

# SCIENTIFIC DATA

## OPEN Five years of *Scientific Data*

### EDITORIAL

*Scientific Data* published its first batch of papers five years ago this week. Here, we reflect on our progress and thank all those that have helped us along the way.

Published online: 28 May 2019

Over the last five years, funders and journals around the globe, and across a wide range of disciplines, have adopted new policies that better recognize and promote data sharing. Researchers themselves have rallied behind the FAIR Data Principles<sup>1</sup>, hosting workshops and training events that aim to make research data more findable, accessible, interoperable and reusable.

There have, of course, also been differences of opinion and obstacles. Researchers who rely heavily on the data of others were accused of being ‘research parasites’<sup>2</sup>, a label that some chose to embrace and transform into a ground-breaking award that celebrates innovative data reuse<sup>3</sup>. Other researchers have had to contend with shifting public views caused by high-profile data scandals and an increasingly divisive political climate.

Against this backdrop, *Scientific Data* has sought to be an advocate not just for open data sharing, but also for responsible and effective data sharing. Five years since our launch, over 750 data descriptors have been published at the journal, releasing and describing datasets across a wide range of fields and topics. We are delighted to see that our papers have been collectively cited by more than 6000 other scholarly papers, many of which are themselves compelling examples of data reuse.

We would like to thank all of our authors who over these last five years have supported *Scientific Data*. We also extend our sincerest thanks to our many dedicated and hardworking Editorial Board members and peer reviewers, along with the members of our new Senior Editorial Board (<https://go.nature.com/2XFatD6>). The journal never would have made it this far without all of your hard work and support.

Equally important has been the support of the many data repositories with whom we work. There are now over 100 repositories on our recommended list (<http://go.nature.com/2eLHBFP>), and we work with an even wider range of institutional and project-specific repositories. Because of the diversity of these systems and their policies, our staff often correspond extensively with submitting authors and the hosting repository to find the right way to host our authors’ data and to share it securely with our referees. We would like to extend our thanks to the many repository managers and curators who have borne with us when things have become complicated, which does happen, especially for the complex datasets that are common at the journal.

Many types of data submitted to the journal do not have a dedicated specialist data repository. For authors of such datasets, we have built strong partnerships with a number of ‘generalist’ data repositories, most notably Figshare (<http://figshare.com>) and Dryad (<https://datadryad.org/>). About a third of the journal’s publications use one of these repositories to help host at least part of their data.

We would also like to extend a very special thanks to the researchers behind ISA-tools (<http://isa-tools.org/>) and FAIRsharing<sup>4</sup>. Both projects are led by the group of Susanna-Assunta Sansone, *Scientific Data*’s Honorary Academic Editor. FAIRsharing (<https://fairsharing.org/>) is a curated portal that tracks and interlinks community reporting standards, databases, repositories and data policies. As a result of a long-standing collaboration, users can browse our recommended repositories through a dedicated collection (<https://fairsharing.org/recommendation/ScientificData>), and use FAIRsharing’s tools to browse key information and discover related reporting standards. *Scientific Data* uses the ISA framework as an integral part of our unique metadata curation process<sup>5</sup>.

Beyond the journals’ own pages, *Scientific Data* has led or been a part of a number of initiatives across the wider publishing and research community that promote best practice on topics like sharing clinical research data<sup>6,7</sup>, data citation<sup>8</sup> and research data policies<sup>9</sup>. The journal has also forged partnerships with other publications to share best practice and enable the publication of more reproducible research, most notably journals in the Nature Research family<sup>10</sup>.

The journal also continues to engage with the research community in creative ways, and our conference Better Science through Better Data, will be holding its sixth event in the fall of 2019.

To celebrate our achievements on our 5<sup>th</sup> anniversary, we have launched a special web page with an interactive history of the journal’s most important milestones. We invite you to explore and share it (<https://www.nature.com/sdata/5th-anniversary>).

Here’s to the next five exciting years of data.

## References

1. Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **3**, 160018, <https://doi.org/10.1038/sdata.2016.18> (2016).
2. Longo, D. L. & Drazen, J. M. Data Sharing. *N Engl J Med* **374**, 276–277, <https://doi.org/10.1056/NEJMe1516564> (2016).
3. Greene, C. S., Garmire, L. X., Gilbert, J. A., Ritchie, M. D. & Hunter, L. H. Celebrating parasites. *Nat. Genet* **49**, 483–484, <https://doi.org/10.1038/ng.3830> (2017).
4. Sansone, S.-A. *et al.* FAIRsharing as a community approach to standards, repositories and policies. *Nat. Biotechnol* **7**, 358–367, <https://doi.org/10.1038/s41587-019-0080-8> (2019).
5. Open data, open curation. *Sci. Data* **5**, 180204, <https://doi.org/10.1038/sdata.2018.204> (2018).
6. Hrynaskiewicz, I., Khodiyar, V., Hufton, A. L. & Sansone, S.-A. Publishing descriptions of non-public clinical datasets: proposed guidance for researchers, repositories, editors and funding organisations. *Res. Integr. Peer Rev* **1**, 6, <https://doi.org/10.1186/s41073-016-0015-6> (2016).
7. Let's be pragmatic about clinical data. *Sci. Data* **2**, 150034, <https://doi.org/10.1038/sdata.2015.34> (2015).
8. Data citation needed. *Sci. Data* **6**, 27, <https://doi.org/10.1038/s41597-019-0026-5> (2019).
9. Jones, L., Grant, R. & Hrynaskiewicz, I. Implementing publisher policies that inform, support and encourage authors to share data: two case studies. *Insights* **32**, 11, <https://doi.org/10.1629/uksg.463> (2019).
10. Data-access practices strengthened. *Nature* **515**, 312, <https://doi.org/10.1038/515312a> (2014).



**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/>.

© Springer Nature Limited 2019