Data citation practices at the Nature Research journals

July 12, 2016

Erika Pastrana, PhD

SPRINGER NATURE

Nature's data availability policies

- Basic principle and minimum standard:
  - ***All authors required to make materials, data, code, and associated protocols promptly available to readers without undue qualifications***
  - Any restrictions on availability must be disclosed to the editors at the time of submission and in submitted manuscript
  - Data sharing mandated for numerous communities
- **2013 update:** Reproducibility checklist, figure source data encouraged
- **2014 update:** Strong preference for data archiving via repositories rather than SI
  - Community repositories preferred, where they exist (use Sci Data's list)
  - General repositories (figshare, Dryad)
  - Authors encouraged to consider preparing Data Descriptors for Sci Data
- **2016 update:**
  - Mandatory DAS (pilot)
  - Data citations for datasets with DOIs (pilot)

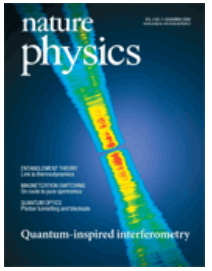
Data citation at Nature journals – key events

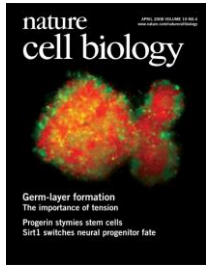
- Accession codes required for various data types and marked up in articles for several years (= data referencing rather than formal citation)
- 2014: Signatory of Joint Declaration of Data Citation Principles http://blogs.nature.com/scientificdata/2014/03/24/endorsing-the-joint-declaration-of-data-citation-principles/
- 2014: Launch Scientific Data
  - Data citation mandated for every article
  - Uses JATS 1.0 with data citations list specifically tagged
- 2016: Data citation policy piloted at Nature journals
  - Strongly encourages datasets with DOIs to be included in reference lists

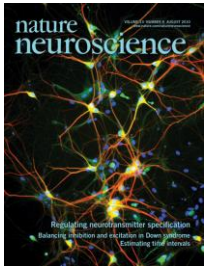
Policy pilot at 5 Nature journals (March 2016 –)

Policy summary

- All manuscripts reporting original research to include a mandatory Data availability statement (DAS)
- Citations of publicly available datasets with DOIs in article reference lists strongly encouraged
- Makes conditions for data availability more transparent to our readers







Motivations

Data availability statements (DASs):

- Effective way to ensure compliance with open data policies (1)
- Association of public data availability with increased citations to papers (2)
- Richer, more consistent reader and reviewer experience
- Supports compliance with funder policies (especially the 7 UK Research Councils)

Data citations:

- Researchers want to know who's using their data!
- Support reproducible research and data reuse
- Increase potential for credit for data sharing
- Helps make data a legitimate, assessable research output

DC1
Data Citation Principles

(1) Vines et al., Mandated data archiving greatly improves access to research data. *The FASEB Journal* doi:10.1096/fj.12-218164 (2013)
(2) Piwowar et al., Sharing Detailed Research Data Is Associated with Increased Citation Rate. *PLOS One* DOI: 10.1371/journal.pone.0000308 (2007)

Data citation at Nature journals following pilot

43. Chen, J. *et al.* Common genetic variation in ETV6 is associated with colorectal cancer susceptibility. Dryad Digital Repository. doi:doi:10.5061/dryad.7dj7t (2016).

Hide context Article

The genotyping data has been deposited in the Dryad Digital Repository (DOI: 10.5061/dryad.7dj7t) (ref. **43**). **in article**

http://www.nature.com/ncomms/2016/160505/ncomms11478/full/ncomms11478.html#ref43

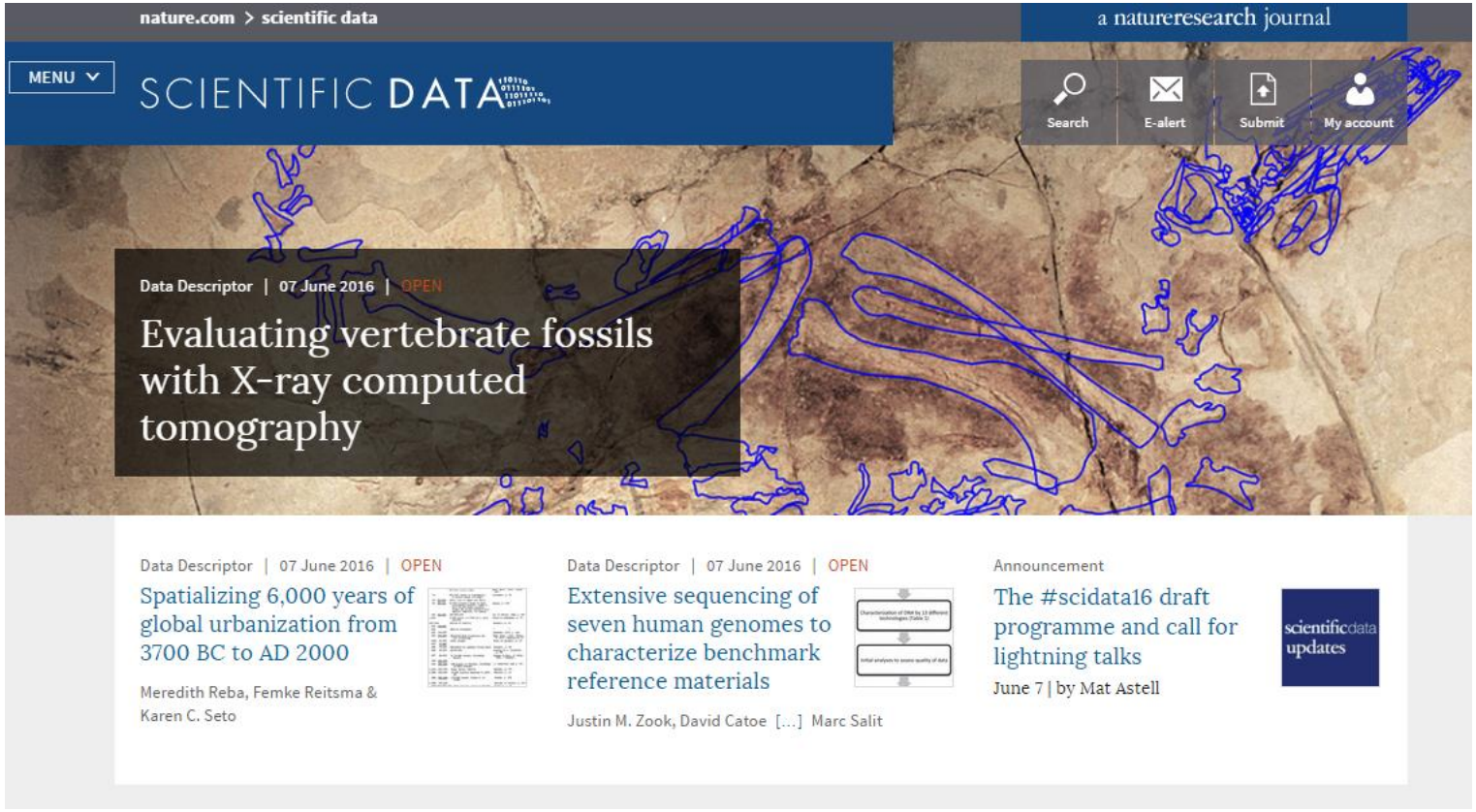
49. Stan, C. A. *et al. Image Data Analyzed in "Liquid Explosions Induced by X-ray Laser Pulses"* (Stanford Digital Repository, 2016); http://purl.stanford.edu/wv179nv3100

Hide context

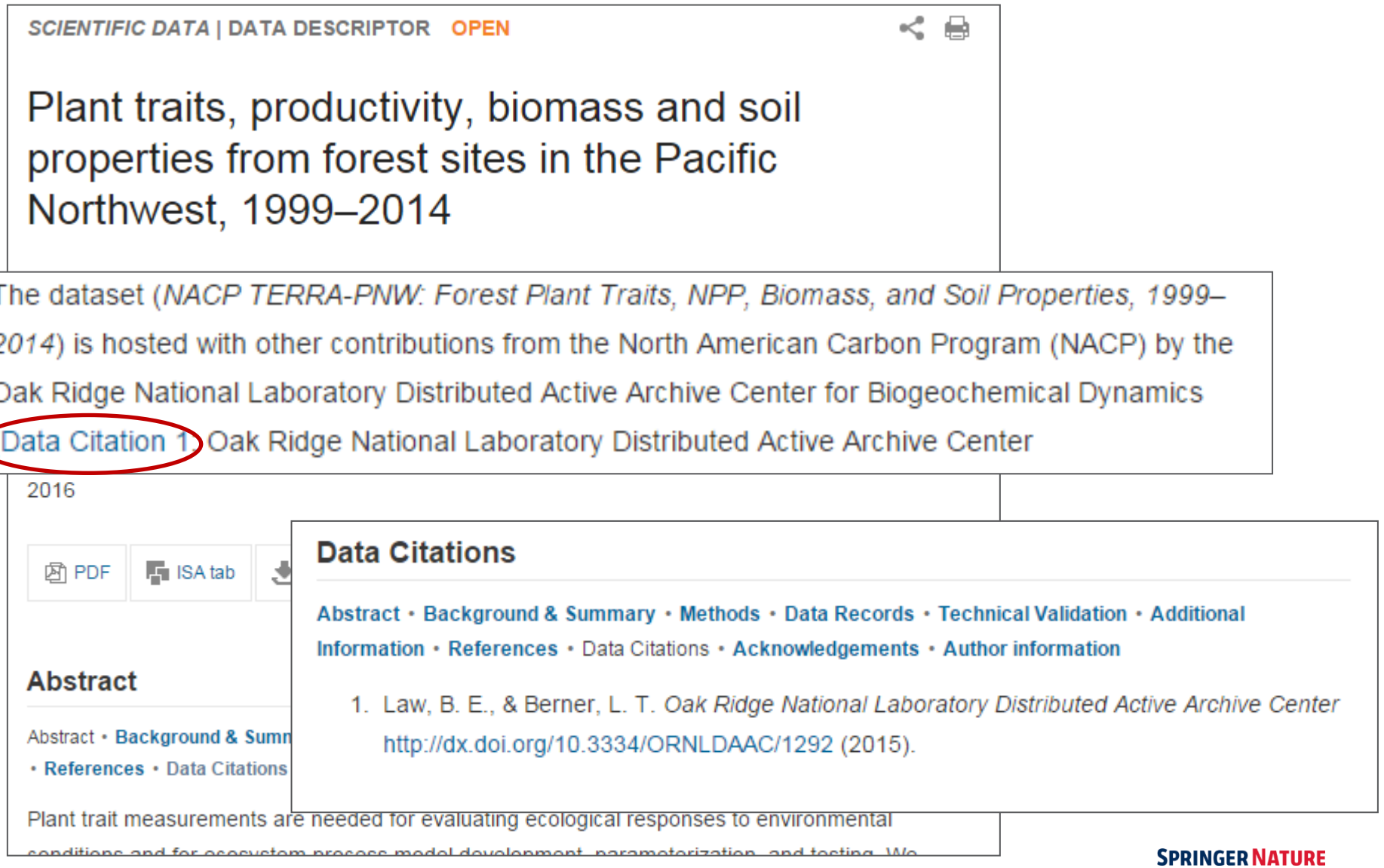
Complete image data sets and corresponding metadata supporting the findings of this study[49] are available from http://purl.stanford.edu/wv179nv3100. **in article**

http://www.nature.com/nphys/journal/vaop/ncurrent/full/nphys3779.html#ref49

Scientific Data (beta)



In-article data citation



SCIENTIFIC DATA | DATA DESCRIPTOR OPEN

Plant traits, productivity, biomass and soil properties from forest sites in the Pacific Northwest, 1999–2014

The dataset (*NACP TERRA-PNW: Forest Plant Traits, NPP, Biomass, and Soil Properties, 1999–2014*) is hosted with other contributions from the North American Carbon Program (NACP) by the Oak Ridge National Laboratory Distributed Active Archive Center for Biogeochemical Dynamics (Data Citation 1). Oak Ridge National Laboratory Distributed Active Archive Center

2016

PDF ISA tab

Abstract

Abstract • Background & Summ • References • Data Citations

Plant trait measurements are needed for evaluating ecological responses to environmental

Data Citations

Abstract • Background & Summary • Methods • Data Records • Technical Validation • Additional Information • References • Data Citations • Acknowledgements • Author information

1. Law, B. E., & Berner, L. T. *Oak Ridge National Laboratory Distributed Active Archive Center* http://dx.doi.org/10.3334/ORNLDAAC/1292 (2015).

Scientific Data's Repository List

Browse our recommended data repositories online.

- *We currently list >80 repositories, across biological, medical, physical and social sciences*
- *When required, we provide guidance to authors on the best place to store their data*
- *Institutional and project specific repositories are supported*

www.nature.com/sdata/data-policies/repositories

Future directions and possibilities

- A standardised research data policy for every journal at Springer Nature
  - All policy types includes data citations in reference lists http://www.springernature.com/gp/group/data-policy
  - Launched 5th July: http://blogs.nature.com/ofschemesandmemes/2016/07/05/promoting-research-data-sharing-at-springer-nature
- More consistent links between data and articles
- Progress on challenging data types (clinical/sensitive, large datasets)
- Wider availability of integrated repositories e.g. figshare
- More discoverable and reusable data supporting publications (better metadata for data and SI)
- Datasets and data citations better captured in manuscript tracking system

Thank you

Erika Pastrana, PhD
Editorial Team Manager
Nature Communications
Springer Nature
e.pastrana@us.nature.com

Iain Hrynaszkiewicz
Head of Data Publishing
Open Research Group
Springer Nature

iain.hrynaszkiewicz@nature.com

SPRINGER NATURE