Lesson 8: Data Citation

What is Data Citation?

“The practice of providing a reference to data in the same way as researchers routinely provide a bibliographic reference to printed resources.”

“A key practice underpinning the recognition of data as a primary research output rather than as a by-product of research.”

— Australian National Data Service

What is a persistent identifier?

A unique web-compatible alphanumeric code that points to a resource to be preserved for the future.

EXAMPLES
- DOI: Digital Object Identifier
- ARK: Archival Resource Key
- UUID: Universally Unique Identifier
- ORCID: Open Researcher & Contributor ID

A few of the benefits of data citation

- Facilitates discovery of links between data and publications, making it easier to validate and build upon previous work
- Ensures proper credit is given when others use your data
- Makes it possible to estimate dataset impact based on the number of citations
- Links datasets to related methodology in articles
- Makes it easier to discover existing data relevant to a particular research question

Data Citation Principles

- Importance
- Credit & Attribution
- Evidence
- Unique Identification
- Access
- Persistence
- Specificity/
Verifiability
- Interoperability/
Flexibility

https://www.force11.org/datacitation

How to cite data -- what to include

- Author/PI/creator
- Release/pub. date
- Title
- Unique identifier
- Version number
- Data format
- 3rd party producer
- Archive/distributor
- Access date & time
- Subset of data used
- Editor/contributor
- Publication place
- Data w/in larger work

Example data citation formats

DataCite:
Creator (Publication Year): Title. Publisher. Identifier

Dryad:
Author (Date of Article Publication) Data from: article name. Dryad Digital Repository. doi: DOI number.

How to get a DOI for your data

Contact a DataCite-registered organization or institution that can create DOIs, such as:
- DataCite Canada
- California Digital Library's EZID service
- Australian National Data Service (ANDS) Cite My Data Service

Create a citation profile that includes metadata like:
- Creator
- Title
- Publisher
- Date

Best practices to support data citation

- Use applications that support metadata creation
- Use standardized keywords
- Use persistent identifiers
- Work with publishers and repositories to archive data during the publication process
- Encourage others to cite data and make their own data available for reuse

Collaborating to support data citation requires the participation of lots of people

Data authors          Data users
Data managers          Data publishers
Journal publishers     Data repositories
Professional orgs