Education and Training in the Discipline

Peter L. Pulsifer
peter.pulsifer@colorado.edu
Presentation Overview

- Establishing an education and training mandate
- Drawing on existing resources
  - Specific to data citation
  - General DM
  - Formal coursework
- Engaging early career researchers
Establishing an Education and Training Mandate
A closely related issue is that the intellectual work going in creating a good data set should be recognized and ideally cited in formal publications. Citation also aids reproducibility, which means that the citation must have a persistent locator or identifier to the data such as digital object identifier. This in turn means the data need to be professionally stewarded over time.

Fair attribution is also fundamental to an ethically open data environment. Data creators and associated institutions deserve formally recognizable credit for their scientific contributions. IASC encourages this recognition through citation, appropriate co-authorship, and other means, and encourages IASC members to require appropriate credit and reference when serving as editors and reviewers.
Broader Efforts

- ARCTIC DATA COMMITTEE
- Collection of national representatives and champions with an interest in integration and education
Confirmation of Key Priorities (WPs)

1. Documenting and understanding the Arctic data management ecosystem

2. Identifying and promoting common metadata elements

3. Engaging in data citation and publication movement

4. Promoting interoperability through action – interoperability experiment
Drawing on existing resources
Reference Existing Materials

Developing Standards for Data Citation & Attribution in Linguistics
Workshop 1 | Boulder, CO | September 18-20 2015
Basic Training

http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/data-citation-and-linking

Developing Standards for Data Citation & Attribution in Linguistics
Workshop 1 | Boulder, CO | September 18-20 2015
Advanced Training

http://www.dcc.ac.uk/resources/how-guides/cite-datasets

Developing Standards for Data Citation & Attribution in Linguistics
Workshop 1 | Boulder, CO | September 18-20 2015
Advanced Training

https://rd-alliance.org/groups/data-citation-wg.html

Developing Standards for Data Citation & Attribution in Linguistics
Workshop 1 | Boulder, CO | September 18-20 2015
Data Citations

Professional associations in the social sciences are increasingly recognizing the importance of properly citing data in their publications to encourage the replication of scientific results, to improve research standards, and to give proper credit to data producers.

The Data-PASS partners are committed to promoting standards and improving practices for the citation of data. This site offers guidelines and best practices for citing social science research data in order to promote vigorous and consistent attribution of datasets.

The *American Sociological Review* has already adopted a set of standards for citing data after an appeal from the Data-PASS partners. As other peer-reviewed journals and data stakeholders follow suit, consistently applied data citation standards will ensure that research data can be: discovered; reused; replicated for verification; credited for recognition; and tracked to measure usage and impact.

In short, accurate citation of data promotes more and better science, and we believe all data stakeholders can do more to improve data citation. Below are guidelines on how to cite data and what you can do to help.

**How to Cite Data**

Citing data is straightforward. Each citation must include the basic elements that allow a unique dataset to be identified over time:

- Title
- Author
- Date
- Version
UK Data Archives

http://www.data-archive.ac.uk/

Developing Standards for Data Citation & Attribution in Linguistics
Workshop 1 | Boulder, CO | September 18-20 2015
ANDS

Australian National Data Service

Our Vision: More valuable data for Australian research

ANDS is enabling the transformation of:

<table>
<thead>
<tr>
<th>Data that are:</th>
<th>to</th>
<th>Structured Collections that are:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unmanaged</td>
<td>to</td>
<td>Managed</td>
</tr>
<tr>
<td>Disconnected</td>
<td>to</td>
<td>Connected</td>
</tr>
<tr>
<td>Invisible</td>
<td>to</td>
<td>Findable</td>
</tr>
<tr>
<td>Single-use</td>
<td></td>
<td>Reusable</td>
</tr>
</tbody>
</table>

Australian Research Data Commons

ANDS is building the Australian Research Data Commons: a cohesive collection of research resources from all research institutions, to make better use of Australia’s research data outputs.

Research Data Australia

ANDS’ flagship service, provides a comprehensive window into the Australian Research Data Commons. It is an internet-based discovery service designed to provide rich connections between data, projects, researchers and institutions, and promote data sharing.

Developing Standards for Data Citation & Attribution in Linguistics
Workshop 1 | Boulder, CO | September 18-20 2015
General DM Resources

http://commons.esipfed.org/datamanagementshortcourse
Formal Coursework

https://www.lis.illinois.edu/

Developing Standards for Data Citation & Attribution in Linguistics
Workshop 1 | Boulder, CO | September 18-20 2015
Engaging Early Career Researchers
### Topic 1: Data Sharing and Open Science in Polar Research

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:30 - 08:45</td>
<td>Welcome</td>
</tr>
<tr>
<td>08:45 - 09:30</td>
<td><strong>Sharing Data - the International Polar Year legacy</strong> (Speaker: David Corrall)</td>
</tr>
</tbody>
</table>

### APECS World Summit 2015

APECS organized its first APECS World Summit in Sofia, Bulgaria from 6 - 8 June 2015. Since its success during the International Polar Year (IPY), APECS has been grown significantly. At the time of the summit, APECS’s membership counted about 5000 early career scientists, mentors and educators with interests in the Polar Regions in over 80 countries. But APECS has also evolved institutionally with currently 26 National Committees (NCs) across all continents — except Antarctica. The APECS World Summit 2015 “The Future of Polar Research” was an initiative to bring representatives of APECS’ leadership and NCs together to discuss the future of polar research, APECS’ experiences in communicating science and new challenges APECS wants to tackle in the near future. The venue for the APECS World Summit 2015 was Sofia University.

#### Panel Discussions

<table>
<thead>
<tr>
<th>Time</th>
<th>Speaker</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:00 - 12:30</td>
<td>Pavel Stoev, Pensett Publishers / National Museum of Natural History, Sofia</td>
<td><strong>Unpacking the concept of open science</strong> (Speaker: Ivo Grigorov, Danish National Institute of Aquatic Resources)</td>
</tr>
<tr>
<td>12:30 - 13:00</td>
<td>Publisher, Constantia Consulting / Antarctic Environments Portal</td>
<td><strong>The Antarctic Environments Portal</strong> (Speaker: Neil Gilbert, Constantia Consulting / Antarctic Environments Portal)</td>
</tr>
</tbody>
</table>
Acknowledgements

This material is based in part upon work supported by the National Science Foundation under Grant Numbers:

- ARC 0856634 and ARC 1231638

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.
The role of LSA/CoSCIL in standardizing citation and attribution practices in linguistics

Geoffrey Nathan
geoffnathan@wayne.edu
Wayne State University
What is CoSCIL?

- Established by LSA Executive Committee in December 2014
- Charge:
  - assess the challenges and opportunities presented by new technologies for the dissemination of linguistic scholarship
- Chair: Geoff Nathan (formerly chair of COMSIT)
- Members: volunteers and ‘volunteers’
Responsibilities

- To advise the LSA Executive Committee on actions to take with respect to the development of new technologies for the dissemination of scholarly linguistic work, in particular where they may impact the Society's existing programs and activities.

- To assess the ways in which the results of linguistic research can be effectively communicated to diverse audiences of stakeholders in language data, including scholars in allied disciplines, speaker communities, government agencies, and standards bodies.

- To provide opportunities for members of the linguistics community to receive training in the use of new methods for the curation, dissemination, and processing of language data in order to enhance their research, teaching, and advising.

- To coordinate with other LSA committees, delegates/liaisons, and the LSA's sister societies in efforts related to scholarly communication, as appropriate.
Standards Categories

- Text
- Interlinear Glossing
- Video
- Audio
Issues

- a volunteer effort
- meetings are all virtual
- wide variety of experience
- we’re all busy (i.e. not everyone does their homework on time, or at all…)
- We shouldn’t be reinventing the wheel