

Making Data Visible: Data Citation Policy and Implementation at NCAR

Matthew S. Mayernik, Karon Kelly, Mary Marlino, & Mike Wright

National Center for Atmospheric Research (NCAR), University Corporation for Atmospheric Research (UCAR), Boulder, Colorado, USA

Interest in data citations at many levels

- Funding agency workshops and reports
- Professional society statements (AGU and AMS)
- Research organization recommendations (ESIP and DCC)

Numerous motivations for data citations

1. Give researchers and data archives credit for their work in collecting and creating data
2. Track the use of data sets over time, and their derivative research products (e.g. publications, tools, and other data sets)
3. Enable easier data discovery, access, and use by bringing data and their resultant publications closer together
4. Make science and data more transparent

Data citation impediments

- Researchers typically do not cite data, but use acknowledgements or in-text attribution instead
- Overcoming inertia: researchers may not...
 - know that they are asked to cite data
 - know how to cite data
 - know that they have data to cite
 - be willing to spend time creating data citations
- What does "peer review" mean with data?

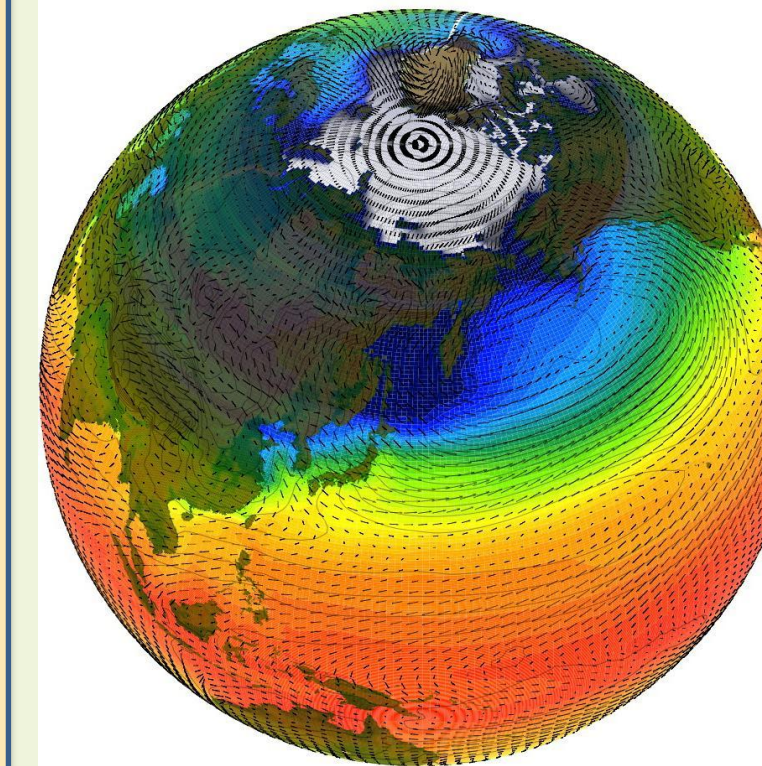
Assigning Digital Object Identifiers (DOIs) to NCAR data

- At what granularity should data be assigned DOIs?
- At what point in the data workflow should a DOI be assigned?
- How should DOIs be assigned to continuously changing data sets?
When does a data set change enough to merit a new DOI assignment?
- Who "owns" DOIs? Who assigns and maintains DOIs in inter-institutional research projects?

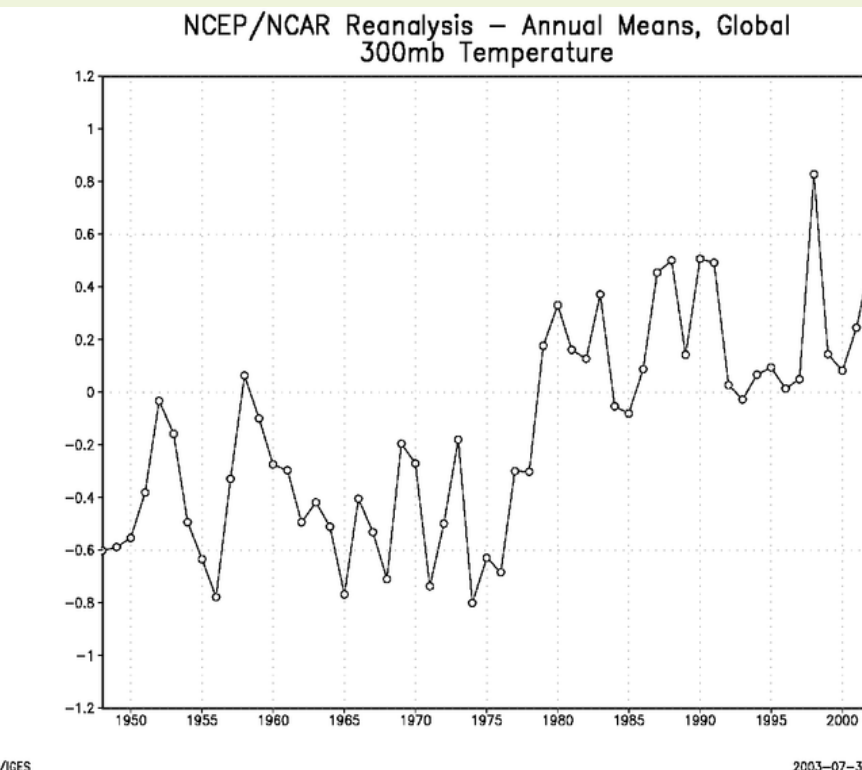
Recommended data citation elements for NCAR data

*Author. Release date. Title. Archive/Provider. Locator/Identifier.
Data access date.*

Data archives at NCAR



Climate model
output data



Observational
time-series data



Data from
field studies

Outreach

Buy-in from data managers
Presentations to NCAR administration
White paper on data citations
Presentations to individual labs

Challenges

How to find and count data citations?
How to encourage institutional acceptance?
How to build services to include data that are not currently archived?

Policies

Who can assign DOIs
What data should be assigned DOIs
How to address granularity and versioning
How to assign and maintain DOIs in collaborative projects

Flexible citation recommendations

Many different kinds of data
Many different journal citation formats