics: Quality and impact’
18 November 2015
Outline

Knowledge Exchange

DCC How-to Guide

CASRAI Data-Level Metrics
The Value of Research Data

Report

Workshop
- ‘Making data count: Research data and research assessment’, 11–12 April 2013, Berlin
## Types of metrics

<table>
<thead>
<tr>
<th>Type</th>
<th>Tools</th>
<th>How used</th>
<th>Paper</th>
<th>Shared dataset</th>
<th>Published dataset</th>
<th>Data paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citations</td>
<td>Data Citation Index,</td>
<td>Size dependent</td>
<td>Y</td>
<td>Hard</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>Google Scholar, Scopus,</td>
<td>Direct average</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>DataCite</td>
<td>performance</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Source-based performance</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social media indicators</td>
<td>ImpactStory, Twitter,</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>Facebook</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Readership counts</td>
<td>Mendeley, CiteULike</td>
<td></td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Downloads, Views</td>
<td>Repositories, Journal</td>
<td></td>
<td>Y</td>
<td>Hard</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>platforms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Challenges

- Vicious cycle
  
  Scholars don’t share because it is not rewarded and not worth the time.

  Without metrics, it is hard to find a basis for rewarding data sharing.

  Data sharing/reuse transactions are few, making metrics hard or impossible.

- Need to take access restrictions and embargoes into account
- Need to take scale of data sharing within disciplines into account
Recommendations

- Share good practice regarding data quality review in disciplines.
- Standardize and formalize data citation
- Ensure robust linking between data and publications
- Immediately start collecting metrics that might be used
- Work out which metrics are appropriate for data
How to Track the Impact of Research Data with Metrics

http://www.dcc.ac.uk/resources/how-guides/track-data-impact-metrics

- Why measure the impact of research data
- Impact measurement concepts
- Impact measurement services
- Current issues and challenges
- Tips for raising research data impact
Issues

Title and scope

- How to Measure the Impact of Research Data
  - But cannot measure impact directly
  - Can only measure suggestive surrogates
  - Need to look at reasons underlying the ‘score’
Issues

Title and scope

▶ How to Measure the Impact of Research Data
  ▶ But cannot measure impact directly
  ▶ Can only measure suggestive surrogates
  ▶ Need to look at reasons underlying the ‘score’

▶ How to Track the Impact of Research Data
  ▶ But metrics don’t cover all ways of having impact
  ▶ Risk of losing focus!
Issues

Title and scope

▶ How to Measure the Impact of Research Data
  ▶ But cannot measure impact directly
  ▶ Can only measure suggestive surrogates
  ▶ Need to look at reasons underlying the ‘score’

▶ How to Track the Impact of Research Data
  ▶ But metrics don’t cover all ways of having impact
  ▶ Risk of losing focus!

▶ How to Track the Impact of Research Data with Metrics
  ▶ Political issues: academic freedom, creativity, censure of work that doesn’t ‘fit’ the metrics used
Issues

Twitter: tweet counts might be
  ▶ inflated by auto-tweets from repositories, publicity engines, or fraudster accounts
  ▶ suppressed if tweeting ‘not the done thing’ in a community

Credit models
  ▶ Need to go beyond download stats for dataset itself
  ▶ Consider the network of reuse: transitive credit

Impact and quality
  ▶ Researchers can do more about quality than impact
  ▶ Both very hard to measure
Use cases

- A **publisher** wanting to display metrics for datasets underlying articles so authors have an idea of the impact of their data
- A **funder** wanting to evaluate funded datasets using a common set of impact-related metrics
- A **university administrator** wanting to evaluate the impact of current and potential staff using a common set of impact-related metrics
- A **data repository** (domain-specific or generalist) wanting to demonstrate to stakeholders the quality and impact of their holdings
- A **general data repository** wanting metrics or criteria for evaluating the quality of incoming datasets
Quality DLMs for Repositories WG: Inputs

- CICSNC Data Stewardship Maturity Matrix
- NERC Data Value Checklist
- FORCE11 FAIR Principles
- UK Data Archive data review
- University of Edinburgh Deposit Checklist
Quality DLMs for Repositories WG: Questions

- What potential impact/reuse do you anticipate from this dataset?
- How has the dataset been licensed?
- Are there legal or access requirements for the data?
- How has your dataset been documented for reuse?
- Is the data structured and labelled?
- To what degree is your dataset reproducible?
- Have you provided additional disciplinary-associated metadata?
- Has the data been peer reviewed (externally)? What was the result?
- What kinds of format(s) is your dataset in?
- Was the data collected according to an accepted methodology?
- Does your submission have integrity?
We need you!

WG Chair  Alex Ball
WG Scribe  Rebecca Lawrence (notes in meetings); David Baker/Chair/Point Person(s) (blog posts)

WG Point Person(s)  TBD

WG Facilitator/Coordinator  CASRAI Secretariat (David Baker)

WG Modeller  CASRAI Secretariat (Paul Kiel)

WG Participants
Thank you for your attention

Alex Ball: http://alexball.me.uk/

CASRAI Quality DLMs for Repositories WG:
http://www.casrai.org/Quality_DLMs_for_Repositories