

Out of Cite, Out of Mind: CODATA Work on Data Citation Practices and Standards

US Data Citation Workshop:
Developing Policy And Practice
23 September 2013

Bonnie C. Carroll
CODATA/ICSTI Data Citation Task Group
On Data Citation Standards and Practices
CODATA Executive Committee Liaison



A Note from Your Sponsor

- CODATA is an interdisciplinary Scientific Committee of the International Council for Science (ICSU), which works to improve the quality, reliability, management and accessibility of data of importance to all fields of science and technology.
- 50 years of service science and society

<http://www.codata.org/uploads/CODATA@45years.pdf>

CODATA Strategy: Mobilizing the Data Revolution

Mobilizing capacity to take advantage of the data revolution is **the** major priority for international science.

CODATA strategy lays out three priorities and a plan that shows we can **deliver benefits** for members on these priorities.

Presents a vision for international promotion and coordination of data ecosystems.

Promote intelligently open data

- **data policies:** supporting implementation of data principles and practice

Adapt to the transformation in research

- **data science:** addressing the frontier issues of data science

Promote data skills, data scientists, data managers

- **research data capacity building** (particularly in LMICs)



CODATA/ICSTI TG on Data Citation - International Context

- Approved at CODATA General Assembly 2010, renewed 2012
- Co sponsored by ICSTI and supported by the US National Committee for CODATA
- Objectives:
 - Examine key issues related to data identification, attribution citation and linking
 - Help coordinate activities internationally
 - Promote common practices and standards

CODATA Data Citation Task Group Participation

Members:

- ❖ 28 Members; 4 Consultants
- ❖ 12 Countries
- ❖ 4 Continents

Co-Chairs:

Jan Brase, (Director, DataCite, and ICSTI representative), Technische Informations Bibliothek (TIB)/German National Library of Science and Technology, **GERMANY**

Martie Van Deventer (Council for Scientific and Industrial Research (CSIR), **SOUTH AFRICA**)

Christine Borgman, University of California, Los Angeles, **USA**

Bonnie C Carroll, CODATA Ex Com Liaison:, Information International Associates, Inc (Chair 2010-2012)

Sarah Callaghan (U.K. CODATA), The NCAS British Atmospheric Data Centre, STFC Rutherford Appleton Laboratory, **UNITED KINGDOM**

Support: **Paul Uhler**, US NC



Activities of the CODATA/ICSTI TG on Data Citation



2010-2012: Background Research

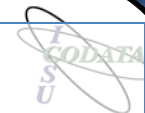
- Bibliographic inventory and analysis (in final report).
- Interviews with a sample of identified stakeholders concerning data citation and attribution practices
 - Data Repositories
 - Publishers
 - Researchers
 - Funding Organizations
- Symposium and workshop held in Berkeley, CA in August 2011.
- Began Report on Current Practices in Data.

2012-2014: Disseminating Results

- Publish Report from August Symposium and Workshop (May 2012). **Symposium and Workshop, Berkeley, CA August 2011: *For Attribution: Developing Data Attribution and Citation Practices and Standards*** 
- Task Group meetings and White Paper drafting sessions (Copenhagen, DK, 06/12, and Taipei, TW, 10/12.)
- 3 Track Sessions at CODATA International Conference in Taipei, TW October/November 2012 on Data Publishing and Data Citation in Cooperation with the WDS
- Publication of final report - **Out of Cite, Out of Mind: Report of the CODATA Task Group on Data Citation** 

Symposium and Workshop, Berkeley, CA August 2011

For Attribution: Developing Data Attribution and Citation Practices and Standards



- I. Why are attribution and citation of data important?
- II. Major technical issues in developing and implementing scientific data citation standards and practices
- III. Major scientific issues in developing and implementing scientific data citation standards and practices
- IV. Major institutional, financial, legal, and socio-cultural issues in developing and implementing scientific data citation standards and practices
- V. Status of data attribution and citation practices in the natural and social sciences in the U.S. and internationally
- VI. Institutional roles and perspectives: similarities and differences across disciplines and countries
- VII. Workshop – Options: Where do we go from here?

For more information on the symposium and workshop see:

http://sites.nationalacademies.org/PGA/brdi/PGA_063656

Out of Cite, Out of Mind: Report of the CODATA Task Group on Data Citation



The report was published by the CODATA Data Science Journal on 13 September 2013

A screenshot of the J-STAGE website. The header includes the J-STAGE logo and navigation links: "About My J-STAGE", "Sign-In", "Register", "Shopping Cart", "Help", and language options for "Japanese" and "English". Below the header is a banner for "DATA SCIENCE Journal" with the CODATA logo and the text "CODATA". The banner also displays "ONLINE ISSN: 1683-1470" and "(As of September 19, 2013) Registered articles: 695". The main content area shows the article title "Out of Cite, Out of Mind: The Current State of Practice, Policy, and Technology for the Citation of Data" by the CODATA+ICSTI Task Group. It includes the DOI link "http://dx.doi.org/10.2481/dsj.OSOM13-043" and the J-STAGE URL "http://www.jstage.jst.go.jp/article/dsj/12/0/12_OSOM13-043". The article is dated "Released on J-STAGE 20130913 [Advance Publication] Released 20130908". Keywords listed are "Data citation, Data management, Data policy, Data publishing, Data centers, Data access, Reuse of data, Metadata, Digital preservation, Information standards, Information infrastructure, Information technologies, Internet, Libraries, Scientific organizations, STM publishers, Research funders, Information metrics". A "Full Text PDF [2943K]" link is provided. The right sidebar contains "Article Tools" such as "Add to Favorites", "Citation Alert", "Authentication Alert", "Additional Info Alert", "Copy the URL", "Mail to Author", "Download Meta of Article", "RIS", "BibTeX", and "Contact us". There is also a "Share this Article" section.

Free and open at:

https://www.jstage.jst.go.jp/article/dsj/12/0/12_OSOM13-043/_article

Chapter 3 EMERGING PRINCIPLES FOR DATA CITATION

Based on a review of emerging practices and analysis of existing literature on citation practices, we have identified the following set of “first principles” for data citation:

1. **Status of Data**
2. **Attribution**
3. **Persistence**
4. **Access**
5. **Discovery**
6. **Provenance**
7. **Granularity**
8. **Verifiability**
9. **Metadata Standards**
10. **Flexibility**

First Principles for Data Citation – Catching Fire!



- Considerable Concurrent International Interest
 - Amsterdam Manifesto on Data Citation Principles from Force11
 - DCC Guide
 - DataCite Metadata Core
 - RDA birds of a feather focus
 - ICSU World Data Service Working Group on Data Publication
- Synthesis Group Formed – over 40 participants
- Consensus Principles In Process

1. Importance

2. Credit and Attribution

3. Evidence

4. Unique Identification

5. Access

6. Persistence

7. Specificity and Verifiability

8. Interoperability and Flexibility

- Developed Broad Dissemination Plan

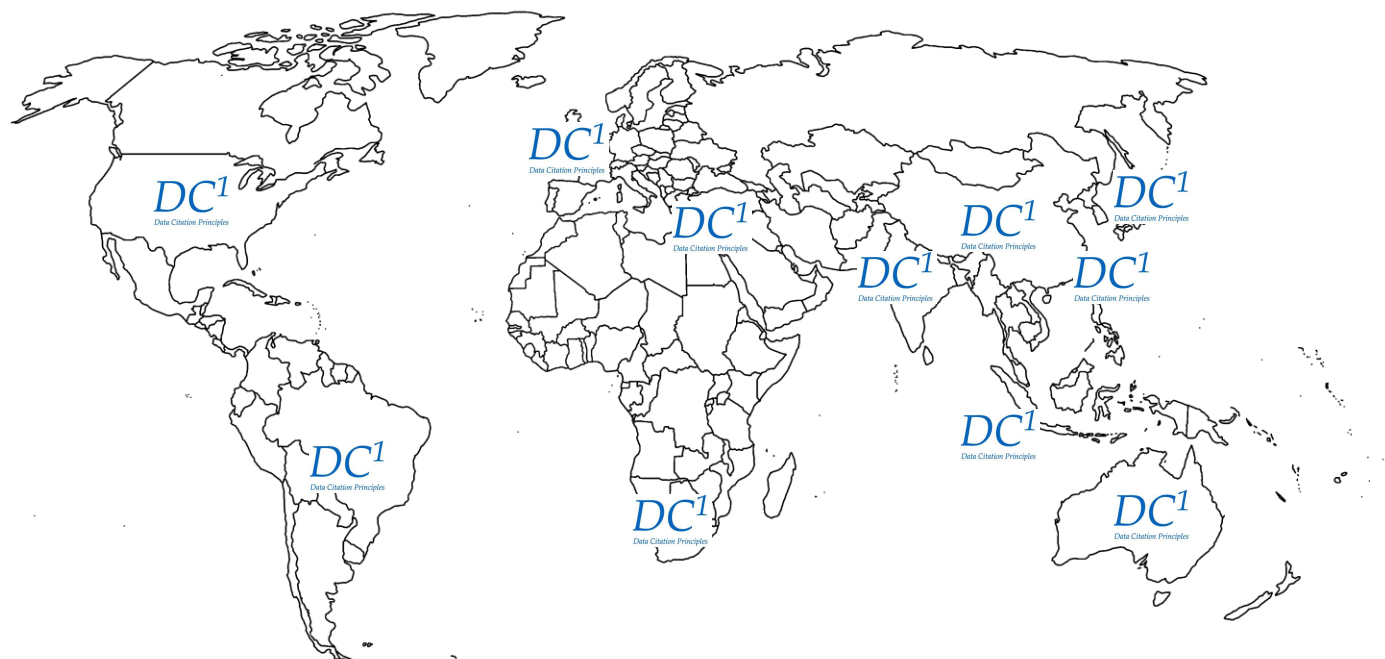
Current endorsements: Individual: 245 Organizational: 110

<https://www.force11.org/datacitation/endorsements>

2014-2016 Data Citation: From Principles to Practice Today

- A **Focus** on the Research Policy and Funding Community':
<http://www.codata.org/task-groups/data-citation-standards-and-practices>
- **Organizing an international series of implementation and adoption workshops.**
- Promote the implementation of data citation principles in the research policy and funding communities throughout the world.
- **Stakeholders** include: government, funders, research performing institutions, research administrators, research librarians, researchers, learned societies, publishers, data archives, journal editors ...
- **Questions to be answered:**
 - What is the policy environment for data citation?
 - What infrastructure is available to support data citation?
 - What are current attitudes to data citation?
 - Benefits and challenges in implementing data citation policies and practice (including economic and financial considerations)?
 - Specific items of concern raised at the workshop?
 - Other aspects also addressed at the workshop?
 - What is the role of the research funding and policy community in data citation practices?

We are taking Data Citation Workshops on a world tour!



2015: China, Australia, Japan, India and South Africa.

2016: Israel, USA, Russia. *Finland*

2017: *France, European Union, Korea, Taipei,*
Indonesia, Brazil...

***To Date Summary Report Draft at
SciDataCon/International Data Week, September 10-17***

The Future: 2016-2018

- Complete the International Summary Report of Workshops on Principles to Practice
- Proposal to Renew TG is a Work In Progress
- Focus on **Impact and Metrics**
- TBD: At CODATA General Assembly, 9/10/16
- US Contact: Franciel Azpurua-Linares
falanares@iiaweb.com

A Thank You to Our Sponsors

We are grateful to the following funders of this project over the past 4 years:

- ❖ Sloan Foundation
- ❖ Institute for Museum and Library Services
- ❖ Library of Congress
- ❖ CODATA
- ❖ Microsoft Research

Thank You!

Bonnie C. Carroll

bcarroll@iiaweb.com