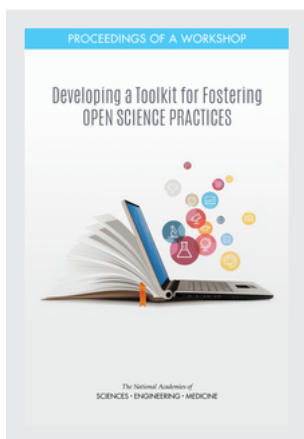


This PDF is available at <http://nap.nationalacademies.org/26308>



Developing a Toolkit for Fostering Open Science Practices: Proceedings of a Workshop (2021)

DETAILS

120 pages | 6 x 9 | PAPERBACK

ISBN 978-0-309-09361-3 | DOI 10.17226/26308

CONTRIBUTORS

Thomas Arrison, Jennifer Saunders, and Emi Kameyama, Rapporteurs; Committee on Developing a Toolkit for Fostering Open Science Practices: A Workshop; Board on Research Data and Information; Policy and Global Affairs; National Academies of Sciences, Engineering, and Medicine

SUGGESTED CITATION

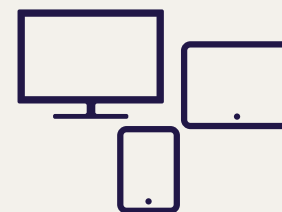
National Academies of Sciences, Engineering, and Medicine. 2021. *Developing a Toolkit for Fostering Open Science Practices: Proceedings of a Workshop*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/26308>.

BUY THIS BOOK

FIND RELATED TITLES

Visit the National Academies Press at nap.edu and login or register to get:

- Access to free PDF downloads of thousands of publications
- 10% off the price of print publications
- Email or social media notifications of new titles related to your interests
- Special offers and discounts



All downloadable National Academies titles are free to be used for personal and/or non-commercial academic use. Users may also freely post links to our titles on this website; non-commercial academic users are encouraged to link to the version on this website rather than distribute a downloaded PDF to ensure that all users are accessing the latest authoritative version of the work. All other uses require written permission. ([Request Permission](#))

This PDF is protected by copyright and owned by the National Academy of Sciences; unless otherwise indicated, the National Academy of Sciences retains copyright to all materials in this PDF with all rights reserved.

Appendix C

Toolkit Elements

This appendix includes examples of draft elements of a toolkit that have been developed by members of working groups of the National Academies of Sciences, Engineering, and Medicine’s Roundtable on Aligning Incentives for Open Science. The following materials were developed to stimulate discussions at the November 5, 2020, workshop on Developing a Toolkit for Fostering Open Science Practices:

- I. **Open Science Imperative.** This essay communicates the benefits of open science using approachable language.
- II. **Open Science Signaling Language Template and Rubric.** These resources provide specific language that can be adapted and adopted to signal an organization’s interest in open science activities at specific points of high leverage (e.g., grant applications, job postings).
- III. **Good Practices Primers.** These concise guides offer policy makers a high-level overview of open sharing.
- IV. **Open Science by the Numbers Infographic.** This infographic communicates the benefits of open science in a graphic form.
- V. **Open Science Success Stories Database.** This database compiles research articles, perspectives, case studies, news stories, and other materials that demonstrate the myriad ways in which open science benefits researchers and society alike.

- VI. **Reimagining Outputs Worksheet.** This table enumerates the range of research products stakeholders may choose to consider as they develop open science policies.

The toolkit is primarily intended to assist university leadership, academic department chairs, research funders, learned societies, and government agencies about how such a toolkit might be used, what additional materials are needed, and how such a toolkit should be disseminated for broad adoption. As a result of the workshop, a few sections in the Open Science Imperative and Good Practices Primers have been revised by the working group authors.

I. OPEN SCIENCE IMPERATIVE¹

Derrick Anderson, Arizona State University
Rachel Bruce, UK Research and Innovation
Ashley Farley, Bill & Melinda Gates Foundation
Robert Hanisch, National Institute of Standards and Technology
Greg Tananbaum, Open Research Funders Group
Thomas Wang, American Heart Association/
University of Texas Southwestern Medical Center

This narrative communicates the benefits of open science using succinct, approachable language. One way to think about its possible deployment is to envision an academic administrator or senior leader at a philanthropy who has a vague notion that open science is something they should better understand. This piece, if successfully executed, will make the affirmative case as to why the open approach to the research endeavor is preferable to the status quo, and what the benefits to society will be if it is adopted at scale.

Over the last 20 years, the research community has grown increasingly interested in and supportive of open science activities. Open science encompasses a range of individual, institutional, and community efforts to broaden access to research outputs. This increased accessibility facilitates better collaboration and outcomes as a function of collective intelligence. By prioritizing shared discovery over individual and institutional agendas, open science practices are spurring the knowledge economy, generating broad social and public benefits, strengthening cultural values for scientific literacy and education, and improving public policy and democracy (Tennant et al., 2016; Zuccala, 2010). Despite the benefits of open science, individual researchers face numerous barriers that are restricting broad uptake of these practices. The current credit and reward systems disincentivize information sharing in favor of siloed, noninclusive modes of knowledge production. Significant, coordinated support within and across research stakeholder groups is necessary to change these incentives to realize the benefits of open science. This white paper, prepared in conjunction with the National Academies of Sciences, Engineering, and Medicine's Roundtable on Align-

¹The views expressed are those of the authors and do not necessarily reflect the official policies or positions of their employing organizations.

ing Incentives for Open Science, briefly sketches the current state of open science, contrasts the diminishing returns of the traditional scientific model with the advantages of emergent open science practices, and suggests possible measures that organizations can individually and collectively undertake to shape the future of research and discovery.

THE STATE OF OPEN SCIENCE

Open science has been conceptualized in philosophical and ideological terms as an affinity for open flows of information to facilitate innovation for the betterment of society (Gold, 2016), but it is most frequently used as an umbrella term to describe active efforts to reduce the barriers to information access for researchers and the public. A commonly used definition of open science is “the idea that scientific knowledge of all kinds should be openly shared as early as is practical in the discovery process” (Nielsen, 2011). Although varying conceptualizations and definitions of open science exist, there is general agreement on the practices that support it, such as open access publication, research preregistration, open access to data and materials, and development of open source software (Berg and Niemeyer, 2018; Gold, 2016; Gold et al., 2019).

Increased adoption of these mutually reinforcing practices by institutions and especially by individual researchers has created a momentum behind open science. This momentum is reflected partly by the choices that researchers make regarding how their data are shared. In one survey, the number of researchers who reported making their data openly available increased from just over 55 percent to 64 percent between 2016 and 2018. From before 1990 through the 2010s, the percentage of researchers who were unaware of the license under which they made their data openly available decreased from 71 percent to 54 percent. During the same time, the percentage of respondents who would feel motivated to make their data openly available for co-author credit increased from 7 percent to 27 percent (Digital Science and Figshare, 2018, 8, 13).

The rise of open access as a widespread publishing practice also indicates greater uptake of open science principles and values. An analysis of 70 million articles published between 1950 and 2019 determined that at least 31 percent of all scholarly publications are available as open access and that the proportion is growing. The same analysis indicated that, given existing trends, 70 percent of all article views will be to open access papers by 2025 (Piwowar et al., 2019). This trend appears to be driven by the values held

by researchers: “Over 90 percent of OA [open access] authors published this way because of the principle of free access” (Swan and Brown, 2004, 5) and because of “their perceptions that these journals reach larger audiences, publish more rapidly and are more prestigious than the toll-access (subscription-based) journals that they have traditionally published in” (Swan and Brown, 2005, ES 1). This momentum toward the open sharing of research papers is further underscored by the spectacular flourishing of preprints, with both readership and authorship growth near 100 percent year-on-year (Abdill and Blekman, 2019).

These data indicate that although open science practices have been adopted by an increasing number of researchers, a large share of researchers remain either unaware of the benefits of these practices or find that the barriers to adoption (including time, resources, lack of clear guidance, and ambiguous incentives) are significant. Enhanced researcher awareness and adoption of open science approaches, combined with proper institutional support and better alignment of credit/reward systems, holds the potential to realize greater knowledge diffusion; improved efficiency, transparency, and interdisciplinarity of scientific exploration; and a more robust, accessible, and replicable body of research (Spellman et al., 2018; Tennant et al., 2016).

THE BENEFITS OF OPEN SCIENCE

Communicating the advantages of open science to researchers and the broader public is essential to greater uptake of these practices. Open science offers an array of benefits across five domains:

1. ***Supporting the growth of the knowledge economy.*** By facilitating freer flows of information among scientists, research institutions, and firms, open science practices can accelerate the discovery process and commercialization of scientific research. The inherently transparent nature of open science also makes testing the reproducibility and replicability of scientific research substantially more efficient.
2. ***Improving the integrity, reliability and transparency of scientific research.*** Science as a process operates with reproducibility as a core objective. Students are trained through replication exercises and scientists are expected to describe their work in ways that facilitate replication. Open science practices make the processes of

science more transparent, which, in turn, makes scientific findings easier to test and to trust.

3. ***Generating social and public benefit.*** By lowering barriers to public participation in science, open science approaches allow social needs articulated by the public to inform a greater share of scientific research and enable citizens to make better-informed decisions.
4. ***Strengthening scientific literacy and education.*** By making scientific research freely available to the public, open science enables nonscientists to become more familiar with scientific methods and encourages greater layperson interest in applying a rigorous, inquisitive approach to their engagement with the world and the pressing issues of the day.
5. ***Improving public policy and democracy.*** By encouraging greater transparency in research and availability of research products, open science allows policy makers and the public to be more informed about research that can be used to shape policy and promote civic action.

Numerous research projects and platforms have realized the benefits of open science approaches, sometimes across all five of these domains, including the following:

- **The Human Genome Project**, completed in 2003, was carried out with an explicit commitment to open science. Participating researchers pledged to make their discoveries available online within 24 hours and provide unrestricted access to information in real time. As a result, the project's public-domain gene sequences generated an estimated 30 percent more genetic diagnostic tests than genes that were first sequenced by private firms and then restricted as intellectual property. The myriad of public and private economic benefits created by the Human Genome Project (estimated at \$965 billion and nearly four million jobs between 1988 and 2012; Tripp and Grueber, 2011) have established it as a model for the effective use of open data, providing a picture of what the future of science and innovation could look like with greater adoption of open science practices (SPARC, n.d.).
- The **Group on Earth Observations (GEO)** is a global network of more than 100 national governments and more than 100 par-

ticipating organizations that enables the collection and sharing of atmospheric, oceanic, and terrestrial data and information to facilitate better decision making and policy formulation. GEO's Global Earth Observation System of Systems (GEOSS) portal was designed according to best practices in open science to facilitate open, coordinated, and sustained data sharing to advance the United Nations 2030 Agenda for Sustainable Development, the Paris Agreement, and the Sendai Framework for Disaster Risk Reduction. In addition to enabling communication between researchers and governments, "data products and information derived from GEO data can be useful for individuals to better understand the environment in which they live and work, and protecting the health of their family, and better educating themselves, and through the positive results of many other generative and even serendipitous applications" (Benkler, 2006; Mayo and Steinberg, 2007; NRC, 2009; and Zittrain, 2006; cited in Uhlir, 2015, 13).

- **The Lab @ DC** is a unit within the Washington, D.C., mayor's administration that works to design public policy and program interventions for the city. The Lab @ DC uses the Open Science Framework to share their methodology, analysis, and evaluations of municipal programs, utilizing transparency to allow their projects to be reproduced and replicated by other community groups. Projects that have been undertaken by this group span from transit, housing, and public safety to customer service and economic prosperity (The Lab @ DC, n.d.).
- **Symbiota** is an exclusively web-based open source content management system that integrates natural history collections and other biological community knowledge and data into a network of databases and tools to increase knowledge of biodiversity. Since 2012, 73 percent of projects funded by the National Science Foundation's Advancing Digitization of Biodiversity Collections have used Symbiota. The platform now hosts 37 million records from 766 universities, museums, and research organizations, including linkages to images, tissues, DNA sequences, and taxonomic and ecological information (Symbiota, n.d.). Importantly, Symbiota's software design philosophy and implementation was driven by its "*user community*—e.g., collections managers, taxonomists, ecologists, data entry personnel, programmers, informaticians, and students" (Gries et al., 2014). Symbiota is freely available to researchers and the public.

- **Global Open Data for Agriculture and Nutrition (GODAN)** is an initiative of the U.S. Department of Agriculture and U.S. Agency for International Development that promotes open data sharing to increase global access to information about agriculture and nutrition. Leveraging data input from a partner network of more than 700 private- and public-sector, nonprofit, and academic organizations, GODAN aims to inform and improve daily decision making for farmers and consumers, with the goal of developing solutions to global hunger (Adams, 2018).
- **Microreact** is a free, real-time tool for visualizing and tracking outbreaks of diseases such as Ebola and Zika, as well as antibiotic-resistant microbes. Developed through a collaboration between researchers from the Wellcome Trust Sanger Institute and Imperial College London, Microreact allows any researcher in the world to upload information on disease outbreaks via its web browser, which can be shared and visualized through Microreact's cloud-based system. Microreact also integrates data submitted for publication in the journal *Microbial Genomics* to encourage greater data availability and access (Wellcome Trust, 2016).
- The **California Policy Lab** is a nonprofit based at the University of California, Los Angeles and Berkeley, that partners with state and local governments to solve social issues, including homelessness, poverty, crime, and educational inequality (California Policy Lab, n.d.). The California Policy Lab utilizes the Open Science Framework and has established data-sharing agreements with more than a dozen county agencies in Los Angeles, Sonoma, and San Francisco covering “medical, mental health, criminal justice, social service, and homeless management information systems” (California Policy Lab, 2018). The lab recently received a \$1.2 million grant to expand to all University of California schools and partner with more public agencies to conduct policy-relevant research and overcome data silos.
- The **International Virtual Observatory Alliance** is an open platform enabling astronomers, educators, and the general public to discover, access, and integrate open data from worldwide (including in orbit) observatories. It links together the vast astronomical archives and databases around the world, together with analysis tools and computational services, into a single, integrated

facility. From its inception in 2002 through late 2020, the Virtual Observatory data have powered more than 2,300 scholarly papers,² covering the entire electromagnetic spectrum, from gamma-rays to radio waves.

- The **COVID-19 Open Research Dataset (CORD-19)** is an open collection of scientific articles and preprints related to COVID-19 and historical coronavirus research. The dataset can be text mined and analyzed using artificial intelligence and natural language processing to generate new insights into combatting the virus. The dataset was downloaded more than 200,000 times in the first 3 months after its release (Wang et al., 2020). This is one of several examples of open science’s centrality in rapidly addressing this era’s most critical public health challenge.

OPEN SCIENCE AND THE STATUS QUO

Historically, academic research environments have incentivized competition between individual researchers, which stymies collaboration and leads to the hoarding of knowledge. These dynamics persist as a function of the pursuit of “excellence” by research institutions, which results in the widespread usage of metrics that decrease transparency and collaboration. For example, measuring success by the number of patents filed and industry spinoffs launched leads to the safeguarding of intellectual property by researchers rather than the sharing of this information with external organizations that can increase the possibility of taking a product to market. Likewise, when academic departments measure their success by the volume of research citations and grant tenure to researchers who are cited most frequently, researchers are pressured to be the first to publish their findings and often operate in isolation, rarely venturing out of their respective research programs and communities (Heenan and Williams, 2018). Researchers become understandably hesitant to make their data and findings openly available out of fear of being “scooped” by other researchers (Berg and Niemeyer, 2018). Although competition between institutions and individual researchers may have been adequate to drive discovery in the 20th century, the “explosive sophistication” of science and engineering fields, in particular, has made it impossible for a single individual to be an

²Data accessed from the SAO/NASA Astrophysics Data System, October 16, 2020.

expert in multiple specialties or even a single subfield. Effective knowledge production now demands teams of researchers with diverse knowledge and skills to facilitate ongoing discovery (Brooks, 2010). Greater collaboration, rather than being an aspirational ideal that might produce better outcomes under the right circumstances, has now become a necessity to contend with the extreme specialization of knowledge production and ensure that discovery continues apace.

Open science practices, in contrast to traditional models of knowledge production, emphasize that open, transparent, and collaborative research dissemination practices more properly balance collective, institutional, and individual benefits. Open science represents a positive evolution of the research endeavor along three dimensions:

- ***Collaboration drives innovation with the potential for broad social impact.*** Open science approaches can reduce barriers between researchers and other stakeholders, including the public (e.g., by better informing and directly involving patients in biosciences) (Gold, 2016). By making data openly accessible between researchers and the public, open science can provide greater opportunities for interdisciplinary, collaborative research across institutions worldwide (Uhlir, 2015). Heightened collaboration can also lead to dynamic new knowledge hubs and remove barriers to upstream research and technology transfer (Gold, 2016).
- ***Greater efficiency and speed.*** Open data practices also drive efficiency by enabling real-time, data-driven decision making (Adams, 2018; SPARC, n.d.). The sharing of data reduces transaction costs; increases reproducibility and reuse of data; decreases redundancy; and drives greater transparency, heightened efficiency, and accelerated sustainable innovation (Gold, 2016; Gold et al., 2019; Tennant et al., 2016).
- ***Replicability enhances trust and research quality.*** By enhancing researchers' ability to verify results, open science practices help to build trust and goodwill among researchers and enhance the legitimacy of research (Popkin, 2019; Uhlir, 2015).

THE ROLE OF RESEARCH STAKEHOLDER ORGANIZATIONS

Open science has been largely pioneered by individual researchers who believe the benefits of this approach—to their work, to the shared understanding of a problem space, to their discipline, and to society—outweigh the reputational benefits that may be derived from the older, competition-based models of knowledge production. However, many researchers continue to face strong disincentives for engaging in open science practices, especially early-career scholars, who face the greatest pressure to conform to the traditional modes of credit and recognition that can lead to tenure. The wider uptake of open science, therefore, requires the organizational stakeholders responsible for reward systems—institutions, government agencies, and philanthropies chief among them—to establish new incentives and processes that prioritize open science activities. Because the competition-based incentives that motivate researchers reflect institutional prerogatives to demonstrate excellence vis-à-vis other institutions, institutions must also convene to identify new approaches toward facilitating interinstitutional collaboration and collectively address external barriers to open science.

Fortunately, the values that underpin open science—such as inclusiveness, collaboration, social impact, and scientific literacy—are mutually reinforcing to the missions of the research institutions, agencies, and funding organizations that support scientific research. Forward-thinking organizations have already begun to implement incentives for open science practices that provide a model for others to follow, which have taken several forms, including the following:

1. ***Creating supportive environments.*** The Tanenbaum Open Science Institute (TOSI) at the Neuro (Montreal Neurological Institute-Hospital) was designed as a “living lab for Open Science” to achieve the goals of accelerating discovery in neuroscience through collaboration, developing global best practices, and delivering innovative treatment to benefit patients afflicted by neurological diseases. TOSI supports four Open Science initiatives, including a biologic imaging and genetic repository, an open research platform, several open neuroinformatics platforms, and an early-stage drug discovery unit that collaborates with academia and industry partners (Gold, 2016; Neuro, n.d.).

2. ***Incentivizing open access publishing.*** The Bill & Melinda Gates Foundation and the Wellcome Trust, which funded \$1.3 billion and \$1.2 billion in global health research, respectively, joined a consortium of 11 European funding agencies that require all funded research to be free immediately upon publication. This incentive effectively requires scientists to publish papers in open access journals rather than those that charge subscriptions (Stokstad, 2018).
3. ***Awards for Open Science innovation.*** In 2017 the National Institutes of Health, Wellcome Trust, and the Howard Hughes Medical Institute hosted the Open Science Prize competition, leveraging public input to determine award finalists (NIH, 2017).

These examples represent the kinds of new incentives critical to instantiating the cultural shift necessary for sustained uptake of Open Science. In designing new incentives, research organizations and funders may also consider topics such as advancing the theory and practice of Open Science; how hiring decisions may contribute to cultures supportive of Open Science; and how funding mechanisms can be evolved to encourage open access publishing, data archiving and sharing, preregistration, and collaboration. The National Academies' Roundtable on Aligning Incentives for Open Science aims to encourage exploration of these topics and a wide range of possibilities for using incentives to realize the full potential for scientific research as a catalyst for discovery, economic growth, and societal benefit.

REFERENCES

- Abdill, R. J., and R. Blekhman. 2019. Meta-Research: Tracking the popularity and outcomes of all bioRxiv preprints. *eLife* 8:e45133. DOI: 10.7554/eLife.45133.
- Adams, J. 2018. Open Data: Enabling Fact-Based, Data-Driven Decisions. U.S. Department of Agriculture (blog). Available at <https://www.usda.gov/media/blog/2018/07/13/open-data-enabling-fact-based-data-driven-decisions>. Accessed January 19, 2021.
- Benkler, Y. 2006. *The Wealth of Networks—How Social Production Transforms Markets and Freedom*. Yale University Press. Available at http://www.benkler.org/Benkler_Wealth_Of_Networks.pdf. Accessed August 30, 2021.
- Berg, D. R., and K. E. Niemeyer. 2018. The case for openness in engineering research. *F1000Research* 7:501.

- Brooks, Jr., F. P. 2010. *The Design of Design: Essays from a Computer Scientist*. London: Pearson Education.
- California Policy Lab. 2018. California Policy Lab Awarded \$1.2M UC Multicampus Research Grant. Press release, December 13, 2018. Available at <https://www.capolicylab.org/wp-content/uploads/2018/12/CPL-Press-Release-re-MRPI-12-12-18-final.pdf>. Accessed January 19, 2021.
- California Policy Lab. n.d. What we do. Available at <https://www.capolicylab.org/what-we-do>. Accessed January 19, 2021.
- Digital Science and Figshare. 2018. *The State of Open Data Report 2018*. Available at https://figshare.com/articles/report/The_State_of_Open_Data_Report_2018/7195058. Accessed January 19, 2021.
- Gold, E. R. 2016. Accelerating translational research through open science: The Neuro Experiment. *PLOS Biology* 14(12):e2001259.
- Gold, E. R., S. E. Ali-Khan, L. Allen, L. Ballell, M. Barral-Netto, D. Carr, D. Chalaud, S. Chaplin, M. S. Clancy, P. Clarke, R. Cook-Deegan, A. P. Dinsmore, M. Doerr, L. Federer, S. A. Hill, N. Jacobs, A. Jean, O. A. Jefferson, C. Jones, L. J. Kahl, T. M. Kariuk, S. N. Kassell, R. Kiley, E. R. Kittrie, B. Kramer, W. H. Lee, E. MacDonald, L. M. Mangravite, E. Marincola, D. Mietchen, J. C. Molloy, M. Namchuk, B. A. Nosek, S. Paquet, C. Pirmez, A. Seyller, M. Skingle, S. N. Spadotto, S. Staniszewska, and M. Thelwall. 2019. An open toolkit for tracking open science partnership implementation and impact. *Gates Open Research* 3:1442. Available at <https://doi.org/10.12688/gatesopenres.12958.1>. Accessed January 19, 2021.
- Gries, C., E. Gilbert, and N. Franz. 2014. Symbiota – A virtual platform for creating voucher-based biodiversity information communities. *Biodiversity Data Journal* 2:e1114. DOI: 10.3897/BDJ.2.e1114.
- Heenan, A., and I. D. Williams. 2018. How open data can help the world better manage coral reefs. *The Conversation* (January 29). Available at <https://theconversation.com/how-open-data-can-help-the-world-better-manage-coral-reefs-88805>. Accessed January 19, 2021.
- Mayo, E., and T. Steinberg. 2007. The Power of Information: An Independent Review. Available at http://ses.library.usyd.edu.au/bitstream/2123/6557/1/PSI_vol2_chapter20.pdf. Accessed August 30, 2021.
- Neuro. n.d. Open science, to accelerate discovery and deliver cures. Available at <https://www.mcgill.ca/neuro/open-science-0>. Accessed January 19, 2021.
- Nielsen, M. 2011. An informal definition of open science. OpenScience Project (website). Available at <http://openscience.org/an-informal-definition-of-openscience>. Accessed January 19, 2021.
- NIH (National Institutes of Health). 2017. Open Science Prize announces epidemic tracking tool as grand prize winner. Press release, February 28, 2017. Available at <https://www.nih.gov/news-events/news-releases/open-science-prize-announces-epidemic-tracking-tool-grand-prize-winner>. Accessed January 19, 2021.

- NRC (National Research Council). 2009. *The Socioeconomic Effects of Public Sector Information on Digital Networks: Toward a Better Understanding of Different Access and Reuse Policies: Workshop Summary*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/12687>.
- Piwowar, H., J. Priem, and R. Orr. 2019. The Future of OA: A large-scale analysis projecting Open Access publication and readership. *bioRxiv* 795310. DOI: <https://doi.org/10.1101/795310>.
- Popkin, G. 2019. Data sharing and how it can benefit your scientific career. *Nature* (Career Feature article), May 13, 2019. Available at <https://www.nature.com/articles/d41586-019-01506-x>. Accessed January 19, 2021.
- SPARC (Scholarly Publishing and Academic Resources Coalition). n.d. “From ideas to industries: Human Genome Project.” Available at <https://sparcopen.org/impact-story/human-genome-project>. Accessed January 19, 2021.
- Spellman, B. A., E. A. Gilbert, and K. S. Corker. 2018. Open science. In *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience: Volume 5 Methodology*, 1–47. New York: John Wiley & Sons.
- Stokstad, E. 2018. In a win for open access, two major funders won't cover publishing and hybrid journals. *Science*. DOI:10.1126/science.aav9422.
- Swan, A., and S. Brown. 2004. Authors and open access publishing. *Learned Publishing* 17:219–224. DOI:10.1087/095315104323159649.
- Swan, A., and S. Brown, S. 2005. Open access self-archiving: An author study. Key Perspectives (website). Available at <http://cogprints.org/4385/1/jisc2.pdf>. Accessed January 19, 2021.
- Symbiota. n.d. Symbiota Introduction. Available at <http://symbiota.org/docs>. Accessed January 19, 2021.
- Tennant, J. P., F. Waldner, D. C. Jacques, P. Masuzzo, L. B. Collister, and C. H. J. Hartgerink. 2016. The academic, economic and societal impacts of Open Access: An evidence-based review. *F1000Research* 5:632. Available at <https://f1000research.com/articles/5-632>. Accessed January 19, 2021.
- The Lab @ DC. n.d. Home. Available at <https://thelab.dc.gov>. Accessed January 19, 2021.
- Tripp, S., and M. Grueber. 2011. *Economic Impact of the Human Genome Project*. Battelle Memorial Institute, Technology Partnership Practice. Available at <https://www.battelle.org/docs/default-source/misc/battelle-2011-misc-economic-impact-human-genome-project.pdf?sfvrsn=6>. Accessed January 19, 2021.
- Uhlir, P. 2015. The Value of Open Data Sharing – A White Paper for the Group on Earth Observations. Group on Earth Observations, GEO-XII Plenary and Mexico City Ministerial Summit, November 11–12, 2015. Available at https://www.earthobservations.org/documents/geo_xii/GEO-XII_09_The%20Value%20of%20Open%20Data%20Sharing.pdf.
- Wang, L. L., K. Lo, Y. Chandrasekhar, R. Reas, J. Yang, D. Burdick, D. Eide, K. Funk, Y. Katsis, R. Kinney, Y. Li, Z. Liu, W. Merrill, P. Mooney, D. Murdick, D. Rishi, J. Sheehan, Z. Shen, B. Stilson, A. Wade, K. Wang, N. X. R. Wang, C. Wilhelm, B. Xie, D. Raymond, D. S. Weld, O. Etzioni, and S. Kohlmeier. 2020. CORD-19: The Covid-19 Open Research Dataset. *ArXiv* [preprint]. April 22, 2020. arXiv:2004.10706v2.

- Wellcome Trust Sanger Institute. 2016. Online epidemic tracking tool embraces open data and collective intelligence to understand outbreaks. News article, November 30, 2016. Available at <https://www.sanger.ac.uk/news/view/online-epidemic-tracking-tool-embraces-open-data-and-collective-intelligence-understand>. Accessed January 19, 2021.
- Zittrain, J. 2006. The Generative Internet. *Harvard Law Review* 119. Available at <http://dash.harvard.edu/handle/1/9385626>. Accessed August 30, 2021.
- Zuccala, J. 2010. Open access and civic scientific information literacy. *Information Research: An International Electronic Journal* 15(1).

II. OPEN SCIENCE SIGNALING LANGUAGE TEMPLATE AND RUBRIC³

Maryrose Franko, Health Research Alliance
 Courtney Brown, Lumina Foundation
 Rachel Bruce, UK Research and Innovation
 Glenn Dillon, American Heart Association
 Randolph Hall, University of Southern California
 Robert Kiley, Wellcome Trust
 Lisa Nichols, Formerly, Office of Science and Technology Policy
 Greg Tananbaum, Open Research Funders Group
 Roger Wakimoto, University of California, Los Angeles

This resource provides specific language that can be adapted and adopted to signal an organization's interest in open science activities at specific points of high leverage (e.g., grant applications, job postings). Even absent adoption of formal open science policies, this language can indicate an organization's values and “nudge” researcher behavior toward open practices.

NOTE: The language below can be customized to reflect the specific research considerations of each participating organization.

FUNDERS AND AGENCIES

Grant Application

1. Foundation XYZ values the open sharing of research outputs. If applicable, describe (1) instances where you have engaged in “open” activities (such as making articles open access and sharing data/code according to FAIR principles [Findability, Accessibility, Interoperability, and Reuse of digital assets]); (2) examples of how your open research outputs have been used by others in your discipline, in other disciplines, and/or outside of academia (include DOIs if possible); and (3) plans to engage in open activities in the future.

³The views expressed are those of the authors and do not necessarily reflect the official policies or positions of their employing organizations.

2. For each of the categories below, provide *representative examples* demonstrating how you have made research outputs resulting from other projects openly accessible. If possible, provide the DOI and license terms under which the materials are available.
 - Open access articles
 - Open access books, book chapters, and/or monographs
 - Copies of your papers, chapters, monographs, or other published materials in institutional or disciplinary repositories
 - Preprints
 - Datasets
 - Software/Code
 - Materials/Reagents
 - Preregistration plans
 - Other outputs (please describe)

Additionally, it is important to include negative and null results, which could be covered in a variety of information formats.

Grant Progress Report

1. Foundation XYZ values the open sharing of research outputs. If applicable, describe, in the context of this funded project, (1) instances where you have engaged in “open” activities (such as making articles open access and sharing data/code according to FAIR principles); (2) examples of how your open research outputs have been used by others in your discipline, in other disciplines, and/or outside of academia (include DOIs if possible); and (3) plans to engage in open activities as the project progresses and concludes.
2. For each of the categories below, provide *representative examples* demonstrating how you have made research outputs resulting from this project openly accessible. If possible, provide the DOI and license terms under which the materials are available.
 - Open access articles
 - Open access books, book chapters, and/or monographs

- Copies of your papers, chapters, monographs, or other published materials in institutional or disciplinary repositories
- Preprints
- Datasets
- Software/Code
- Materials/Reagents
- Preregistration plans
- Other outputs (please describe)

Additionally, it is important to include negative and null results, which could be covered in a variety of information formats.

Grant Final Report

1. Foundation XYZ values the open sharing of research outputs. If applicable, describe, in the context of this funded project, (1) instances where you have engaged in “open” activities (such as making articles open access and sharing data/code according to FAIR principles); (2) examples of how your open research outputs have been used by others in your discipline, in other disciplines, and/or outside of academia (include DOIs if possible); and (3) plans to engage in open activities for any future outputs pertaining to this project.
2. For each of the categories below, provide *representative examples* demonstrating how you have made research outputs resulting from this project openly accessible. If possible, provide the DOI and license terms under which the materials are available.
 - Open access articles
 - Open access books, book chapters, and/or monographs
 - Copies of your papers, chapters, monographs, or other published materials in institutional or disciplinary repositories
 - Preprints
 - Datasets
 - Software/Code
 - Materials/Reagents

- Preregistration plans
- Other outputs (please describe)

Additionally, it is important to include negative and null results, which could be covered in a variety of information formats.

UNIVERSITIES

Faculty Annual Report

1. For each of the categories below, provide *representative examples* demonstrating how (where appropriate) you have made outputs resulting from your research openly accessible. If possible, provide the DOI and license terms under which the materials are available.
 - Open access articles
 - Open access books, book chapters, and/or monographs
 - Copies of your papers, chapters, monographs, or other published materials in institutional or disciplinary repositories
 - Preprints
 - Datasets
 - Software/Code
 - Materials/Reagents
 - Preregistration plans
 - Other outputs (please describe)

Additionally, it is important to include negative and null results, which could be covered in a variety of information formats.

2. If known, describe how others have made use of these open research outputs, and include relevant DOIs if possible. This can include use in other disciplines and outside of academia.
3. Describe the impact that your openly available research outputs from this evaluation period have had from the research, public policy, pedagogic, and/or societal perspectives.

University Job Posting/Application

1. University XYZ values transparent, replicable, and reproducible research and open science principles (the open sharing of research outputs, including, but not limited to, open access and open data). How have you engaged in “open” activities during your career and how do you plan to do so in the future?

Or

2. University XYZ values transparent, replicable research and open science principles (the open sharing of research outputs, including, but not limited to, open access and open data). Describe the impact that your openly available research outputs have had from the research, public policy, pedagogic, and/or societal perspectives.

SENDING SIGNALS RUBRIC

This rubric complements the “Suggested Open Science Signaling Language” document produced by the same authors, which can be used by universities, agencies, philanthropies, and other stakeholders to highlight an organization’s interest in open science activities at specific points of high leverage (such as grant applications, job postings). The rubric can be used by tenure and promotion committees, program managers, department chairs, hiring committees, and others tasked with evaluating the absolute and relative merits of responses to the signaling questions.

This workbook contains four sheets—one each with language pertaining specifically to articles, data, and other forms of research outputs at both application and reporting stages. The first sheet (Tables 1 and 2) is the amalgamated version, the second sheet (Tables 3 and 4) includes the articles version, and the third sheet (Tables 5 and 6) provides the data version. The fourth sheet (Tables 7 and 8) is the other output version that provides combined language encompassing all of these types of open science activities.

Please note that both the Sending Signals Language and the Sending Signals Rubric can be adapted to address the unique considerations, priorities, and norms of a specific community.

Table 1 Amalgamated Version – Application Stage

<i>Application Stage (e.g., jobs, grants)</i>	Beginning 1	Developing 2	Accomplished 3	Exemplary 4
Describe instances where you have engaged in “open” activities (such as making articles open access and sharing data/principles), including representative examples	The researcher has not, in their recent research (<5 years), demonstrably engaged in open science practices such as making articles, data, and other research outputs openly available for access and reuse.	The researcher has sometimes engaged in open science practices. This is defined as occasionally making recent research (<5 years) available openly for access and reuse. Specific activities include (a) making at least one of their articles available in open access journals or repositories; (b) to the extent that the researcher has generated research data, making at least one of these datasets available in accessible repositories under adherence to the FAIR principles; and (c) to the extent that the researcher has generated research outputs beyond articles and data, making at least one of these materials openly available for access and reuse. Additionally, the researcher demonstrates at least some open science hygiene (e.g., use of DOIs, Creative Commons licenses).	The researcher has frequently engaged in open science practices. This is defined as often making recent research (<5 years) available openly for access and reuse. Specific activities include (a) making some (more than one, but less than most) of their articles available in open access journals or repositories; (b) to the extent that the researcher has generated research data, making some (more than one dataset, but less than most) of these data available in accessible repositories under adherence to the FAIR principles; and (c) to the extent that the researcher has generated research outputs beyond articles and data, making the majority of these materials openly available for access and reuse. Additionally, the researcher consistently demonstrates good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).	The researcher has consistently engaged in open science practices. This is defined as making the majority of recent research (<5 years) available openly for access and reuse. Specific activities include (a) making the majority of their articles available in open access journals or repositories; (b) to the extent that the researcher has generated research data, making the majority of these data available in accessible repositories under adherence to the FAIR principles; and (c) to the extent that the researcher has generated research outputs beyond articles and data, making the majority of these materials openly available for access and reuse. Additionally, the researcher consistently demonstrates good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).

Table 1 Continued

<i>Application Stage (e.g., jobs, grants)</i>	Beginning 1	Developing 2	Accomplished 3	Exemplary 4
Provide examples of how your open research outputs have been used by others in your discipline, in other disciplines, and/or outside of academia (include DOIs, if possible)	The researcher cannot provide qualitative and/or quantitative evidence that any of their recent (<5 years) open research outputs have been used by others.	The researcher can provide qualitative and/or quantitative evidence that at least one of their recent (<5 years) open research outputs has been used by others.	The researcher can provide qualitative and/or quantitative evidence that (a) some of their recent (<5 years) open research outputs have been used by others; and/or (b) a narrower range of their recent (<5 years) open research outputs have been used deeply within a specific community.	The researcher can provide qualitative and/or quantitative evidence that (a) a wide range of their recent (<5 years) open research outputs have been used by others; and/or (b) a narrower range of their recent (<5 years) open research outputs have been used deeply within a specific community.

Enumerate your plans to engage in open activities in the future	The researcher has articulated a clear plan to make at least some research outputs (including, but not limited to, articles and data) available openly for access and reuse. Specific activities include (a) making their articles available in open access journals or repositories; (b) to the extent that the researcher has generated research data, making most of these data available in accessible repositories under adherence to the FAIR principles; and (c) to the extent that the researcher has generated research outputs beyond articles and data, making at least some of these materials openly available for access and reuse. Additionally, the researcher has articulated a plan that demonstrates an awareness of at least some aspects of good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).	The researcher has articulated a clear plan to make most research outputs (including, but not limited to, articles and data) available openly for access and reuse. Specific activities include (a) making most of their articles available in open access journals or repositories; (b) to the extent that the researcher has generated research data, making most of these data available in accessible repositories under adherence to the FAIR principles; and (c) to the extent that the researcher has generated research outputs beyond articles and data, making most of these materials openly available for access and reuse. Additionally, the researcher has articulated a plan that demonstrates an intent to engage in good open science hygiene in most instances (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).	The researcher has articulated a clear plan to make all appropriate research outputs (including, but not limited to, articles and data) available openly for access and reuse. Specific activities include (a) making their articles available in open access journals or repositories; (b) to the extent that the researcher has generated research data, making these data available in accessible repositories under adherence to the FAIR principles; and (c) to the extent that the researcher has generated research outputs beyond articles and data, making these materials openly available for access and reuse. Additionally, the researcher has articulated a clear and consistent plan to engage in good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).
---	---	---	---

NOTE: FAIR – Findability, Accessibility, Interoperability, and Reuse of digital assets; ORCID – Open Researcher and Contributor ID; DOI – Digital Object Identifier.

Table 2. Amalgamated Version – Reporting Stage

<i>Reporting Stage (e.g., faculty tenure and promotion reviews, interim and final grant reports)</i>	Beginning 1	Developing 2	Accomplished 3	Exemplary 4
For your work (related to this grant/during this time period), describe instances where you have engaged in “open” activities (such as making articles open access and sharing data/code according to FAIR principles), including representative examples	The researcher has not, in their research (for this project/period), demonstrably engaged in open science practices such as making articles, data, and other research outputs openly available for access and reuse.	The researcher has sometimes engaged in open science practices. This is defined as occasionally making research (for this project/period) available openly for access and reuse. Specific activities include (a) making at least one of their articles available in open access journals or repositories; (b) to the extent that the researcher has generated research data, making at least one of these datasets available in accessible repositories under adherence to the FAIR principles; and (c) to the extent that the researcher has generated research outputs beyond articles and data, making at least one of these materials openly available for access and reuse. Additionally, the researcher demonstrates at least some open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).	The researcher has frequently engaged in open science practices. This is defined as often making research (for this project/period) available openly for access and reuse. Specific activities include (a) making more than one, but less than most) of their articles available in open access journals or repositories; (b) to the extent that the researcher has generated research data, making some (more than one dataset, but less than most) of these data available in accessible repositories under adherence to the FAIR principles; and (c) to the extent that the researcher has generated research outputs beyond articles and data, making some (more than one, but less than most) of these materials openly available for access and reuse. Additionally, the researcher frequently demonstrates good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).	The researcher has consistently engaged in open science practices. This is defined as making the majority of research (for this project/period) available openly for access and reuse. Specific activities include (a) making the majority of their articles available in open access journals or repositories; (b) to the extent that the researcher has generated research data, making the majority of these data available in accessible repositories under adherence to the FAIR principles; and (c) to the extent that the researcher has generated research outputs beyond articles and data, making the majority of these materials openly available for access and reuse. Additionally, the researcher consistently demonstrates good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).

<p>For your work (related to this grant/during this time period), provide examples of how your open research outputs have been used by others in your discipline, in other disciplines, and/or outside of academia (include DOIs, if possible)</p>	<p>The researcher cannot provide qualitative and/or quantitative evidence that any of their open research outputs (for this project/period) have been used by others.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that at least one of their open research outputs (for this project/period) has been used by others.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that (a) some of their open research outputs (for this project/period) have been used by others; and/or (b) a narrower range of their open research outputs (for this project/period) have been used deeply within a specific community.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that (a) a wide range of their open research outputs (for this project/period) have been used by others; and/or (b) a narrower range of their open research outputs (for this project/period) have been used deeply within a specific community.</p>
--	---	--	---	---

continued

Table 2 Continued

<i>Reporting Stage (e.g., faculty tenure and promotion reviews, interim and final grant reports)</i>	Beginning 1	Developing 2	Accomplished 3	Exemplary 4
<p>For your work (related to this grant/during this time period), enumerate your plans to engage in open activities in the future</p>	<p>The researcher has not articulated a clear plan to make at least some research outputs (including, but not limited to, articles and data) available openly for access and reuse.</p>	<p>The researcher has articulated a clear plan to make at least some research outputs (including, but not limited to, articles and data) available openly for access and reuse. Specific activities include (a) making at least some of their articles available in open access journals or repositories; (b) to the extent that the researcher has generated research data, making most of these data available in accessible repositories under adherence to the FAIR principles; and (c) to the extent that the researcher has generated research outputs beyond articles and data, making at least some of these materials openly available for access and reuse. Additionally, the researcher has articulated a plan that demonstrates an awareness of at least some aspects of good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>	<p>The researcher has articulated a clear plan to make most research outputs (including, but not limited to, articles and data) available openly for access and reuse. Specific activities include (a) making most of their articles available in open access journals or repositories; (b) to the extent that the researcher has generated research data, making most of these data available in accessible repositories under adherence to the FAIR principles; and (c) to the extent that the researcher has generated research outputs beyond articles and data, making most of these materials openly available for access and reuse. Additionally, the researcher has articulated a plan that demonstrates an intent to engage in good open science hygiene in most instances (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>	<p>The researcher has articulated a clear plan to make all appropriate research outputs (including, but not limited to, articles and data) available openly for access and reuse. Specific activities include (a) making their articles available in open access journals or repositories; (b) to the extent that the researcher has generated research data, making these data available in accessible repositories under adherence to the FAIR principles; and (c) to the extent that the researcher has generated research outputs beyond articles and data, making these materials openly available for access and reuse. Additionally, the researcher has articulated a clear and consistent plan to engage in good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>

NOTE: FAIR – Findability, Accessibility, Interoperability, and Reuse of digital assets; ORCID – Open Researcher and Contributor ID; DOI – Digital Object Identifier.

Table 3 Articles Version – Application Stage

<i>Application Stage (e.g., jobs, grants)</i>	Beginning 1	Developing 2	Accomplished 3	Exemplary 4
Describe instances where you have engaged in making articles open access, including representative examples	The researcher has not, in their recent research (<5 years), demonstrably engaged in making articles openly available for access and reuse.	The researcher has sometimes engaged in open access practices. This is defined as occasionally making recent research articles (<5 years) available openly for access and reuse. Specific activities include (a) making at least one of their articles available in open access journals or repositories; and (b) demonstrating at least some open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).	The researcher has frequently engaged in open access practices. This is defined as often making recent research articles (<5 years) available openly for access and reuse. Specific activities include (a) making some (more than one, but less than most) of their articles available in open access journals or repositories; and (b) frequently demonstrating good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).	The researcher has consistently engaged in open access practices. This is defined as making the majority of recent research articles (<5 years) available openly for access and reuse. Specific activities include (a) making the majority of their journals or repositories; and (b) consistently demonstrating good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).

continued

Table 3 Continued

<i>Application Stage (e.g., jobs, grants)</i>	Beginning 1	Developing 2	Accomplished 3	Exemplary 4
Provide examples of how your open access articles have been used by others in your discipline, in other disciplines, and/or outside of academia (include DOIs if possible)	The researcher cannot provide qualitative and/or quantitative evidence that any of their recent (<5 years) open access articles have been used by others.	The researcher can provide qualitative and/or quantitative evidence that at least one of their recent (<5 years) open access articles has been used by others.	The researcher can provide qualitative and/or quantitative evidence that (a) some of their recent (<5 years) open access articles have been used by others; and/or (b) a narrower range of their recent (<5 years) open access articles have been used deeply within a specific community.	The researcher can provide qualitative and/or quantitative evidence that (a) a wide range of their recent (<5 years) open access articles have been used by others; and/or (b) a narrower range of their recent (<5 years) open access articles have been used deeply within a specific community.
Enumerate your plans to engage in open access activities in the future	The researcher has not articulated a clear plan to make at least some research articles available openly for access and reuse.	The researcher has articulated a clear plan to make at least some research articles available openly for access and reuse. Specific activities include (a) making at least some of their articles available in open access journals or repositories; and (b) articulating a plan that demonstrates an awareness of at least some aspects of good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).	The researcher has articulated a clear plan to make most research articles available openly for access and reuse. Specific activities include (a) making most of their articles available in open access journals or repositories; and (b) articulating a plan that demonstrates an intent to engage in good open science hygiene in most instances (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).	The researcher has articulated a clear plan to make all appropriate research articles available openly for access and reuse. Specific activities include (a) making their articles available in open access journals or repositories; and (b) articulating a clear and consistent plan to engage in good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).

NOTE: DOI – Digital Object Identifier; ORCID – Open Researcher and Contributor ID.

Table 4 Articles Version – Reporting Stage

<i>Reporting Stage (e.g., faculty tenure and promotion reviews, interim and final grant reports)</i>	Beginning 1	Developing 2	Accomplished 3	Exemplary 4
<p>For your work (related to this grant/during this time period), describe instances where you have engaged in open access activities, including representative examples</p>	<p>The researcher has not, in their research (for this project/period), demonstrably engaged in making research articles openly available for access and reuse.</p>	<p>The researcher has sometimes engaged in open access practices. This is defined as occasionally making research articles (for this project/period) available openly for access and reuse. Specific activities include (a) making at least one of their articles available in open access journals or repositories; and (b) demonstrating at least some open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>	<p>The researcher has frequently engaged in open access practices. This is defined as often making research articles (for this project/period) available openly for access and reuse. Specific activities include (a) making some (more than one, but less than most) of their articles available in open access journals or repositories; and (b) frequently demonstrating good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>	<p>The researcher has consistently engaged in open access practices. This is defined as making the majority of research articles (for this project/period) available openly for access and reuse. Specific activities include (a) making the majority of their articles available in open access journals or repositories; and (b) consistently demonstrating good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>

continued

Table 4 Continued

<i>Reporting Stage (e.g., faculty tenure and promotion reviews, interim and final grant reports)</i>	Beginning 1	Developing 2	Accomplished 3	Exemplary 4
<p>For your work (related to this grant/during this time period), provide examples of how your open access articles have been used by others in your discipline, in other disciplines, and/or outside of academia (include DOIs, if possible)</p>	<p>The researcher cannot provide qualitative and/or quantitative evidence that any of their open access articles (for this project/period) have been used by others.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that at least one of their open access articles (for this project/period) has been used by others.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that (a) some of their open access articles (for this project/period) have been used by others; and/or (b) a narrower range of their open access articles (for this project/period) have been used deeply within a specific community.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that (a) a wide range of their open access articles (for this project/period) have been used by others; and/or (b) a narrower range of their open access articles (for this project/period) have been used deeply within a specific community.</p>

<p>For your work (related to this grant/during this time period), enumerate your plans to engage in open access activities in the future</p>	<p>The researcher has not articulated a clear plan to make at least some research articles (including, but not limited to, articles and data) available openly for access and reuse.</p>	<p>The researcher has articulated a clear plan to make at least some research articles available openly for access and reuse. Specific activities include (a) making at least some of their articles available in open access journals or repositories; and (b) articulating a plan that demonstrates an awareness of at least some aspects of good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>	<p>The researcher has articulated a clear plan to make most research articles available openly for access and reuse. Specific activities include (a) making most of their articles available in open access journals or repositories; and (b) articulating a clear and consistent plan to engage in good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>
--	--	---	--

NOTE: DOI – Digital Object Identifier; ORCID – Open Researcher and Contributor ID.

Table 5 Data Version – Application Stage

<i>Application Stage</i> <i>(e.g., jobs, grants)</i>	Beginning 1	Developing 2	Accomplished 3	Exemplary 4
Describe instances where you have engaged in open data activities (such as sharing data according to FAIR principles), including representative examples	The researcher has not, in their recent research (<5 years), demonstrably engaged in making data available for access and reuse according to FAIR principles.	The researcher has sometimes engaged in open data practices. This is defined as occasionally making research data (<5 years) available for access and reuse according to FAIR principles. Specific activities include (a) making at least one of their datasets available in accessible repositories under adherence to the FAIR principles; and (b) demonstrating at least some open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).	The researcher has frequently engaged in open data practices. This is defined as often making recent research data (<5 years) available openly for access and reuse according to FAIR principles. Specific activities include (a) making some (more than one dataset, but less than most) of their research data available in accessible repositories under adherence to the FAIR principles; and (b) frequently demonstrating good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).	The researcher has consistently engaged in open data practices. This is defined as making the majority of recent research data (<5 years) available openly for access and reuse according to FAIR principles. Specific activities include (a) making the majority of their research data available in accessible repositories under adherence to the FAIR principles; and (b) consistently demonstrating good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).

<p>Provide examples of how your open datasets have been used by others in your discipline, in other disciplines, and/or outside of academia (include DOIs, if possible)</p>	<p>The researcher cannot provide qualitative and/or quantitative evidence that any of their recent (<5 years) open datasets have been used by others.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that at least one of their recent (<5 years) open datasets has been used by others.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that (a) some of their recent (<5 years) open datasets have been used by others; and/or (b) a narrower range of their recent (<5 years) open datasets have been used deeply within a specific community.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that (a) a wide range of their recent (<5 years) open datasets have been used by others; and/or (b) a narrower range of their recent (<5 years) open datasets have been used deeply within a specific community.</p>
<p>Enumerate your plans to engage in open data activities in the future</p>	<p>The researcher has not articulated a clear plan to make at least some research data available for access and reuse according to FAIR principles.</p>	<p>The researcher has articulated a clear plan to make at least some research data available for access and reuse according to FAIR principles. Specific activities include (a) making most of their research data available in accessible repositories under adherence to the FAIR principles; and (b) articulating a plan that demonstrates an awareness of at least some aspects of good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>	<p>The researcher has articulated a clear plan to make most research data available for access and reuse according to FAIR principles. Specific activities include (a) making most of their research data available in accessible repositories under adherence to the FAIR principles; and (b) articulating a clear and consistent plan to engage in good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>	<p>The researcher has articulated a clear plan to make all appropriate research data available for access and reuse according to FAIR principles. Specific activities include (a) making their research data available in accessible repositories under adherence to the FAIR principles; and (b) articulating a clear and consistent plan to engage in good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>

NOTE: FAIR – Findability, Accessibility, Interoperability, and Reuse of digital assets; DOI – Digital Object Identifier; ORCID – Open Researcher and Contributor ID.

Table 6 Data Version – Reporting Stage

<i>Reporting Stage (e.g., faculty tenure and promotion reviews, interim and final grant reports)</i>	Beginning 1	Developing 2	Accomplished 3	Exemplary 4
For your work (related to this grant/during this time period), describe instances where you have engaged in open data activities (such as sharing data according to FAIR principles), including representative examples	The researcher has not, in their research (for this project/period), demonstrably engaged in making data available for access and reuse according to FAIR principles.	The researcher has sometimes engaged in open data practices. This is defined as occasionally making research data (for this project/period) available for access and reuse according to FAIR principles. Specific activities include (a) making at least one of their datasets available in accessible repositories under adherence to the FAIR principles; and (b) demonstrating at least some open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).	The researcher has frequently engaged in open data practices. This is defined as often making research data (for this project/period) available openly for access and reuse according to FAIR principles. Specific activities include (a) making some (more than one dataset, but less than most) of their research data available in accessible repositories under adherence to the FAIR principles; and (b) frequently demonstrating good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).	The researcher has consistently engaged in open data practices. This is defined as making the majority of research data (for this project/period) available openly for access and reuse according to FAIR principles. Specific activities include (a) making the majority of their research data available in accessible repositories under adherence to the FAIR principles; and (b) consistently demonstrating good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).

<p>For your work (related to this grant/during this time period), provide examples of how your open datasets have been used by others in your discipline, in other disciplines, and/or outside of academia (include DOIs if possible)</p>	<p>The researcher cannot provide qualitative and/or quantitative evidence that any of their open datasets (for this project/period) have been used by others.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that at least one of their open datasets (for this project/period) has been used by others.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that (a) some of their open datasets (for this project/period) have been used by others; and/or (b) a narrower range of their open datasets (for this project/period) have been used deeply within a specific community.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that (a) a wide range of their open datasets (for this project/period) have been used by others; and/or (b) a narrower range of their open datasets (for this project/period) have been used deeply within a specific community.</p>
<p>For your work (related to this grant/during this time period), enumerate plans to engage in open activities in the future</p>	<p>The researcher has not articulated a clear plan to make at least some research data available for access and reuse according to FAIR principles.</p>	<p>The researcher has articulated a clear plan to make at least some research data available for access and reuse according to FAIR principles. Specific activities include (a) making most of their research data available in accessible repositories under adherence to the FAIR principles; and (b) articulating a plan that demonstrates an awareness of at least some aspects of good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>	<p>The researcher has articulated a clear plan to make most research data available for access and reuse according to FAIR principles. Specific activities include (a) making most of their research data available in accessible repositories under adherence to the FAIR principles; and (b) articulating a plan that demonstrates an intent to engage in good open science hygiene in most instances (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>	<p>The researcher has articulated a clear plan to make all appropriate research data available for access and reuse according to FAIR principles. Specific activities include (a) making their research data available in accessible repositories under adherence to the FAIR principles; and (b) articulating a clear and consistent plan to engage in good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>

NOTE: FAIR – Findability, Accessibility, Interoperability, and Reuse of digital assets; DOI – Digital Object Identifier; ORCID – Open Researcher and Contributor ID.

Table 7 Other Outputs Version – Application Stage

<i>Application Stage (e.g., Jobs, grants)</i>	Beginning 1	Developing 2	Accomplished 3	Exemplary 4
Describe instances where you have engaged in “open” activities beyond sharing articles and data, including representative examples	The researcher has not, in their recent research (<5 years), demonstrably engaged in making research outputs beyond articles and data openly available for access and reuse.	The researcher has (a) occasionally made recent (<5 years) research outputs beyond articles and data available openly for access and reuse; and (b) demonstrated at least some open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).	The researcher has frequently made recent (<5 years) research outputs beyond articles and data available openly for access and reuse. Specific activities include (a) making some (more than one, but less than most) of these outputs available for access and reuse; and (b) frequently demonstrating good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).	The researcher has (a) consistently made the majority of recent (<5 years) research outputs beyond articles and data available openly for access and reuse; and (b) consistently demonstrated good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).

<p>Provide examples of how your open research outputs beyond articles and data have been used by others in your discipline, in other disciplines, and/or outside of academia (include DOIs if possible)</p>	<p>The researcher cannot provide qualitative and/or quantitative evidence that any of their recent (<5 years) open research outputs beyond articles and data have been used by others.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that at least one of their recent (<5 years) open research outputs beyond articles and data has been used by others.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that (a) a wide range of their recent (<5 years) open research outputs beyond articles and data have been used by others; and/or (b) a narrower range of their recent (<5 years) open research outputs beyond articles and data have been used deeply within a specific community.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that (a) a wide range of their recent (<5 years) open research outputs beyond articles and data have been used by others; and/or (b) a narrower range of their recent (<5 years) open research outputs beyond articles and data have been used deeply within a specific community.</p>
<p>Enumerate your plans to engage in open activities in the future, beyond the open sharing of articles and data</p>	<p>The researcher has not articulated a clear plan to make at least some research outputs beyond articles and data available openly for access and reuse.</p>	<p>The researcher has articulated a clear plan to (a) make at least some research outputs beyond articles and data available openly for access and reuse; and (b) engage in at least some aspects of good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>	<p>The researcher has articulated a clear plan to (a) make most research outputs beyond articles and data available openly for access and reuse; and (b) engage in good open science hygiene in most instances (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>	<p>The researcher has articulated a clear plan to (a) make all appropriate research outputs beyond articles and data available openly for access and reuse; and (b) engage in consistent good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>

NOTE: DOI – Digital Object Identifier; ORCID – Open Researcher and Contributor ID.

Table 8 Other Outputs Version – Reporting Stage

<i>Reporting Stage (e.g., faculty tenure and promotion reviews, interim and final grant reports)</i>	Beginning 1	Developing 2	Accomplished 3	Exemplary 4
<p>For your work (related to this grant/during this time period), describe instances where you have engaged in “open” activities (beyond sharing articles and data), including representative examples</p>	<p>The researcher has not, in their research (for this project/period), demonstrably engaged in making research outputs beyond articles and data openly available for access and reuse.</p>	<p>The researcher has (a) occasionally made research outputs (for this project/period) beyond articles and data available openly for access and reuse; and (b) demonstrated at least some open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>	<p>The researcher has frequently made research outputs research (for this project/period) beyond articles and data available openly for access and reuse. Specific activities include (a) making some (more than one, but less than most) of these outputs available for access and reuse; and (b) frequently demonstrating good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>	<p>The researcher has (a) consistently made the majority of research outputs research (for this project/period) beyond articles and data available openly for access and reuse; and (b) consistently demonstrated good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>

<p>For your work (related to this grant/during this time period), provide examples of how your open research outputs beyond articles and data have been used by others in your discipline, in other disciplines, and/or outside of academia (include DOIs, if possible)</p>	<p>The researcher cannot provide qualitative and/or quantitative evidence (for this project/period) that any of their open research outputs beyond articles and data have been used by others.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that at least one of their open research outputs (for this project/period) beyond articles and data has been used by others.</p>	<p>The researcher can provide qualitative and/or quantitative evidence that (a) some of their open research outputs beyond articles and data have been used by others; and/or (b) a narrower range of their open research outputs (for this project/period) beyond articles and data have been used deeply within a specific community.</p>
<p>For your work (related to this grant/during this time period), enumerate your plans to engage in open activities beyond sharing articles and data in the future</p>	<p>The researcher has not articulated a clear plan to make at least some research outputs beyond articles and data available openly for access and reuse.</p>	<p>The researcher has articulated a clear plan to (a) make most research outputs beyond articles and data available openly for access and reuse; and (b) engage in good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>	<p>The researcher has articulated a clear plan to (a) make all appropriate research outputs beyond articles and data available openly for access and reuse; and (b) engage in consistent good open science hygiene (e.g., use of DOIs, ORCID iDs, Creative Commons licenses).</p>

Table 8 Notes:

- The rubric can and should be adapted to reflect the questions being asked of researchers (e.g., if a grant report form does not ask about data sharing, the data sharing elements of the rubric can be excised).
- The “Reporting” language can be customized for grant reporting vs. departmental reporting.
- Researchers who generate data with personal identifiable information (PII) or other sensitive details that cannot be openly shared may indicate as such in their response.
- While the FAIR (Findable, Accessible, Interoperable, Reusable) data principles support open research, data can be FAIR without being open. The FAIR principles can accommodate legitimate exceptions to open sharing practices such as data with PII, as mentioned above.
- “Other Outputs” include a range of research products such as the National Academies Roundtable on Aligning Incentives for Open Science list enumerated in VI. Reimagining Outputs Worksheet.
- DOI – Digital Object Identifier; ORCID – Open Researcher and Contributor ID.

III. GOOD PRACTICES PRIMERS⁴

Nicholas Gibson, John Templeton Foundation
 Jerry Sheehan, National Institutes of Health
 Stuart Buck, Formerly, Arnold Ventures
 J. C. Burgelman, Vrije Universiteit Brussel
 Anne-Marie Coriat, Wellcome
 Anne Korolova, Helmsley Trust
 Heather Pierce, Association of American Medical Colleges
 Dawid Potgieter, Templeton World Charity Foundation
 Greg Tananbaum, Open Research Funders Group

Many organizations, particularly those that perform or fund research, are in the information-gathering stage with respect to their open science policies and practices. These concise primers are intended to provide decision makers with a high-level overview of the *what's* and *how's* of open sharing of various research outputs. Each primer (1–2 pages) addresses a different output type, delving into exemplars, dependencies, resourcing, and a range of other considerations. The following drafts provide a sense of what the primers will encompass. They do not provide a detailed rationale for adopting an open science policy, an analysis of the barriers, or a comprehensive guide to implementation, including the pros and cons of various approaches.

ARTICLES

Relevance to Open Ecosystem

Unrestricted access to, and reuse of, published journal articles benefits the research community by facilitating the dissemination of new information, thus maximizing opportunities for that work to lead to new insights and discoveries.

Considerations

Among the key issues that organizations will wish to address in developing a policy to make articles open are the following:

⁴The views expressed are those of the authors and do not necessarily reflect the official policies or positions of their employing organizations.

- **Fulfillment.** Can researchers adhere to the policy by publishing in a fully open access journal, a “hybrid” journal (a subscription-based journal that allows authors to make individual articles open access immediately on payment of an article publication charge), or by posting a copy of a paper in an open, trusted repository? If the latter is permissible, must a certain version (e.g., version of record, approved manuscript) be posted?
- **Timing.** Does the policy require that the articles be made openly available immediately, or is some embargo (e.g., 6 months) permissible?
- **Financial Support.** Will the policy maker provide funding to defray costs of open access (e.g., article processing charges)? If so, is there a cap on the amount? Must the researcher explicitly account for these expenses at the time of project design? Is there a mechanism for the researcher to have such costs covered after grant close?
- **Discoverability.** How will potential readers discover the openly available content? Will it be picked up by major indexing services or be made available in leading disciplinary repositories?
- **Licensing and Reuse.** What type of licensing requirements will the policy include to facilitate reuse? Free to read, preferably permanent, is often the primary focus of open access policies, but reuse considerations (including, but not limited to, text and data mining) also merit consideration.

Approaches

The practical implementation of a policy requiring access to published articles can take different forms (see Box 1). Some policies require publication in an open access journal or a hybrid journal. This can introduce a modest restriction on researchers’ choice of publication venue, although thousands of journals are open access or offer a hybrid option.

Some policies promote deposit of a copy of the paper (which may not be the final, formatted version, depending on publisher or funder requirements) in a trusted repository. As virtually all journals allow some form of self-archiving, this approach places fewer restrictions on authors (see Box 2). It does require authors to proactively identify and deposit the paper in an appropriate repository. Some journals will, however, deposit articles or final submitted manuscripts in a selected repository on behalf of authors.

SPARC (Scholarly Publishing and Academic Resources Coalition) maintains a succinct resource for tracking, comparing, and understanding

BOX 1

Examples of Open Access Policies Requiring Publication in Open Access Journals

The Bill & Melinda Gates Foundation and the Wellcome Trust require funded researchers to publish their articles in open access journals, with no embargo period.^a The option to publish in hybrid journals is being phased out by both organizations in 2021.

^a See <https://www.gatesfoundation.org/how-we-work/general-information/open-access-policy> and <https://wellcome.org/news/wellcome-updates-open-access-policy-align-coalition-s>.

BOX 2

Examples of Self-Archiving Open Access Policies

- All U.S. federal science funding agencies require submission of the author's final manuscript or final published article to a designated repository such as PubMed Central, with public access provided no later than 12 months after publication.^a
- Harvard University is among the many universities that asks faculty to deposit a version of their articles ("the accepted author manuscript") in Harvard's institutional repository.^b
- The Academic Senate of the University of California adopted a systemwide open access policy in 2013 designed to make research articles authored by faculty available to the public at no charge.^c

^a See https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf.

^b See <https://osc.hul.harvard.edu/policies>.

^c See <https://osc.universityofcalifornia.edu/for-authors/open-access-policy>.

U.S. federal funder article-sharing policies;⁵ ROARMAP (Registry of Open Access Repository Mandates and Policies) provides similar information about funders and universities;⁶ and the federal interagency group CENDI posts information about federal agency public-access policies.⁷ These sites can be used to compare and contrast different approaches that stakeholders are taking to open access policies.

Resourcing

Once open policies are implemented, organizations can undertake a range of activities to manage them. At the low-touch end of the spectrum, organizations can require researchers to document how they intend to comply. Depending on internal resources, some organizations spot-check these plans, while others simply rely on the honor system. Other organizations take a more engaged approach, requiring proof of compliance from researchers and checking this against internal expectations and guidelines. Additionally, funders are increasingly able to rely on emerging research infrastructure such as author and funder registries to automate aspects of the reporting process. Organizations without open policies may view administration and compliance as daunting tasks. However, each organization can make its own appropriate determination about the resources it is able to devote to these activities. Compliance monitoring can often be embedded within other regular research-reporting processes without adding significant burden on researchers or administrative staff.

Next Steps

The Open Research Funders Group (ORFG) can provide support and insight into best practices and available resources.⁸ The ORFG Incentivization Blueprint provides model language that can be adapted and adopted by funders and other organizations.⁹ It offers a stepwise approach to deploying a policy that can grow to encompass not only open access articles but also data, code, and other research outputs.

⁵ See <http://researchsharing.sparcopen.org/articles>.

⁶ See <https://roarmap.eprints.org>.

⁷ See https://www.cendi.gov/projects/Public_Access_Plans_US_Fed_Agencies.html.

⁸ See <http://www.orfg.org>.

⁹ See <http://www.orfg.org/incentivization-blueprint>.

DATA

Relevance to Open Ecosystem

The ability to independently confirm results and conclusions is critical for evaluating scientific rigor and informing future research activities. Openly shared data can support reanalysis and confirmation of research findings. They can also shed light on research that is not published, which can occur when tested hypotheses are not confirmed or research is considered unproductive, thereby mitigating publication bias and improving the efficiency of the research process, and can lead to novel lines of inquiry. In particular, shared data can be reused for new analyses, whether independently or in combination with other data.

Considerations

Several issues merit consideration by organizations developing open data policies, including the following:

- **Scope.** What data are needed for the independent verification of research results? Which data are most valuable to preserve for reuse? What is the appropriate balance between making available large volumes of raw data versus smaller amounts of more processed data?
- **Metadata.** What documentation and descriptive details are necessary to allow others to use the data properly and without confusion? How does the policy ensure that information about the methodology and procedures used to collect the data, details about codes, definitions of variables, variable field locations, frequencies, and the like are properly collected and disseminated? Are there disciplinary-specific metadata schemas that should be used to facilitate discovery and reuse?
- **Timing.** Starting with the baseline expectation that data underlying reported results will be made available concurrent with the posting of research findings, are there legitimate exceptions? Should researchers be given a period of exclusivity to analyze research data additional to those directly supporting reported findings before sharing them with the community? If data are

not reported in a publication, what is an appropriate time line for sharing the data?

- **Financial Support.** Who will provide funding to defray costs of preparing and/or depositing the data? What costs are recoverable? If so, is there a cap on the amount? Must the researcher explicitly account for these expenses at the time of project design?
- **Licensing.** What type of licensing requirements will the policy include to facilitate reuse of the data?
- **Proprietary Software.** To the extent that the data can only be accessed or analyzed through software that is not open source, what steps can be taken to reduce restrictions on its reuse?
- **Data Management Plans.** What support and guidance will the organization provide to help the researcher clearly articulate at the outset of a project what, how, and where data will be shared? What mechanisms are in place to ensure that the researcher adheres to the data management plan?
- **Data Standards.** For the study type in question, or for the field in which the work is centered, are there best practices for how the data should be formatted, to enable wider and more efficient reuse and interoperability?
- **Preservation.** What constitutes an appropriate deposit location for the data? Is there a repository that is appropriate for the subject matter in question, and/or has emerged within a specific research community as the default resource in that field? Is the repository secure, stable, and open for all to access?
- **Discoverability.** How will data be discoverable? Even if it is deposited in a particular repository, how will other possible users know where to look? Will the data be assigned a unique persistent identifier, and will that identifier be promulgated through related publications?
- **Privacy/Confidentiality.** Some datasets may contain human subject details that cannot be fully disseminated, due to the Health Insurance Portability and Accountability Act of 1996 (HIPAA) (Public Law 104-191; 104th Congress), the Family Educational Rights and Privacy Act (FERPA) (20 U.S.C. § 1232g; 34 C.F.R. Part 99), the European Union's General Data Protection Regulation 2016/679 (EU GDPR) (O.J. L. 119, 04.05.2016; cor. O.J. L. 127, 23.5.2018), or other privacy restrictions. Such datasets, however, can often be shared after anonymization or deidentifica-

tion techniques (including adding statistical noise, suppression of small cells, etc.), or under protected mechanisms such as a virtual data warehouse accessible only with a confidentiality agreement in place. How will such datasets be handled in a way that maximizes sharing while protecting privacy? Can analytic opportunities be made openly available while the confidential aspects of the data remain restricted?

- **Compliance monitoring.** How can compliance with data management and sharing requirements/expectations be easily monitored, for example, by funders, other institutions, or individuals?

Approaches

One common approach to facilitate data sharing is to develop policies requiring data to be findable, accessible, interoperable, and reusable, that is, to meet the Findability, Accessibility, Interoperability, and Reuse of digital assets (FAIR) data principles. While data can be FAIR without necessarily being publicly open, the FAIR principles broadly support open science. Specific definitions and operationalizations of each of these principles, together with practical guidance on how to satisfy each requirement, have been prepared by the GO FAIR Initiative.¹⁰ To render data FAIR, metadata and datasets should be prepared in a standardized, descriptive manner that makes it easier for both humans and machines to find and use.

With respect to data accessibility, a common rule of thumb in the open science community is that data should be shared in a manner that promotes reuse and transparency while recognizing that certain safeguards may be required to protect sensitive information that could compromise subject privacy or other norms and regulations. While the default position needs to shift to “open,” legitimate restrictions on access need to be taken into account.

Many U.S. federal science agencies require researchers to submit a data management plan either as part of a grant application or before issuing an award. These plans provide general information about the types of data to be collected in a research study, the repository into which they will be deposited, and the time lines and other conditions of access. For certain types of research studies, federal science agencies have developed more

¹⁰ See <https://www.go-fair.org/fair-principles>.

specific guidance or requirements (see the National Institutes of Health [NIH] example in Box 3).

Some organizations, such as the National Science Foundation, provide a general set of guidelines on data sharing, articulating to researchers that they are expected to share their data with their peers under reasonable cir-

BOX 3 **Examples of Open Data Policies**

- The National Institutes of Health (NIH) issued a Data Management and Sharing Policy that applies to all data generated by funded research^a as well as specific policies that apply to genomic data, clinical trial data, and other specific research programs and data types.^b NIH has also provided information for selecting a data repository.^c
- The American Heart Association requires grant applicants to include a data-sharing plan as part of the application process. Any research data that are needed for independent verification of research results must be made freely and publicly available within 12 months of the end of the funding period (and any no-cost extension).
- The European Open Science Cloud (EOSC) has developed a strategic implementation plan for the creation of a data commons housing interoperable, machine-readable data across domains, consistent with FAIR (findable, accessible, interoperable, and reusable) principles.^d
- The Yale University Open Data Access (YODA) Project facilitates clinical trial data access to promote independent analyses of the data. It also provides a formal vetting of the data to ensure consistency with informed consent and confidentiality requirements.^e

^a See <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html>.

^b For genomic data, see <https://osp.od.nih.gov/scientific-sharing/genomic-data-sharing>; for clinical trial data, see <https://grants.nih.gov/policy/clinical-trials/reporting/understanding/nih-policy.htm>; and for other specific research programs and data types, see https://www.nlm.nih.gov/NIHbmic/nih_data_sharing_policies.html.

^c See <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-016.html>.

^d See https://ec.europa.eu/info/research-and-innovation/strategy/goals-research-and-innovation-policy/open-science/european-open-science-cloud-eosc_en.

^e See <https://yoda.yale.edu>.

cumstances.¹¹ Others, such as the NIH, have overarching data management and sharing policies that apply to all funded research, while also having more focused policies that provide explicit guidance as to the timing, licensing, and dissemination of data of particular types (e.g., genomic data) or associated with particular research programs (e.g., the Cancer Moonshot).¹²

Resourcing

For data specifically, it is important to ensure that appropriate metadata and documentation are provided so that datasets are properly contextualized. Organizations will also benefit from in-house or outsourced expertise to assess the appropriateness of data management plans and informed consents, to ensure these allow data sharing to the extent that the organization desires.

Once open policies are implemented, organizations can undertake a range of activities to manage them. At the low-touch end of the spectrum, organizations can require researchers to document how they intend to comply. Depending on internal resources, some organizations spot-check these plans, while others simply rely on the honor system. Other organizations take a more engaged approach, requiring proof of compliance from researchers and checking this against internal expectations and guidelines. Additionally, funders are increasingly able to rely on emerging research infrastructure, such as author and funder registries, to automate aspects of the reporting process. Organizations without open policies may view administration and compliance as daunting tasks. However, each organization can make its own appropriate determination about the resources it is able to devote to these activities.

Next Steps

There are a range of resources that can contribute to a detailed understanding of policy options and approaches, including the following:

- GO FAIR provides a starter kit with a wealth of information on data management plans, license options, and repositories.¹³

¹¹ See https://www.nsf.gov/pubs/policydocs/pappg19_1/pappg_11.jsp#XID4.

¹² See <https://www.cancer.gov/research/key-initiatives/moonshot-cancer-initiative/funding/public-access-policy#requirement>.

¹³ See <https://www.go-fair.org/resources/rdm-starter-kit>.

- The Transparency and Openness Promotion (TOP) Guidelines provide sample language for three levels of open data policies.¹⁴ This wording can be adapted and adopted to suit the specific circumstances of various organizations.
- The Open Research Funders Group Incentivization Blueprint offers sample open data policy language that can be adapted for a range of use cases.¹⁵
- The American Heart Association's website contains a detailed FAQ page that articulates questions commonly asked by researchers subject to an open data policy.¹⁶
- The DMPTool site is an excellent resource for both browsing the data policies of hundreds of organizations and generating data management plans to fit a range of requirements and circumstances.¹⁷
- NIH is developing various resources to assist researchers in complying with its Data Management and Sharing Policy, including clarifications about the contents of a data management and sharing plan, selection of data repositories, and allowable costs.¹⁸

¹⁴ See <https://osf.io/bcj53>.

¹⁵ See <http://www.orfg.org/incentivization-blueprint>.

¹⁶ See <https://professional.heart.org/en/research-programs/aha-research-policies-and-awardee-hub/open-science-frequently-asked-questions#:~:text=The%20AHA%20open%20data%20policy%20requires%20any%20data%20needed%20for,but%20the%20most%20exceptional%20circumstances>.

¹⁷ See <https://dmptool.org>.

¹⁸ See <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html>.

PROTOCOLS AND PREREGISTRATION ANALYSIS PLANS

Relevance to Open Ecosystem

Unreported flexibility in data analysis can reduce the credibility of reported results and invalidate common tools of statistical inference. By submitting a detailed study protocol and statistical analysis plan to a public registry prior to conducting the work (i.e., preregistering with an analysis plan), the scientist makes a clearer distinction between planned hypothesis tests (i.e., confirmatory tests) and unplanned discovery research (i.e., screening or exploratory research). Preregistration of laboratory protocols—detailed descriptions of the methods used in the experiment, including equipment and reagents—is becoming more common and facilitates replicability. Preregistration is particularly important for studies that make an inferential claim from a sampled group or population, as well as studies that are reporting and testing hypotheses. After a project is completed, protocols and preregistration analysis plans can be used in conjunction with the final study and analysis by researchers seeking to replicate, reproduce, and build upon findings.

Considerations

- **Scope.** Should preregistration address the study protocol (how a study or experiment will be conducted), the laboratory protocol (detailed description of methods), the analysis plan (how the collected data will be organized and evaluated), or all three? Of primary interest in ensuring the integrity of the research outcome is information about the prespecified outcome measures/end-points. However, decisions made during analysis can also affect the integrity of the reported findings, so many registries encourage preregistration of both.
- **Documentation.** Should preregistration include disclosure of the full-study protocol or just summary information about the protocol and statistical analysis plan? Submission of summary information can be more time consuming, but it also allows for structured data entry to facilitate searching and cross-study comparison. If a summary is submitted, then what specific information needs to be provided?
- **Data Privacy.** Protocols and analysis plans can contain proprietary or other protected information (e.g., names of study person-

nel). To what extent can information be redacted without undermining the benefits of access? The desire to promote meaningful preregistration must be balanced against the provision of necessary protections/redactions of information.

- **Deposit Location.** Where and how should a scientist register their protocol and/or analysis plan? There are a limited number of established public repositories. For clinical trials of health-related interventions, NIH's ClinicalTrials.gov is the default system.¹⁹ Within the social, behavioral, and preclinical sciences, the Open Science Framework is becoming a default registry.²⁰ Some public repositories tend to be disciplinarily focused.
- **Timing.** How long before or after a study begins must it be registered? When should a preregistration be updated? Earlier may be better, but additional information may be needed about its status (e.g., has Institutional Review Board approval been received). The timing of an update is also linked to the degree to which a change has implications on the full preregistration (e.g., challenges in recruiting a full sample may necessitate moving from a single cohort to a multicohort design). Protocols shared at study initiation can more clearly establish a project's aims and plan. Does the registry support time-stamped versioning?
- **Discoverability.** Are preregistrations automatically made public after a fixed period of time? Does the registry support public searches for preregistrations?
- **Scope.** To date, the majority of registries are for causal impact studies, typically carried out either in a small-scale experiment or a large randomized clinical/field trial. However, there may be a strong rationale to consider preregistering exploratory studies at the time of funding or at the beginning of a study so as to capture strong theory-driven exploratory questions as opposed to post hoc "fishing" analyses.
- **Results.** To what extent should a funder require the ultimate posting of a study's results in a way that can be compared to whatever was preregistered? Federal law requires the posting of results at ClinicalTrials.gov for certain clinical trials; should this be a broader expectation?

¹⁹ See <https://clinicaltrials.gov>.

²⁰ See <https://osf.io>.

Approaches

There are a range of different preregistration locations available, primarily driven by discipline. All NIH-funded clinical trials and most clinical trials of Food and Drug Administration (FDA) regulated drugs, biologics, and devices must be preregistered at NIH's ClinicalTrials.gov not later than 21 days after first recruitment. Summary information is provided in a highly structured format. Final protocols for NIH-funded clinical trials and most FDA-regulated clinical trials of drugs, biologics, and devices must be submitted to NIH's ClinicalTrials.gov as part of summary data reporting after a trial has completed. These policies also require that the statistical analysis plan be submitted if it is not considered part of the protocol. (See Box 4 for examples of preregistration and protocols policies.)

BOX 4 Examples of Preregistration and Protocols Policies

- The Chan Zuckerberg Initiative (CZI) requires grantees to make laboratory protocols publicly available and has nurtured dedicated protocol communities of CZI-funded investigators.^a
- The American Economic Association (AEA) encourages researchers to register their randomized controlled trials (including research designs and analysis plans) in the AEA Randomized Controlled Trials (RCT) Registry.^b
- CHDI Foundation has established an Independent Statistical Standing Committee (ISSC) to provide unbiased evaluation and expert advice on developing protocols and statistical analysis plans, and evaluation of prepared study protocols.^c
- Arnold Ventures requires all funded empirical studies that involve statistical inference to be preregistered before the start of intervention or data collection on the Open Science Framework.^d

^a See <https://cziscience.medium.com/power-to-the-protocols-388fe92001be> and <https://www.protocols.io/workspaces/neurodegeneration-method-development-community1>.

^b See <https://www.socialscienceregistry.org>.

^c See <https://chdifoundation.org/independent-statistical-standing-committee>.

^d See <https://www.arnoldventures.org/guidelines-for-investments-in-research> and <https://osf.io>.

Other disciplines have their own community-promoted repositories. Researchers carrying out causal studies in education have the opportunity to preregister their work in the Registry of Efficacy and Effectiveness Studies.²¹ Researchers in the social, behavioral, and cognitive sciences often use the Open Science Framework platform.²² The Registry for International Development Impact Evaluations hosts impact evaluations related to development in low- and middle-income countries.²³

Resourcing

Organizations considering preregistration will need to consider whether resources are needed to support a preregistration repository for collecting preregistration reports and protocols. It is also important that there is a transparent link among any disseminated findings (preprints, articles, etc.), data, and preregistrations to determine whether there are significant deviations from the intended analysis.

Organizations and publishers will also need to ascertain how to indicate where preregistration records and protocol information exist for a published article. Multiple publishers and other organizations offer modalities for publishing study protocols, laboratory protocols, and registered reports. To be most effective, preregistrations and protocols should be closely linked to associated publications and other study information so that they can be easily discovered and accessed by those examining the study results.

Next Steps

The TOP Guidelines provide sample language for three levels of policies for study preregistration and analysis plan preregistration.²⁴ This wording can be adapted and adopted to suit the specific circumstances of a range of organizations. The TOP recommendations include (1) disclosing whether work was preregistered or not, (2) verifying that any preregistered work adheres to the prespecified plans, and (3) requiring preregistration for relevant research studies (typically inferential and hypothesis-testing work).

²¹ See <https://sreereg.icpsr.umich.edu/sreereg>.

²² See <https://osf.io/prereg>.

²³ See <https://ridie.3ieimpact.org>.

²⁴ See <https://osf.io/bcj53>.

The Center for Open Science provides multiple resources on how to preregister studies and analytic plans, including templates.²⁵ NIH provides a number of resources to facilitate the development of protocols, including the National Institutes of Health e-Protocol Writing Tool and protocol templates for clinical trials and behavioral/social science research.²⁶

²⁵ See <https://www.cos.io/initiatives/prereg> and <https://osf.io/zab38/wiki/home/?view>.

²⁶ See <https://e-protocol.od.nih.gov/#/home> and <https://grants.nih.gov/policy/clinical-trials/protocol-template.htm>.

REGISTERED REPORTS

Relevance to Open Ecosystem

Peer review of study protocols with analysis plans, along with dissemination of findings regardless of outcome, addresses publication bias against null results. It also provides the benefits of preregistration by making a clearer distinction between hypothesis tests and discovery research. By submitting funded studies to journals as a registered report, the scientist improves study planning, increases study rigor, and improves scientific credibility. Funders who support this process anticipate that peer-review feedback could change study processes that result in budget changes and are prepared to consider such amendments in response to journal reviewer feedback. Funders can also partner with journals to coordinate review for funding and publishing decisions.

Considerations

- **Scope.** Registered reports are most appropriate for specific experiments or studies, not for grants that fund a research program over several years. Such grants could still include one or more registered reports, but it would likely not cover the entire program.
- **Research Scope.** Registered reports are best for studies that test hypotheses and in disciplines that could suffer from publication bias (typically against null results). Registered reports are not appropriate for purely exploratory or discovery science, until those studies are ready to use traditional hypothesis tests.
- **Timing.** By design, registered reports include additional time at the beginning of a project. Project plans should account for this. Additional time devoted to peer review in the early stages of the project is also required to ensure that the study methods are as rigorous as possible and that results will be disseminated regardless of outcome.

Approaches

There are a number of ways in which an organization can promote registered reports. On the low end of engagement, a funder or agency can ask grantees to specifically state whether all or part of the work would be

BOX 5
**Examples of Funders Encouraging/
Requiring Registered Reports**

- The Flu Lab is partnering with PLOS and the Center for Open Science to promote replications and registered reports of influenza research.
- Cancer Research UK is collaborating with the journal *Nicotine & Tobacco Research* on an integrated review process for grant proposals and preregistered reports.^a

^aSee <https://academic.oup.com/ntr/article/19/7/773/3106460>.

appropriate for a registered report. This will remind grantees that registered reports are a valued addition to a proposed study. Principal investigators can be encouraged to notify their communities—via social media, their websites, CVs, and other appropriate channels—when their precollection hypotheses and data analysis plans have been reviewed and registered. Organizations may also wish to educate researchers on the benefits of registered reports, particularly researchers in domains where the practice is not currently widespread.

For specific grants, programs, or initiatives where projects are appropriate for the format, agencies and funders may elect to make registered report submissions to a journal before data collection a requirement. If a study does not receive an in-principle acceptance offer from a journal, the plan can still be preregistered by the authors on a platform like the Open Science Framework and submitted for publication after the study is completed.

Some funders are partnering directly with discipline-appropriate journals to integrate the registered reports model in the grant application process. One example is the Children's Tumor Foundation,²⁷ which is partnering with the journal *PLOS ONE* to concurrently evaluate grant proposals and the ethics and rigor of the experimental design. Accepted proposals will simultaneously receive both funding and a commitment to publication of the study results in *PLOS ONE*. (See Box 5.)

²⁷ See <https://grants.nih.gov/policy/clinical-trials/protocol-template.htm>.

Resourcing

Given the relative novelty of registered reports, organizations may need to educate grantees about the merits and mechanics of this approach. Organizations that seek to integrate grant proposals and registered reports will need to establish a review process that allows for independent evaluation of the latter along a timescale and workflow that supports the former. This may also require negotiation of a direct partnership with a journal or publisher.

Absent this type of embedded relationship, researchers may require guidance to evaluate the growing number of journals that accept and publish registered reports. The Comparison of Registered Reports site provides an interactive tool to assist in this process.²⁸ Policies that require registered reports will also require some form of monitoring, ranging from spot-checking to soliciting proof of compliance.

Next Steps

The Center for Open Science provides a comprehensive registered reports resource,²⁹ including FAQs, workflow suggestions, and other foundational materials. The Center for Open Science also provides a simple Q&A tutorial to assist authors in the drafting of registered reports.³⁰ The Open Science Framework provides a searchable database of registered reports across a range of disciplines.³¹ These may offer useful guidance to better understand the core elements of a well-constructed registered report.

²⁸ See https://katedrax.shinyapps.io/cos_registered_reports.

²⁹ See <https://www.cos.io/initiatives/registered-reports>.

³⁰ See https://osf.io/93znh/?_ga=2.100491997.298846709.1580837996-1159488863.1580234077.

³¹ See <https://osf.io/registries/discover?provider=OSF&type=Registered%20Report%20Protocol%20Preregistration>.

SOFTWARE AND CODE

Relevance to Open Ecosystem

Research projects may generate code that is used as a means to run, analyze, or interpret research data. The ability to independently confirm results and conclusions is critical for evaluating scientific rigor and informing future research activities. To extract maximum value from research findings and available data, any code deployed to process these data must therefore be widely and freely available. Research findings are not fully open unless the tools necessary to understand and test them are also made available. Research projects may also generate software that is the product of the project rather than the byproduct, a specified deliverable designed to perform a specific task. Making the underlying code for this type of research output open source can encourage collaboration, further development, community engagement, and enhanced return on funders' investment.

Considerations

As organizations develop open science policies pertaining to code and software, among the issues they must consider are the following:

- **Software/Code Maintenance.** What are the expectations for the duration and extent to which code should be kept up to date? Should the version used to produce the reported findings be maintained?
- **Proprietary Software.** To the extent that some or all of the code base upon which an experiment relies is not open source, what steps can be taken to reduce restrictions on its reuse?
- **Timing.** Does the policy require that the code or software be made openly available immediately upon the posting of research findings (e.g., publication of an article, deposit of a dataset), or is some embargo (e.g., 6 months) permissible? If research findings are not published or posted, should code and software be made publicly available no later than grant close?
- **Financial Support.** Will the policy maker provide funding to defray costs of preparing and/or depositing the code or software? If so, is there a cap on the amount? Must the researcher explicitly account for these expenses at the time of project design? If code

or software is made publicly available after the conclusion of the grant, does the grantee have a mechanism to request additional financial support?

- **Licensing.** What type of licensing requirements will the policy include to facilitate reuse? Do the grantee and/or the funder retain any stake in the intellectual property?
- **Metadata.** What documentation and descriptive details are needed to understand and execute the code or run the software program? How will the computational environment in which software or code was originally executed be described and archived? Should researchers establish virtual environments (e.g., Docker)?
- **Preservation.** What constitutes an appropriate deposit location for the code or software? Is there a repository that is appropriate for the subject matter in question and/or has emerged within a specific research community as the default resource in that field? Is the repository secure, stable, and open for all to access? Does the repository assign persistent digital identifiers to code?

Approaches

The TOP Guidelines advise that researchers should “provide program code, scripts for statistical packages, and other documentation sufficient to allow an informed researcher to precisely reproduce all published results ... through a trusted digital repository.”³² More funder-specific TOP guidance may be found at <https://www.cos.io/initiatives/top-funders>.

Some agencies within the U.S government use open source code as a matter of policy. For example, the Consumer Financial Protection Bureau unequivocally states, “When we build our own software or contract with a third party to build it for us, we will share the code with the public at no charge.”³³ Other agencies, such as the Department of Education, make the source code for their prominent public-facing initiatives (in ED’s case, the College Scorecard)³⁴ openly available. Both of these organizations deposit these research outputs (software as a product, not a byproduct, of the grant) on GitHub. When code is developed to interpret or analyze research findings (code as a secondary output of the grant), organizations such as

³² See <https://osf.io/bcj53>.

³³ See <https://www.consumerfinance.gov/about-us/blog/the-cfpbs-source-code-policy-open-and-shared>.

³⁴ See <https://collegescorecard.ed.gov>.

BOX 6

Examples of Open-Code and Software Policies

- NASA's Earth Science Data Systems (ESDS) Program requires that all software developed through research and technology awards be made available to the public as open source.^a All funding proposals must include software development plans that are vetted as part of the application process.
- The U.S. government's Federal Source Code Policy includes a pilot program that "requires agencies, when commissioning new custom software, to release at least 20 percent of new custom-developed code as Open Source Software for three years."^b
- Several learned societies that publish flagship disciplinary journals, including the American Geophysical Union and the American Astronomical Society, require or strongly encourage authors to make openly available any code used to generate results or analyses reported in their papers.^c

^a See <https://earthdata.nasa.gov/collaborate/open-data-services-and-software/esds-open-source-policy>.

^b See <https://www.cio.gov/2016/08/11/peoples-code.html>.

^c See <https://www.agu.org/Publish-with-AGU/Publish/Author-Resources/Policies/Data-policy> and <https://journals.aas.org/news/policy-statement-on-software>.

the Wellcome Trust typically require the code to be shared at the time the primary research is published.³⁵ (See Box 6 for examples of open-code and software policies.)

Resourcing

For code specifically, some technical expertise may be required to ensure that the code and software are operable and can be accessed and used by the wider community.

Once open policies are implemented, organizations can undertake a range of activities to manage them. At the low-touch end of the spectrum, organizations can require researchers to document how they intend to

³⁵ See <https://wellcome.org/news/our-new-policy-sharing-research-data-what-it-means-you>.

comply. Depending on internal resources, some organizations spot-check these plans, while others simply rely on the honor system. Other organizations take a more engaged approach, requiring proof of compliance from researchers and checking this against internal expectations and guidelines.

Next Steps

The TOP Guidelines provide sample language for three levels of open-code policies.³⁶ This wording can be adapted and adopted to suit the specific circumstances of a range of organizations. For a deeper dive into policy formulation, interested parties can download the National Academies of Sciences, Engineering, and Medicine's report *Open Source Software Policy Options for NASA Earth and Space Sciences*.³⁷ This comprehensive document provides a deep dive into the established approaches, best practices, and practical considerations that can help effectively shape an open code policy.

³⁶ See <https://osf.io/bcj53>.

³⁷ See <https://www.nap.edu/catalog/25217>.

IV. OPEN SCIENCE BY THE NUMBERS INFOGRAPHIC

Open Science by the Numbers

Open Science posits that research has its widest impact and is most trustworthy when all of its elements (including articles, data, protocols, and code) can be openly accessed, tested, and built upon.



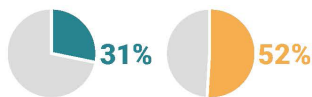
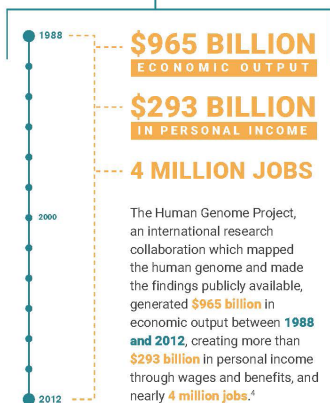
Researchers estimate that **\$3.2 trillion** in economic output could be added to global GDP through Open Data across all sectors, with scientific and scholarly data playing an important role.¹



The Symbiota open source platform, funded by NSF, hosts **37 million biological records** from **766 universities, museums, and research organizations**.²



NIH The Human Genome Project timeline



In 2019, **31% of all journal articles were available as Open Access**, and **52% were viewed Open Access articles**.⁵

Authors: Derrick Anderson, Arizona State University; Rachel Bruce, UK Research and Innovation; Ashley Farley, Bill & Melinda Gates Foundation; Robert Hanisch, National Institute of Standards and Technology; Greg Tarantoum, Open Research Funders Group; Thomas Wang, American Heart Association/University of Texas Southwestern Medical Center. The views expressed are those of the authors and do not necessarily reflect the official policies or positions of their employing organizations.

Source:
1. https://www.oandp.com/sites/default/files/019_archive/insights/01920Report_061114_FINAL.pdf
2. <https://ymbiota.org/oc/>
3. <https://www.usda.gov/media/blog/2018/07/13/open-data-enabling-fact-based-data-driven-decisions>

4. <https://spar.open.org/impact-story/human-genome-project/>
5. <https://www.biorxiv.org/content/10.1101/179531v1>

V. OPEN SCIENCE SUCCESS STORIES DATABASE³⁸

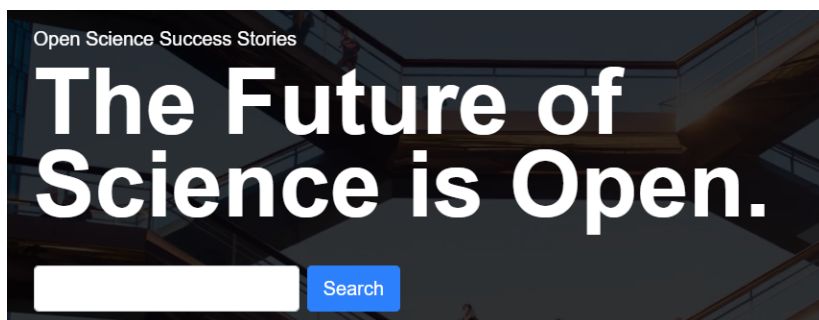
Derrick Anderson, Arizona State University
Greg Tananbaum, Open Research Funders Group

The Open Science Success Stories Database compiles articles, perspectives, case studies, news stories, and other materials that demonstrate the myriad ways in which open science benefits researchers and society alike.³⁹

Scientists, scholars, librarians, department chairs, university administrators, philanthropic program officers, government agency representatives, policy makers, publishers, journalists, and other stakeholders can use the curated resources to understand how open science is positively influencing specific disciplines and communities, as well as how these lessons can be applied to the global scientific endeavor.

The database is being developed by Arizona State University in collaboration with the Open Research Funders Group. An initial version is being made available as part of the background material for the November 5, 2020, National Academies workshop on Developing a Toolkit for Fostering Open Science Practices.

Open Science Success Stories Database:
<https://projectopen.io/>



³⁸ The views expressed are those of the authors and do not necessarily reflect the official policies or positions of their employing organizations.

³⁹ See <https://projectopen.io>.

VI. REIMAGINING OUTPUTS WORKSHEET⁴⁰

Boyana Konforti, Formerly, Howard Hughes Medical Institute

Elizabeth Albro, U.S. Department of Education

Anurupa Dev, Association of American Medical Colleges

Josh Greenberg, Alfred P. Sloan Foundation

Ross Mounce, Arcadia Fund

Brian Quinn, Robert Wood Johnson Foundation

Greg Tananbaum, Open Research Funders Group

Richard Wilder, Coalition for Epidemic Preparedness Innovations

The following table (organized alphabetically) represents the authors' perspective about the range of research products that should be accounted for as the science community thinks about the behaviors and activities that should be rewarded. What are the outputs that are consistent with the values the science community collectively espouses? What outputs encourage open dialog and the tackling of big questions, build upon and enhance the work of others, and advance the research endeavor? As the community enumerates these research products, what considerations must be contemplated and addressed to create appropriate alignment between values and activities? The authors believe it will be crucial to ensure that the science community takes an expansive view of the types of research products that should be "open"—available for access and reuse without gatekeeping or payment.

⁴⁰The views expressed are those of the authors and do not necessarily reflect the official policies or positions of their employing organizations.

Reimagining Outputs Worksheet Table

Research Output Type	Exemplar Open Practices	Importance to Open Ecosystem	Concerns/Considerations
Articles	All primary research articles should be made immediately available (open access with no embargo period) and reusable via an expansive license such as CC BY.	Unrestricted access to, and reuse of, published articles benefit the research community by facilitating the discovery of new information, thus maximizing opportunities for that work to lead to new insights and discoveries.	<ul style="list-style-type: none"> Free to read is often the primary focus of open access policies, but reuse considerations (including, but not limited to, text and data mining) also merit consideration. Distinctions between versions (version-of-record accepted manuscript) may be more important within certain disciplines. Stewardship/ownership of repositories—ensuring these are open and sustainable.
Code and Software	To the greatest extent allowable by copyright, all software, code, lab notebooks, and executables necessary to independently verify research results should be curated and made freely available in an open repository no later than the publication of the first paper running this code.	The independent confirmation of results and conclusions is critical for understanding scientific soundness and informing future research activities. To extract maximum value from research findings, both the raw data that underpin the results and any code deployed to process these data must be widely and freely available to any interested party. Succinctly, research findings are not fully open unless the tools necessary to understand and test them are also made available.	

<p>Commentaries and Analyses</p>	<p>Commentaries, analyses, and other summary works that place research developments into context should be made immediately available (open access with no embargo period) and reusable via an expansive license such as CC BY.</p>	<p>With millions of research articles published annually, the need for filtering, selection, and curation has never been greater. Commentaries and analyses, including (but not limited to) review articles and research summaries, provide context for the findings described in primary articles. These materials extend the utility of primary research and widen the prospective audience to include policy makers and the general public.</p>	<ul style="list-style-type: none"> • Commentaries and summaries are an important way for learned societies to add value and continue to earn some subscription income.
<p>Data</p>	<p>Subject to personal privacy, regulatory, and legal restrictions, data underlying specific claims in a research project should be deposited with the necessary metadata into a repository, with efforts taken to maximize findability, accessibility, interoperability, and reuse. Deposits should be made no later than the publication of the first paper based on the data. Data should be considered legitimate, citable products of research.</p>	<p>The independent confirmation of results and conclusions is critical for understanding scientific soundness and informing future research activities. Openly shared data can shed light on negative results and attempted research directions, with the potential to improve efficiency of the research process as well as lead to novel analyses and conclusions.</p>	<ul style="list-style-type: none"> • Stewardship/ownership of repositories—ensuring these are open and sustainable. • Timing of data release. • Restrictions on data reuse (e.g., text and data mining).

continued

Table Continued

Research Output Type	Exemplar Open Practices	Importance to Open Ecosystem	Concerns/Considerations
Digital Scholarship	Multimedia, digital media, and audiovisual outputs should be made immediately available (open access with no embargo period) and reusable via an expansive license such as CC BY.	Digital scholarship encompasses a range of research outputs in several disciplines (particularly in the humanities). These materials are critical to the scholarly record, particularly when they are made available under a license that permits reuse and remixing.	<ul style="list-style-type: none"> Stewardship/ownership of repositories—ensuring these are open and sustainable. Ensuring that materials are “future proofed” and viable for access and reuse for an extended period of time.
Monographs, Books, Book Chapters, and/or Edited Volumes	All monographs, books, book chapters, and/or edited volumes should be made immediately available (open access with no embargo period) and reusable via an expansive license such as CC BY.	Unrestricted access to, and reuse of, monographs, books, book chapters, and/or edited volumes benefits the research community because it facilitates the discovery of new information, and thus maximizes opportunities for that work to lead to new insights and discoveries.	<ul style="list-style-type: none"> Open access for books and longer form content is less developed than journals. Few options/models.
Non-Peer-Reviewed Reports, Posters, and Presentations	All non-peer-reviewed outputs that are appropriate to be shared with the research community (e.g., reports and presentations) should be made immediately available (open access with no embargo period) and reusable via an expansive license such as CC BY.	Unrestricted access to, and reuse of, non-peer-reviewed outputs benefits the research community because it facilitates the discovery of new information, and thus maximizes opportunities for that work to lead to new insights and discoveries.	<ul style="list-style-type: none"> Grantees/faculty members may require additional guidance as to what constitutes an appropriate research output.

Peer Reviews

Peer reviews should be published with the article (so-called open reports). They can be anonymous or not. The author's response to the reviews should be published as well.

Publishing referee reports makes the process more transparent. Peer reviews contain arguments and ideas that can reveal how thinking in a field evolves. This material should be preserved and made available to others. Additionally, readers have a right to understand the level of scrutiny that a paper has undergone, and it provides them with a window into the editorial process. Because peer reviews are an essential component of the research endeavor, publishing referee reports helps create a pathway for formally crediting this activity.

- Infrastructure limitations. Right now, less than 3 percent of scientific journals allow peer reviews to be published.
- Ownership considerations. Who has the right to disseminate referee reports? Authors? Reviewers? Publishers?
- Providing credit for peer reviews without compromising anonymity (see ORCID *PLOS*⁶ collaboration).
- Several initiatives are emerging to support peer-review experiments. For example, ASAPbio has launched ReimagineReview a directory of peer-review trials, inside and outside the journal system.⁶

continued

Table Continued

Research Output Type	Exemplar Open Practices	Importance to Open Ecosystem	Concerns/Considerations
Preprints	<p>Scientists should share preprints (paper drafts that have not yet been peer reviewed for formal publication) by posting in a repository or preprint server that codifies free, unrestricted, and perpetual access to the preprint.</p> <p>Preprints should be posted in a timely manner, ideally at the time of first submission to a journal.</p>	<p>Preprints allow research findings to be quickly and easily available to all and allows researchers to claim priority of discovery, receive community input, and demonstrate evidence of progress for funders and others.</p>	<ul style="list-style-type: none"> The growing visibility of preprints may render double-blind peer review more challenging, as prospective referees are exposed to preprints (and their authors) prior to the journal submission and review stages. Several initiatives are emerging to support preprints. For example, ASAPbio.org is a comprehensive resource for information on preprints, peer reviews, transparency, and so forth.^c Transpose is a directory of journal policies, co-reviewing, and preprints.^d

<p>Preregistration Analysis Plans</p>	<p>Indicate in grant proposals, progress reports, and published articles of funded research that the research will be preregistered with an analysis plan. Provide a URL link to preregistration in reports and articles when completed. When results are reported, make a clear distinction between the planned research and any unplanned research or analysis that was conducted. Disclose any deviations from the planned procedures.</p>	<p>Unreported flexibility in data analysis decreases scientific credibility and invalidates common tools of statistical inference. By submitting a detailed study protocol and statistical analysis plan to a registry prior to conducting the work (i.e., preregistering with an analysis plan) the scientist makes a clearer distinction between planned hypothesis tests (i.e., confirmatory tests) and unplanned discovery research (i.e., screening or exploratory research). Preregistration is particularly important for studies that make an inferential claim from a sampled group or population, as well as studies that are reporting hypotheses.</p>	<ul style="list-style-type: none"> • May not be appropriate for all types of research, such as studies that do not claim to make inferences, that are purely discovery, that do not test hypotheses, or that generate computational models. • Also see the Registered Reports section, below, as a way to practice prospective registration.
<p>Protocols</p>	<p>Descriptions of the design and implementation of experiments should be made freely available in an open repository that facilitates the sharing, editing, forking (copying and adopting/modifying), and further development. These include study protocols (description of the study plan), and laboratory protocols (detailed description of experimental methods).</p>	<p>Understanding the starting point for work—including assumptions—along with the final study and analysis can provide guidance to other researchers as to additional research avenues to explore. Protocols provide the context to interpret and understand how research results are derived. They can convey exactly what was done and the decisions/compromises that were made on route to a scientific discovery.</p>	<ul style="list-style-type: none"> • Protocols can be shared prior to conducting work, which provides insights into research that does not ultimately get published; this is uncommon at present.

continued

Table Continued

Research Output Type	Exemplar Open Practices	Importance to Open Ecosystem	Concerns/Considerations
Registered Reports	Indicate in grant proposals, progress reports, and published articles which parts of the funded research will be submitted as a registered report. In project time-line documentation, add the appropriate time (e.g., 2 to 4 months) for the peer-review process at the beginning of the relevant project phases. Communicate with the funder on any procedural changes that occur as a result of peer-review feedback. If the funder partners with journals to combine reviewer feedback to jointly offer funding and publishing, submit to such solicitations.	Peer review of study protocols with analysis plans, along with dissemination of findings regardless of outcome, addresses publication bias against null results. It also provides the benefits of preregistration by making a clearer distinction between hypothesis tests and discovery research. By submitting funded studies to journals as a registered report, the scientist improves study planning, increases study rigor, and improves scientific credibility. Funders who support this process anticipate that peer-review feedback could change study processes that result in budget changes and are prepared to consider such amendments in response to journal reviewer feedback. Funders can also partner with journals to coordinate review for funding and publishing decisions.	<ul style="list-style-type: none"> • In some disciplines and in some types of research, infrastructure (including, but not limited to, participating journals) to support registered reports activities is limited. • May not be appropriate for all types of research, such as studies that do not claim to make inferences, that are purely discovery, that do not test hypotheses, or that generate computational models.

continued

<p>Research Materials</p>	<p>Biological and other physical samples (in particular starting materials), research tools (including reagents, animal models, and the like), and other materials (including metadata and identifiers) necessary to reproduce or extend research findings should be made freely available in an open repository no later than the publication of the first paper based on the materials.</p>	<p>Similar to code and data, it allows the independent confirmation of results. Also similar to code and data, broader access to research materials can accelerate research more broadly and allow comparisons across research project or products. Biological materials, such as cell lines, are fundamentally different from data and even software as they may embody a type of “machine” that, through cell expression and the like, can be used to make desirable products, such as a particular valuable protein.</p>	<ul style="list-style-type: none"> • Cost of maintaining/sharing certain types of samples and quality control. • Stewardship/ownership of repositories—ensuring these are open and sustainable, including detailed descriptions of samples.
---------------------------	---	---	---

Table Continued

Research Output Type	Exemplar Open Practices	Importance to Open Ecosystem	Concerns/Considerations
Theses and Dissertations	All theses and dissertations should be made available (open access with as short an embargo period as possible) and reusable via an expansive license such as CC BY.	Theses and dissertations represent significant contributions to the advancement of knowledge and the scholarly record. The open sharing of these materials offers a particularly unique insight into the research perspective of the emerging generation of scholars.	<ul style="list-style-type: none"> Because students often try to publish portions of their theses and dissertations as articles, and because some journals still consider posted electronic theses/dissertations to be “prior publication,” a reasonable embargo period may be both necessary and appropriate. An embargo of substantial length may create an added burden if the author has graduated and left the institution. Authors of theses and dissertations that disclose a novel process or invention for which a patent may be sought may require longer embargoes. Some theses and dissertations incorporate works of other copyright owners; this may require additional intellectual property guidance.

^a See <https://theplosblog.plos.org/2019/06/youve-completed-your-review-now-get-credit-with-orcid>.

^b See <https://reimaginerreview.asapbio.org>.

^c See <https://asapbio.org>.

^d See <https://transpose-publishing.github.io/#>.

NOTE: CC BY – Creative Commons Attribution License.