

Growth in policies for data sharing



80+

Funders internationally now have policies

54

that require

31

or encourage data sharing

Most of these policies have been announced in the past 5 years.



29

funders - including the European Commission, National Institutes of Health (NIH) and the Wellcome Trust - cover the costs of data sharing and data management plans.¹

Growth in data sharing



72%

of researchers are curating their data for sharing.²

64%

The number of researchers who routinely share their data has also grown year on year, to 64% in 2018.²

The case for data sharing



There is evidence that good data practice leads to:

- Increased citation³
- Increased reproducibility (quality/robustness) of research⁴
- Increased productivity⁵
- Reduced harm and costs of biased/non-transparent research⁶
- Support for career progression⁷
- Better return on investment in research funding⁸

Researchers are increasingly aware of open data and are actively re-using open datasets for research



80%

of researchers are aware of open data available to reuse².

79%

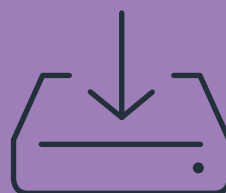
of these researchers are now willing to reuse open data sets in their own research².

Researchers are positive about data sharing



of researchers rate the importance of making their data discoverable highly – with an average rating of 7.3 out of 10⁹

The majority of data sharing practice is not optimal



- The most common method of data sharing is still as supplementary information in a journal article²
- The most common ways to store data are on personal hard drives, external hard drives and institutional servers²
- 30% of researchers have lost research data²
- 60% of researchers have rarely, or never, created a data management plan²

Top barriers to sharing



Researchers are hampered by time constraints and a lack of knowledge around data sharing. Organizing data in a presentable and useful way was the biggest challenge to data sharing identified (46%). 37% were unsure about copyright and licensing. 33% did not know which repository to use. Time was the biggest factor for 26% and cost for 19%⁹.

SOURCES AND REFERENCES

1. Internal analysis, global funder policies
2. Digital Science, *The State of Open Data Report 2018* (2018). doi: <https://doi.org/10.6084/m9.figshare.7195058.v1>
3. Henneken & Accomazzi (2011) <https://arxiv.org/abs/1111.3618>; Belter (2014) <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0092590>; Piwowar & Vision (2013) <https://doi.org/10.7717/peerj.175>; Piwowar et al (2007) <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0000308>; Benevise et al (2016) <http://link.springer.com/article/10.1007%2Fs11192-016-1868-7>; Leitner et al (2016) <http://journal.frontiersin.org/article/10.3389/fnins.2016.00419/full>
4. <http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002165>
5. <http://www.ebi.ac.uk/about/news/press-releases/value-and-impact-of-the-european-bioinformatics-institute>
6. Eyding et al (2010) <http://www.bmj.com/content/341/bmj.c4737>; Cochrane Database Syst Rev. 2012 DOI: 10.1002/14651858.CD008965.pub3
7. McKiernan et al (2016) <http://dx.doi.org/10.7554/eLife.16800>
8. <http://www.nature.com/news/economic-return-from-human-genome-project-grows-1.13187>
9. Stuart, D; Baynes, G; Hrynaszkiwicz, I; Allin, K; Penny, D; Lucraft, M; et al. (2018): *Whitepaper: Practical challenges for researchers in data sharing* <https://doi.org/10.6084/m9.figshare.5975011.v1>