

SCIENTIFIC DATA

Editorial: What's in a name?

OPEN

Shakespeare's Juliet famously did not consider names as being of any importance. Had she ever needed to track data back to their source, however, she may have felt very differently.

When a scholarly publication refers to a dataset hosted at a public repository, it is best practice to acknowledge the hosting repository by including its name and other key information in a formal data citation (<https://go.nature.com/2qBoCU6>). During our standard editorial checks of data citations prior to publication, we routinely encounter inconsistencies in repository naming. We strive to fix these before publication, but the official name of a data repository is not always as clear as might be expected. Naming inconsistencies also pose a challenge as we expand and maintain our list of recommended data repositories (<http://go.nature.com/2eLHBFP>).

Published: 22 May 2018

Today, we are calling on the repositories with whom we work to check that their repository's name is consistently recorded on their webpages, and in any indices that track their repository or datasets. Repository names should be consistently reflected on individual data records, along with other citation information, in both human readable and machine accessible formats. A roadmap describing how repositories can best support data citation is currently open for public comment¹. For repositories that assign DataCite DOIs to their archived data², the repository name should be recorded in the 'publisher' field of each DOI they register, in exactly the form that it should be included in formal data citations.

Repository names should also be consistently recorded in any lists, indexing services or databases that track the repository or its contents. Repositories included on our recommended list should check that we have listed the name they want to be included in data citations at the journal, and that this matches exactly the name they have registered with DataCite, if relevant.

We encourage repository managers to create and maintain a record for their resource at FAIRsharing (<https://fairsharing.org/>)³ and re3data (<https://re3data.org/>)⁴, two repository indexing services that provide central locations for researchers to search for relevant repositories, and access key metadata about the resource. The curators working at FAIRsharing additionally index standards and data sharing policies, and interconnect these with relevant data repositories. Several publishers and journals maintain collections of the standards and repositories they endorse at FAIRsharing (view ours at <https://fairsharing.org/recommendation/ScientificData>). These endorsements can be viewed on the individual standard and repository record pages, and can serve as a rough measure – independent of literature citations – of how widely adopted or used a particular standard or resource may be.

Scientific Data authors (and ideally anyone formally citing data), should ensure that they use the repository name that has been registered in the DataCite metadata for the DOI they are citing. If the dataset does not have a DOI, authors may refer to FAIRsharing or re3data for the repository name. Authors should contact the host repository if they encounter any inconsistencies between these records.

Ensuring consistency in data citations will help to increase the stability of research data infrastructure, leading to more reliable data for all stakeholders. We think even Juliet would have appreciated the importance of repository name consistency between repositories, publications, and repository registries for data citations!

(Not up on your Shakespeare? see ref. 5).

References

1. Fenner, M. *et al.* A Data Citation Roadmap for Scholarly Data Repositories. Preprint at <https://doi.org/10.1101/097196> (2017).
2. Neumann, J. & Brase, J. DataCite and DOI names for research data. *J. Comput. Aided Mol. Des.* **28**, 1035–1041 (2014).

3. Sansone, S.-A. *et al.* FAIRsharing: working with and for the community to describe and link data standards, repositories and policies. Preprint at <https://doi.org/10.1101/245183> (2018).
4. Pampel, H. *et al.* Making Research Data Repositories Visible: The re3data.org Registry. *PLoS ONE* **8**, e78080 (2013).
5. A rose by any other name would smell as sweet. *Wikipedia* https://en.wikipedia.org/wiki/A_rose_by_any_other_name_would_smell_as_sweet.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>