

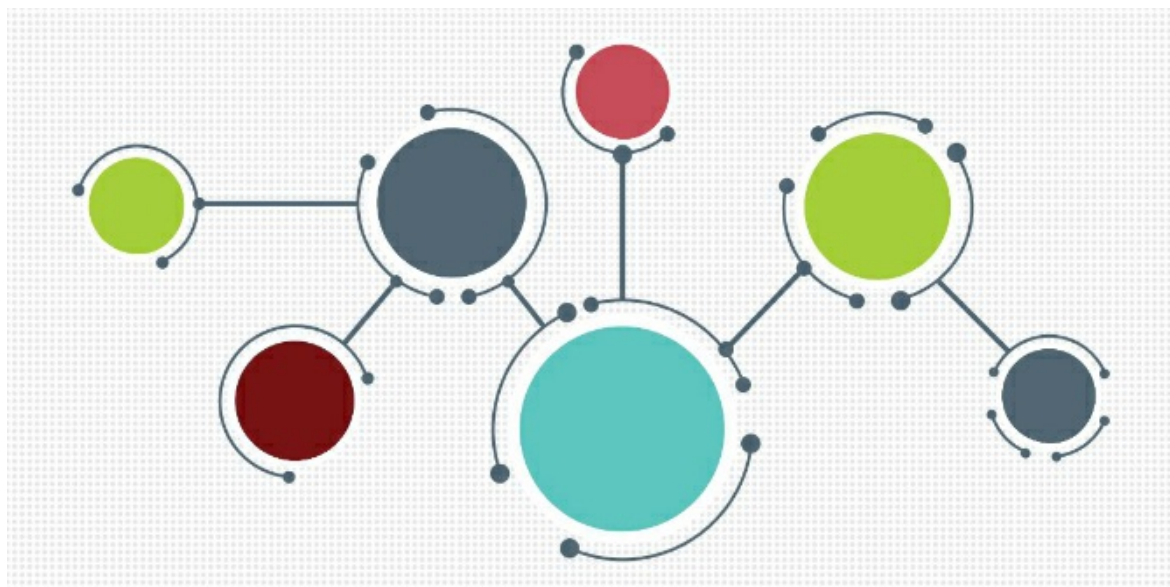
Collaboration and concerted action are key to making open data a reality



*The case for open data is increasingly inarguable. Improved data practice can help to address concerns about reproducibility and research integrity, reducing fraud and improving patient outcomes, for example. Research also shows good data practice can lead to improved productivity and increased citations. However, as **Grace Baynes** reports, recent survey data shows that while the research community recognises the value of open data, uptake remains slow, with good data practice and data sharing far from the status quo. To effect change, government, funders, institutions, publishers, and researchers themselves all have an important role to play.*

The case for good research practice and open data to research outputs is increasingly inarguable. Open access to research data can help speed the pace of advancing discovery and deliver more value by enabling reuse and reducing duplication. Good data practice also makes research more efficient, effective and fulfilling for researchers. As the data in the [Digital Science Open Data survey 2017](#) reveals, the research community recognises the value of open data, yet good data practice and data sharing are still far from the status quo.

Springer Nature and its publications have been advocating for good data practice for over a decade. Recent efforts have focused on growing data publishing options to provide credit, and strengthening and simplifying our data policies. Our future focus is on support and incentives to enable data sharing, data management and open data, built in collaboration with the research community.



The case for data

The argument for better data practice is made stronger by global concerns about reproducibility and research integrity, reducing fraud and improving patient outcomes. As much as 50% of preclinical research done in the US, at a cost of US\$56.4b a year, cannot be reproduced, estimates [a 2015 study](#). In the same year, a [Nature survey](#) found that 70% of over 1,500 respondents had tried and failed to replicate the work of others. More shocking was that 50% of respondents had failed to reproduce their own work. There is evidence that data availability increases reproducibility, as reported in a [review of Nature Genetics papers](#) and elsewhere.

There is also a proven productivity benefit to good data practice. Data archiving can double the publication output of research projects, according to [a study of 7,000 National Science Foundation and National Institutes of Health-funded research projects in social sciences](#). Citation impact of research papers has also been shown to increase when data is made available – by as much as [50% in astrophysics](#), and between 9-35% in [gene expression microarrays](#), [astronomy](#), and [paleoceanography](#).

The data in this survey shows that researchers are using others' research data (49%), or would be willing to do so (80%). Yet only 60% of respondents make their data openly available "frequently" or "sometimes". The most common ways of sharing data are still supplementary information in a journal article or peer-to-peer. Perhaps more concerning is data storage and data management. Only 20% of respondents had prepared a data management plan, and the most common ways to store active and archived data were personal hard drives, external hard drives, and institutional servers.

Researchers are intelligent, responsible, motivated people. They are also time-poor, and do not necessarily want to become data or licensing experts. So they need clear information, simple policies and advice. They also understandably prioritise advancing their field, their own research, and building their careers. So they need tools to make data sharing and management easier, and credit and incentives to make good research data practice and open data worthwhile.

To effect change, government, funders, institutions, libraries, publishers, and researchers themselves all have a role to play. Here are areas this survey has prompted us to think more about:

The role of government

It is interesting to see the support for national mandates for open data in this survey (55% of respondents). Many countries have now made government data open, providing the best use cases to date for economic and social impact of open data. When it comes to research data, national approaches and infrastructures will continue to need similar long-term commitment, and to be balanced with fostering international collaboration, including through global discipline-specific data repositories.

The role of the funder

The results of this survey would suggest that funder mandates are not a key motivator for open data. This contradicts the findings of [other studies](#), and is contrary to what we see as funders' crucial role in effecting change. The growth of open access publishing was driven in part by funders issuing clear and specific mandates, explicitly making funds available and making compliance a requirement. Springer Nature tracks funder policies on data to help provide advice to authors on compliance. Encouragingly, more than 50 funders now mandate or encourage data sharing, compared to 28 in 2015. As yet, only a few funders have requirements for data management plans or data availability statements, or explicitly make funding available for data management, storage, and curation.

The role of the institution

Institutions and libraries have a key role to play in supporting researchers: helping them understand and comply with funder requirements, training, and establishing local [research data management solutions and support](#) where needed. Partnering with data initiatives, repositories, and other useful parties, including publishers, will help reduce potential duplication of effort and ensure sustainability.

The role of the publisher

Publishers work closely with researchers at many stages of the research process, particularly when they are writing up and sharing their findings. Here are five actions publishers can take:

1. Continue to advocate for good data practice across different communities.
2. Encourage good research data practice and open data through journal policies and author information: see,

for example, [Springer Nature's standardised research data policies](#), Research Data Support Helpdesk, and recommended repositories list.

3. Provide credit mechanisms for good data management and open data: through data publishing, registered reports, data citation and linking, and new mechanisms such as badges for open practices.
4. Offer solutions to help researchers share their own data, and discover and use data: for example, our pilot [Data Support Services](#), which help researchers deposit and curate data in partnership with Figshare.
5. Partner with the research community to build shared solutions: for example, the global Research Data Alliance (RDA) interest group to improve [research data policy standards](#), data linking and citation.

A number of other publishers including PLOS, Wiley, and Elsevier are also taking some or all of these steps.

Concerted efforts by governments, funders, research institutions, publishers and researchers themselves are needed to make widespread open data a reality, and make research data management the new normal. Collaboration and partnerships between these groups will make that happen faster, and more effectively. Springer Nature looks forward to further playing its part.

This post originally appeared as part of Digital Science's "[The State of Open Data Report 2017](#)", and is published under a [CC BY 4.0](#) license. The full report can be found on [Figshare](#).

Note: This article gives the views of the author, and not the position of the LSE Impact Blog, nor of the London School of Economics. Please review our [comments policy](#) if you have any concerns on posting a comment below.

About the author

Grace Baynes is Director of Data & New Product Development for Open Research at Springer Nature. She is responsible for promoting open data and good research data practice; data publishing, including the journal *Scientific Data*; data services; and new product development across open science and open research. Grace's passion for open science dates back to joining open access publisher BioMed Central in 2003, and has flourished over 14 years at BMC, Nature Publishing Group, and now Springer Nature. Her ORCID iD is [0000-0002-4933-3186](#).