DC1 Data Citation Principles

Joint Declaration of Data Citation Principles

https://www.force11.org/datacitation

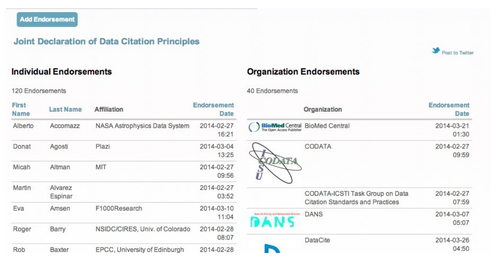
Created by a Synthesis Group that represents more than 25 organizations and consolidates previous efforts on data citation.
https://www.force11.org/datacitationsynthesisgroup

Significance and Scope

- Sound, reproducible scholarship rests upon a foundation of robust, accessible data.
- Data should be considered legitimate, citable products of research.
- Data citation, like the citation of other evidence and sources, is good research practice.
- The Joint Principles cover purpose, function and attributes of citations.
- Specific practices vary across communities and technologies – we recommend communities develop practices for machine and human citations consistent with these general principles.

Endorse the Principles!

http://www.force11.org/datacitation/endorsements


Add Endorsement

Joint Declaration of Data Citation Principles

Post to Twitter

Individual Endorsements

120 Endorsements

First Name	Last Name	Affiliation	Endorsement Date
Alberto	Accomazz	NASA Astrophysics Data System	2014-02-27 16:21
Donat	Agosti	Plazi	2014-03-04 13:25
Micah	Altman	MIT	2014-02-27 09:56
Martin	Alvarez Espinar		2014-02-27 03:52
Eva	Amsen	F1000Research	2014-03-10 11:04
Roger	Barry	NSIDC/CIRES, Univ. of Colorado	2014-02-28 08:07
Rob	Baxter	EPCC, University of Edinburgh	2014-02-28

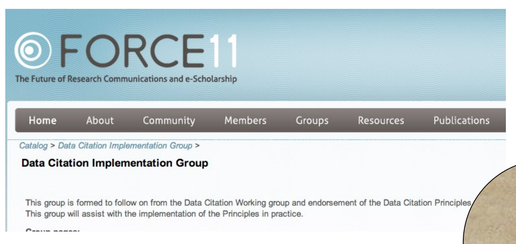
Organization Endorsements

40 Endorsements

Organization	Endorsement Date
BioMed Central	2014-03-21 01:30
CODATA	2014-02-27 09:59
CODATA-ICSTI Task Group on Data Citation Standards and Practices	2014-02-27 07:59
DANS	2014-03-07 05:07
DataCite	2014-03-26 04:50

Join the Implementation Effort!

http://www.force11.org/node/4849


FORCE11

The Future of Research Communications and e-Scholarship

Home About Community Members Groups Resources Publications

Catalog > Data Citation Implementation Group >

Data Citation Implementation Group

This group is formed to follow on from the Data Citation Working group and endorsement of the Data Citation Principles. This group will assist with the implementation of the Principles in practice.

The Noble Eight-Fold Path to Citing Data

Purpose

1. Importance. Data should be considered legitimate, citable products of research. Data citations should be accorded the same importance in the scholarly record as citations of other research objects, such as publications [1].

2. Credit and attribution: Data citations should facilitate giving scholarly credit and normative and legal attribution to all contributors to the data, recognizing that a single style or mechanism of attribution may not be applicable to all data [2].

3. Evidence. In scholarly literature, whenever and wherever a claim relies upon data, the corresponding data should be cited [3].

Function

4. Unique Identification. A data citation should include a persistent method for identification that is machine-actionable, globally unique, and widely used by a community [4].

5. Access. Data citations should facilitate access to the data themselves and to such associated metadata, documentation, code, and other materials, as are necessary for both humans and machines to make informed use of the referenced data [5].

Attributes

6. Persistence. Unique identifiers, and metadata describing the data and its disposition, should persist -- even beyond the lifespan of the data they describe [6].

7. Specificity and verifiability. Data citations should facilitate identification of, access to, and verification of the specific data that support a claim. Citations or citation metadata should include information about provenance and fixity sufficient to facilitate verifying that the specific timeslice, version and/or granular portion of data retrieved subsequently is the same as was originally cited [7].

8. Interoperability and flexibility. Data citation methods should be sufficiently flexible to accommodate the variant practices among communities, but should not differ so much that they compromise interoperability of data citation practices across communities [8].

Principles are supplemented with a glossary, references and examples
http://force11.org/datacitation

Notes

[1] CODATA 2013: sec 3.2.1; Uhlir (ed.) 2012, ch 14; Altman & King 2007

[2] CODATA 2013, Sec 3.2; 7.2.3; Uhlir (ed.) 2012,ch. 14

[3] CODATA 2013, Sec 3.1; 7.2.3; Uhlir (ed.) 2012, ch. 14

[4] Altman-King 2007; CODATA 2013, Sec 3.2.3, Ch. 5; Ball & Duke 2012

[5] CODATA 2013, Sec 3.2.4, 3.2.5, 3.2.8

[6] Altman-King 2007; Ball & Duke 2012; CODATA 2013, Sec 3.2.2

[7] Altman-King 2007; CODATA 2013, Sec 3.2.7, 3.2.8

[8] CODATA 2013, Sec 3.2.10

References

M. Altman & G. King, 2007. A Proposed Standard for the Scholarly Citation of Quantitative Data, *D-Lib*

Ball, A., Duke, M. (2012). 'Data Citation and Linking'. DCC Briefing Papers. Edinburgh: Digital Curation Centre.

CODATA-ICSTI Task Group on Data Citation, 2013; Out of Cite, Out of Mind: The Current State of Practice, Policy, and Technology for the Citation of Data. *Data Science Journal*

P. Uhlir (ed.),2011. For Attribution -- Developing Data Attribution and Citation Practices and Standards. National Academies of Sciences

Placement of Citations

Intra-work:

- *Should provide sufficient information to identify cited data reference within included reference list.*
- *Citation to data should be in close proximity to claims relying on data.* ***[Principle 3]***
- *May include additional information identifying specific portion of data related supporting that claim.* **[Principle 7]**

Example*:* The plots shown in Figure X show the distribution of selected measures from the main data [Author(s), Year, portion or subset used].

Full Citation:

Citation may vary in style, but should be included in the full reference list along with citations to other types works.

Example:

References Section

Author(s), Year, Article Title, Journal, Publisher, DOI.

Author(s), Year, Dataset Title, Data Repository or Archive, Version, Global Persistent Identifier.

Author(s), Year, Book Title, Publisher, ISBN.

Generic Data Citation

(as it appears in printed reference list)

Principle 2: Credit and Attribution (e.g. authors, repositories or other distributors and contributors)

Principle 4: Unique Identifier (e.g. DOI, Handle.). **Principle 5, 6 Access, Persistence:** A persistent identifier that provides access and metadata

Author(s), Year, Dataset Title, Data Repository or Archive, Version, Global Persistent Identifier

Principle 7: Specificity and verification (e.g. the specific version used).
Versioning or timeslice information should be supplied with any updated or dynamic dataset.

Note:
●Neither the format nor specific required elements are intended to be defined with this example. Formats, optional elements, and required elements will vary across publishers and communities. **[Principle 8: Interoperability and flexibility**].
●As illustrated in the previous examples, intra-work citations may be accompanied with information including the specific portion used. **[Principles 7,8]**.
●As illustrated in the next example, printed citations should be accompanied by metadata that support credit, attribution, specificity, and verification. **[Principles 2, 5 and 7].**

Citation Metadata

Author(s), Year, Dataset Title, Data Repository or Archive, Version, Global Persistent Identifier.

Metadata retrieval

```
<!--- CONTRIBUTOR METADATA -->
<contributor role="
ORCIDid="">Name</contributor>

<!-- FIXITY and PROVENANCE --
<fixity type="MD5">XXXX</fixity>
<fixity
type="UNF">UNF:XXXX</fixity>

<!-- MACHINE
UNDERSTANDABILITY -->
<content type>data</content type>
<format>HDF5</format>
```

EXAMPLE METADATA

Note:
●Metadata location, formats, and elements will vary across publishers and communities. **[Principle 8]**
●Citation metadata is needed in addition to the information in the printed citation.
●Metadata describing the data and its disposition should persist beyond the lifespan of the data. **[Principle 6]**
●Citation metadata should support attribution and credit **[Principle 2]**; machine use **[Principle 5];** specificity and verification **[principle 7]**
●For example, additional citation metadata may be embedded in the citing document; attached to the persistent identifier for the citation, through its resolution service; stored in a separate community indexing service (e.g. DataCite, CrossRef); or provided in a machine-readable way through the surrogate ("landing page") presented by the repository to which the identifier is resolved.

For more detail, see the **References** section.
http://www.force11.org/node/4772