Bibliography on data sharing

Created: 2020-05-15

This is a bibliography of articles related to data sharing. I am developing this as part of an effort to write a research article and a research grant. Some of the articles may or may not be appropriate for the bibliographies of the grant or article, but it is better to have too many references and cull the ones that you don’t like than having to find some after the fact.

### References

Agency for Healthcare Research and Quality. Data Management Plan Policy. NOT-HS-20-011. 2020-05-26. Available in [html format](https://grants.nih.gov/grants/guide/notice-files/NOT-HS-20-011.html).

Allison DB, Brown AW, George BJ, Kaiser KA. Reproducibility: A tragedy of errors. Nature 2016-02-03. 530(7588); 27-29. Availabe in [html format](https://www.nature.com/news/reproducibility-a-tragedy-of-errors-1.19264) or [pdf format](https://www.nature.com/news/polopoly_fs/1.19264!/menu/main/topColumns/topLeftColumn/pdf/530027a.pdf).

Baggerly KA, Coombes KR. Deriving Chemosensitivty from cell lines: Forensic bioinformatics and reproducible research in high-throughput biology. The Annals of Applied Statistics, 2009; 3(4): 1309–1334. Available in [pdf format](https://arxiv.org/pdf/1010.1092.pdf).

Baggerly KA, Morris JS, Coombes KR. Reproducibility of SELDI-TOF protein patterns in serum: comparing datasets from different experiments. Bioinformatics. 2004;20(5):777‐785. <doi:10.1093/bioinformatics/btg484>.

Baggerly KA, Morris JS, Edmonson SR, Coombes KR. Signal in noise: evaluating reported reproducibility of serum proteomic tests for ovarian cancer. J Natl Cancer Inst. 2005;97(4):307‐309. <doi:10.1093/jnci/dji00>. Available in [html format](https://academic.oup.com/jnci/article/97/4/307/2544099) or [pdf format](https://academic.oup.com/jnci/article-pdf/97/4/307/9644348/dji008.pdf).

Bill and Melinda Gates Foundation. Bill and Melinda Gates Foundation Open Access Policy [Internet]. Available in [html format](https://www.gatesfoundation.org/How-We-Work/General-Information/Open-Access-Policy).

Bollen K, Cacioppo JT, Kaplan R, Krosnick J, Olds JL. Social, Behavioral, and Economic Sciences Perspectives on Robust and Reliable Science (National Science Foundation, Arlington, VA, 2015). Available in [pdf format](https://www.nsf.gov/sbe/AC_Materials/SBE_Robust_and_Reliable_Research_Report.pdf).

Brazma A, Hingamp P, Quackenbush J et al. Minimum information about a microarray experiment (MIAME)—toward standards for microarray data. Nat Genet 29, 365–371 (2001). doi 10.1038/ng1201-365. Available in [html format](https://www.nature.com/articles/ng1201-365) and [pdf format](https://www.nature.com/articles/ng1201-365.pdf).

Collins FS, Tabak LA. Policy: NIH plans to enhance reproducibility. Nature, 2014-01-27, 505(7485). Available in [html format](https://www.nature.com/news/policy-nih-plans-to-enhance-reproducibility-1.14586).

Davison A. Automated capture of experiment context for easier reproducibility in computational research. Computing in Science and Engineering 14(4): 48-56, July-Aug. 2012. 10.1109/MCSE.2012.41. Article is behind a pay wall.

Desai T, Ritchie F, Welpton R. The Five Safes: Designing Data Access for Research. Economic Working Paper Series 1601, University of the West of England. Available in [pdf format](http://www2.uwe.ac.uk/faculties/BBS/Documents/1601.pdf).

Elsevier. Research Data [Internet]. 2017. Available in [html format](https://www.elsevier.com/about/our-business/policies/research-data).

Fane B, Ayris P, Hahnel M, Hrynaszkiewicz I, Baynes G, Farrell E. The State of Open Data Report 2019. Digital Science. Available in [pdf format](https://digitalscience.figshare.com/ndownloader/files/18090485).

Federer LM, Belter CW, Joubert DJ, Livinski A, Lu Y-L, Snyders LN, et al. (2018) Data sharing in PLOS ONE: An analysis of Data Availability Statements. PLoS ONE 13(5): e0194768. 10.1371/journal.pone.0194768. Available in [html format](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0194768) or [pdf format](https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0194768&type=printable).

Fladie IA, Adewumi TM, Vo NH, Tritz DJ, Vassar MB. An Evaluation of Nephrology Literature for Transparency and Reproducibility Indicators: Cross-Sectional Review. Kidney Int Rep. 2019;5(2):173‐181. Published 2019 Nov 13. <doi:10.1016/j.ekir.2019.11.001>. Available in [html format](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7000850/) and [pdf format](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7000850/pdf/main.pdf).

Goodman SN, Fanelli D, Ioannidis JPA. What does research reproducibility mean? Science Translational Medicine 2016-06-01, 8(341), 341ps12. DOI: 10.1126/scitranslmed.aaf5027. Available in [html format](https://stm.sciencemag.org/content/8/341/341ps12/tab-article-info) or [pdf format](https://stm.sciencemag.org/content/scitransmed/8/341/341ps12.full.pdf).

Greenberg J, White HC, Carrier S, Scherle R. A Metadata Best Practice for a Scientific Data Repository, Journal of Library Metadata. 2009, 9:3-4, 194-212, DOI: 10.1080/19386380903405090. Available in [html format](https://www.tandfonline.com/doi/full/10.1080/19386380903405090).

Hagedorn G, Mietchen D, Morris RA, et al. Creative Commons licenses and the non-commercial condition: Implications for the re-use of biodiversity information. Zookeys. 2011;(150):127‐149. <doi:10.3897/zookeys.150.218>. Available in [html format](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3234435/) and [pdf format](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3234435/pdf/ZooKeys-150-127.pdf).

Ioannidis, J., Allison, D., Ball, C. et al. Repeatability of published microarray gene expression analyses. Nat Genet 41, 149–155 (2009). <https://doi.org/10.1038/ng.295>. Available in [html format](https://www.nature.com/articles/ng.295).

Kolata G. How Bright Promise in Cancer Testing Fell Apart. The New York Times, 2011 Jul 07. Available in [html format](https://www.nytimes.com/2011/07/08/health/research/08genes.html).

Longo D, Drazen DM. Data Sharing. N Engl J Med 2016; 374:276-277 DOI: 10.1056/NEJMe1516564. Available in [html format](https://www.nejm.org/doi/full/10.1056/NEJMe1516564) or [pdf format](https://www.nejm.org/doi/pdf/10.1056/NEJMe1516564?articleTools=true).

National Institutes of Health. NIH Sharing Policies and Related Guidance on NIH-Funded Research Resources [Internet]. 2018-02-07. Available in [html format](https://grants.nih.gov/policy/sharing.htm)

National Institutes of Health. NIH Strategic Plan for Data Science. Available in [html format](https://datascience.nih.gov/sites/default/files/NIH_Strategic_Plan_for_Data_Science_Final_508.pdf).

National Science Foundation. Dissemination and Sharing of Research Results [Internet]. 2010. Available in [html format](http://www.nsf.gov/bfa/dias/policy/dmp.jsp).

National Science Foundation. Digital Research Data Sharing and Management [Internet]. 2011. Available in [pdf format](https://www.nsf.gov/nsb/publications/2011/nsb01211.pdf).

Open Science Collaboration. Estimating the reproducibility of psychological science. Science 2015-08-28, 349(6251), aac4716. DOI: 10.1126/science.aac4716. Available in [html format](https://science.sciencemag.org/content/349/6251/aac4716) or [pdf format](https://science.sciencemag.org/content/349/6251/aac4716/tab-pdf).

Peng RD, Dominici F, Zeger SL. Reproducible Epidemiologic Research, American Journal of Epidemiology, Volume 163, Issue 9, 1 May 2006, Pages 783–789, <https://doi.org/10.1093/aje/kwj093>. Available in [html format](https://academic.oup.com/aje/article/163/9/783/108733).

Pollack A. F.D.A. Says Cancer Test Failed to Get Its Approval. The New York Times, 2008 Oct 08. Available in [html format](https://www.nytimes.com/2008/10/09/business/09cancer.html).

Potti A, Dressman HK, Bild A, et al. Genomic signatures to guide the use of chemotherapeutics. Nature Medicine, 2006; 12: 1294–1300. Available in [html format](https://www.nature.com/articles/nm1491).

Powers SM, Hampton SE. Open science, reproducibility, and transparency in ecology. Ecol Appl. 2019;29(1):e01822. <doi:10.1002/eap.1822>. Available in [pdf format](https://esajournals.onlinelibrary.wiley.com/doi/epdf/10.1002/eap.1822).

Scherle R, Carrier S, Greenberg J, Lapp H, Thompson A, Vision T and White H. “Building support for a discipline-based data repository”. In Proceedings of the 2008 International Conference on Open Repositories. Available in [pdf format](http://pubs.or08.ecs.soton.ac.uk/35/1/submission_177.pdf).

Science. Science: editorial policies [Internet]. 2017. Available in [html format](http://www.sciencemag.org/authors/science-editorial-policies).

Silva L. PLOS’ New Data Policy: Public Access to Data. In: EveryONE: PLOS ONE Community Blog [Internet]. 2014. Available in [html format](http://blogs.plos.org/everyone/2014/02/24/plos-new-data-policy-public-access-data-2/).

Singer N. Duke Suspends Researcher and Halts Cancer Studies. The New York Times Prescription blog, 2010 Jul 20. Available in [html format](https://prescriptions.blogs.nytimes.com/2010/07/20/duke-suspends-researcher-halts-cancer-studies/).

Starr J, Castro E, Crosas M, Dumontier M, Downs RR, Duerr R, Haak LL, Haendel M, Herman I, Hodson S, Hourclé J, Kratz JE, Lin J, Nielsen LH, Nurnberger A, Proell S, Rauber A, Sacchi S, Smith A, Taylor M, Clark T. 2015. Achieving human and machine accessibility of cited data in scholarly publications. PeerJ Computer Science 1:e1 Available in [html format](https://doi.org/10.7717/peerj-cs.1) or [pdf format](https://peerj.com/articles/cs-1.pdf).

Vickers, A.J. Sharing raw data from clinical trials: what progress since we first asked “Whose data set is it anyway?”. Trials 17, 227 (2016). doi: 10.1186/s13063-016-1369-2. Available in [html format](https://trialsjournal.biomedcentral.com/articles/10.1186/s13063-016-1369-2) or [pdf format](https://trialsjournal.biomedcentral.com/track/pdf/10.1186/s13063-016-1369-2).

Welcome Trust. Sharing research data and findings relevant to the novel coronavirus (COVID-19) outbreak. Wellcome Foundation press release, 2020-01-31. Available in [html format](https://wellcome.ac.uk/coronavirus-covid-19/open-data).

Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR Guiding Principles for scientific data management and stewardship [published correction appears in Sci Data. 2019 Mar 19;6(1):6]. Sci Data. 2016;3:160018. Published 2016 Mar 15. <doi:10.1038/sdata.2016.18>. Available in [html format](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4792175/) or [pdf format](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4792175/pdf/sdata201618.pdf).

Collins FS, Tabak LA. Policy: NIH plans to enhance reproducibility. Nature, 2014-01-27, 505(7485). Available in [html format](https://www.nature.com/news/policy-nih-plans-to-enhance-reproducibility-1.14586).

NIH Strategic Plan for Data Science. Available in [html format](https://datascience.nih.gov/sites/default/files/NIH_Strategic_Plan_for_Data_Science_Final_508.pdf).

van Panhuis, W.G., Paul, P., Emerson, C. et al. A systematic review of barriers to data sharing in public health. BMC Public Health 14, 1144 (2014). <https://doi.org/10.1186/1471-2458-14-1144>. Available in [html format](https://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-14-1144) or [pdf format](https://bmcpublichealth.biomedcentral.com/track/pdf/10.1186/1471-2458-14-1144).

Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR Guiding Principles for scientific data management and stewardship [published correction appears in Sci Data. 2019 Mar 19;6(1):6]. Sci Data. 2016;3:160018. Published 2016 Mar 15. <doi:10.1038/sdata.2016.18>. Available in [html format](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4792175/) or [pdf format](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4792175/pdf/sdata201618.pdf).

Ziemann M, Eren Y, El-Osta A. Gene name errors are widespread in the scientific literature. Genome Biology (2016) 17:177 DOI 10.1186/s13059-016-1044-7.

### Where to find this document and related documents.

* Bibliography in [html format](http://www.pmean.com/post/r25-bibliography.html) or in [R Markdown code](https://github.com/pmean/peer-reviewed-data-sets/blob/master/src/r25-bibliography.Rmd).
* R25 grant draft in [html format](http://www.pmean.com/post/r25-grant.html) or in [R Markdown code](https://github.com/pmean/peer-reviewed-data-sets/blob/master/src/r25-grant.Rmd).
* An earlier brief overview of the R25 grant in [html format](http://www.pmean.com/post/training-module-summary.html) or in [R Markdown code](https://github.com/pmean/peer-reviewed-data-sets/blob/master/src/training-module-summary.Rmd).
* Draft publication on figshare in [html format](http://www.pmean.com/post/data-sharing-publication.html) or in [R Markdown code](https://github.com/pmean/peer-reviewed-data-sets/blob/master/src/data-sharing-publication.Rmd).
* R25 specific aims in [html format](http://www.pmean.com/post/r25-specific-aims.html) or [R Markdown code](https://github.com/pmean/peer-reviewed-data-sets/blob/master/src/r25-specific-aims.Rmd).