

Implementing Data Citation

Dr. Helena Cousijn
July 12, 2016

What do we want to achieve?

highly site specific, potentially limiting their wider value. However, applying the approach as conducted in this paper to data such as that presented by Barnett et al. (2013) to derive relative values for different organisms should provide a more generic set of 'reference data'. In taking the REML approach forward it will be beneficial to target

References

- Barnett et al., 2013 C.L. Barnett, N.A. Beresford, L.A. Walker, M. Baxter, C. Wells, D. Copplestone
Element and radionuclide concentrations in representative species of the ICRP's reference animals and plants and associated soils from a forest in North-west England
NERC — Environmental Information Data Centre (2013) <http://dx.doi.org/10.5285/e40b53d4-6699-4557-bd55-10d196ece9ea>

How do we get there?



DATA CITATION IMPLEMENTATION PILOT (DCIP)



EG3 DCIP: PUBLISHER EARLY ADOPTERS

GOALS

- Our primary goal is to lay out a Publishers' Data Citation Roadmap to aid in implementation of data citation in articles

Following the 'life of a paper'

Presubmission > providing guidance



Submission > capturing the right information



Production > developing the underlying infrastructure



Publication > displaying the data citation

Presubmission

- Specify a policy for data citation
- Include this policy in author information and provide suggestions how to adhere to the policy
- Provide authors with clear instructions on how to cite data (which elements should be included) and examples of data references

We strongly encourage authors to include relevant datasets within their reference list (e.g., Pihowar HA, Vision TJ, Whitlock MC. Data from: data archiving is a good investment. Dryad Digital Repository. 2011. doi:10.5061/dryad.j1fd7); when including datasets in your reference list, please remember that they should be accompanied by an in-text citation.

- Provide information for editors and reviewers
- Update guidelines for internal customer services, so that questions can be answered

Submission

Different ways to capture the relevant information:

1) *Within the submission system*

Have you already uploaded ALL of your primary data files to a publicly accessible repository? *

Yes No

Please enter all details for datasets uploaded to publicly accessible repositories in the fields below.

Repository Name *	Title *	Accession Number *	Url *	Password	
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="button" value="Clear"/>
<input type="button" value="Additional Data Files"/>					

2) *Within the reference list of the manuscript*

[1] Pavlicek,A., Noskov,V.N., Kouprina,N., Barrett,J.C., Jurka,J. and Larionov,V. Pan troglodytes BRCA1 (BRCA1) gene, complete cds, GenBank (2003), AY365046.1, <http://www.ncbi.nlm.nih.gov/nuccore/37953274>

3) *Through accession numbers in the manuscript*



Production

```
242 <!--Data reference: Dryad dataset -->
243 <ref id="bib8">
244   <element-citation publication-type="data">
245     <person-group person-group-type="author">
246       <name>
247         <surname>Kok</surname>
248         <given-names>K</given-names>
249       </name>
250       <name>
251         <surname>Ay</surname>
252         <given-names>A</given-names>
253       </name>
254       <name>
255         <surname>Li</surname>
256         <given-names>L</given-names>
257       </name>
258       <name>
259         <surname>Arnosti</surname>
260         <given-names>DN</given-names>
261       </name>
262     </person-group>
263     <year iso-8601-date="2015">2015</year>
264     <data-title>Data from: Genome-wide errant targeting by Hairy</data-title>
265     <source>Dryad Digital Repository</source>
266     <pub-id pub-id-type="doi" assigning-authority="Dryad Digital Repository" xlink:href="http://datadryad.org/resource/doi:
267     </element-citation>
268 </ref>
```

Publication

- Display data citation in article html & PDF; in the reference list, dataset identifiers including DOIs should be expressed as full URLs
- Dataset persistent identifiers in the text should resolve to the landing page of the dataset.
- Publishers may choose to include reference citations to datasets in a separate or sub-list to the main references.

Data Citations

1. Haroon, M. F. *National Center for Biotechnology Information (NCBI) BioProject database PRJNA289734* (2015).

2. Haroon, M. F. *Figshare*
<https://dx.doi.org/10.6084/m9.figshare.3362899.v1> (2016).

- Next to a citation to the dataset, underlying data can be highlighted and/or visualized within or next to the article

DCIP EG3: PUBLISHER WORKSHOP - LONDON

July 22, 2016 - London, England (INVITATION ONLY)

Add to Calendar

RSVP Form

Please RSVP to the workshop using the form here: [LINK TO FORM](#)

Description

The DCIP Publishers Expert Group is holding a one-day workshop for publishers on the Implementation of the FORCE11 Data Citation Principles at the Springer Nature Offices in Kings Cross on July 22, 2016 from 9:00am to 17:00pm (BST). This workshop will provide all publishers interested in implementing the data citation principles the opportunity to meet and develop a joint roadmap that will provide a clear overview of the steps that will need to be taken.

Thank you!

h.cousijn@elsevier.com

<https://www.force11.org/group/dcip/eg3publisherearlyadopters>

<https://www.force11.org/group/6956/workshops-events>