The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Developing a Toolkit for Fostering Open Science Practices: A Workshop
November 5, 2020

AGENDA BOOK CONTENTS

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Developing a Toolkit for Fostering Open Science Practices: A Workshop

Roundtable on Aligning Incentives for Open Science

November 5, 2020 (Thursday)

DRAFT AGENDA

Join from PC, Mac, Linux, iOS or Android:
https://nasem.zoom.us/j/97365225047?pwd=dEFqVzN6cDExbVNQSytaMEVQV2RrUT09
Password: 709837
Or Telephone: US: +1 312 626 6799 or +1 470 250 9358
Meeting ID: 973 6522 5047; Password: 709837
International numbers available: https://nasem.zoom.us/u/abY1pNV1pB

PUBLIC WORKSHOP

11:00 AM EST **Opening Remarks and Goals for the Workshop from Roundtable Co-Chairs and Planning Committee Chair**

Keith Yamamoto, University of California, San Francisco; Loretta Parham, Atlanta University Center; Tom Kalil, Schmidt Futures

11:15 AM **Adopting and Utilizing a Toolkit for Open Science: Stakeholder Perspectives**

University Perspectives (2 x 10 minute talks and 5 minutes discussion)

Thought Leaders: Michael Crow, President, Arizona State University; and Philip E. Bourne, Founding Dean, University of Virginia School of Data Science
Researchers: Tatiana Bryant, Research Librarian for Digital Humanities, History, and African American Studies at University of California-Irvine; and Camille Thomas, Technology and Digital Scholarship Scholarly Communications Librarian, Florida State University

Disciplinary Perspectives (2 x 10 minute talks and 5 minutes discussion)

Thought Leader: Lauren Collister, Chair of Committee on Scholarly Communication in Linguistics, Linguistic Society of America
Researcher: Sanjay Srivastava, University of Oregon Department of Psychology Undergraduate Education Chair

Research Funder Perspectives (2 x 10 minute talks and 5 minutes discussion)

Thought Leader: Ekemini Riley, Managing Director, Aligning Science Across Parkinson's
Researcher: Julia Stewart Lowndes, Senior Fellow at the National Center for Ecological Analysis and Synthesis (NCEAS) of University of California-Santa Barbara and founder of Openscapes

12:35 PM **Break and Log-in to Breakout Sessions**

12:45 PM **Stakeholder Breakout Sessions**

1:30 PM **Breakout Reports and Concluding Discussion**

2:00 PM **Adjourn Public Workshop**

Developing a Toolkit for Fostering Open Science Practices: A Workshop

Roundtable on Aligning Incentives for Open Science
National Academies of Sciences, Engineering, and Medicine

Online Workshop
(Please register to receive information on how to participate)

November 5, 2020

Statement of Task: A committee will organize a public workshop to discuss the information and resource needs of researchers, research institutions, research funders, professional societies, and other stakeholders interested in fostering open science practices. Workshop participants will also discuss approaches to meeting those needs, such as development of a toolkit that could be used by various groups of stakeholders. The workshop will be held in conjunction with the Fall 2020 meeting of the Roundtable on Aligning Incentives for Open Science. The workshop will include presentations of commissioned papers that describe and provide examples of possible elements of a toolkit, a discussion of how available information resources are currently being used, additional resources that could be developed, and future use scenarios aimed at achieving closer alignment across research values, practices, and incentives related to open science. The workshop discussion will also explore possible mechanisms for disseminating a toolkit and other information resources. A Proceedings of a Workshop summarizing the discussion and including the individually-authored papers will be prepared by a designated rapporteur and distributed broadly.

Committee Membership Information

Loretta Parham (Chair), Chief Executive Officer and Library Director, Robert W. Woodruff Library, Atlanta University Center
Stuart Buck, Vice President of Research, Arnold Ventures
Sarah Nusser, Professor, Department of Statistics, Iowa State University
Roger Wakimoto, Vice Chancellor for Research, University of California, Los Angeles

Developing a Toolkit for Fostering Open Science Practices: A Workshop

Roundtable on Aligning Incentives for Open Science

November 5, 2020 (Thursday)

Speaker Biographical Information

LORETTA PARHAM (Workshop Chair) is Chief Executive Officer and Library Director of the Robert W. Woodruff Library of the Atlanta University Center, an independent entity operating as the single library shared by its four member institutions—Clark Atlanta University, the Interdenominational Theological Center, Morehouse College and Spelman College. She is responsible for the strategic agenda transforming the Woodruff Library into the best choice for information for its community. With more than 30 years in the profession her experience includes: Director of the Hampton University Library, Deputy Director of the Carnegie Library of Pittsburgh, Pa., District Chief of the Chicago Public Library and other public service positions with the Chicago Public Schools and the City Colleges of Chicago. An active leader, scholar and engaging speaker, Ms. Parham was named a 2004 "Mover & Shaker" by Library Journal and was also honored as the 2017 Academic/Research Librarian of the Year by the Association of College and Research Libraries (ACRL). She has authored articles on Historically Black Colleges and Universities (HBCU) libraries and archives, and is co-editor of the book, Achieving Diversity: A How-To-Do-It Manual for Librarians. Ms. Parham is co-founder and past chair of the HBCU Library Alliance, past chair of the Georgia Humanities Council, former board member of ACRL, the Wayne State University School of Library Science Advisory Board, and past treasurer of the Oberlin Group. Ms. Parham holds the Master of Library Science degree from the University of Michigan-Ann Arbor and the Bachelor of Science in Communications from Southern Illinois University-Carbondale. She is currently serving on Board of Directors of EDUCAUSE and EDUCOPIA.

THOMAS KALIL (Roundtable Co-Chair) is Chief Innovation Officer at Schmidt Futures. In this role, Mr. Kalil leads initiatives to harness technology for societal challenges, improve science policy, and identify and pursue 21st-century moonshots. Prior to Schmidt Futures, Mr. Kalil served in the White House for two Presidents (Obama and Clinton), helping to design and launch national science and technology initiatives in areas such as nanotechnology, the BRAIN initiative, data science, materials by design, robotics, commercial space, high-speed networks, access to capital for startups, high-skill immigration, STEM education, learning technology, startup ecosystems, and the federal use of incentive prizes. From 2001 to 2008, Mr. Kalil was Special Assistant to the Chancellor for Science and Technology at University of California, Berkeley. Mr. Kalil received a B.A. in political science and international economics from the University of Wisconsin at Madison and completed graduate work at the Fletcher School of Law and Diplomacy.

KEITH R. YAMAMOTO (Roundtable Co-Chair) (NAS/NAM) is University of California, San Francisco (UCSF) vice chancellor for science policy and strategy, director of precision medicine for UCSF, and professor of cellular and molecular pharmacology at UCSF. He is a leading researcher investigating transcriptional regulation by nuclear receptors, which mediate the actions of essential hormones and cellular signals; he uses mechanistic and systems approaches to pursue these problems in pure molecules, cells and whole organisms. He has led or served on numerous national committees focused on public and scientific policy, public understanding and support of biological research, and science education; he chairs the Coalition for the Life Sciences, and sits on the National Academy of Medicine Council and the National Academy of Sciences (NAS) Division of Earth and Life Studies Advisory Committee. As Chair of the NAS Board on Life Sciences, he created the study committee that produced "Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease," the report that enunciated the precision medicine concept, and he has helped to lead efforts in the White House, in Congress, in Sacramento and at UCSF to implement it. He has chaired or served on many committees that oversee training and the biomedical workforce, research funding, and the process of peer review and the policies that govern it at the National Institutes of Health. He is a member of the advisory board for Lawrence Berkeley National Laboratory and the board of directors of Research!America. He was elected to the National Academy of Sciences, the National Academy of Medicine, the American Academy of Arts and Sciences, and the American Academy of Microbiology, and is a fellow of the American Association for the Advancement of Science.

PHILIP E. BOURNE is the Founding Dean of the School of Data Science and Professor of Biomedical Engineering at University of Virginia (UVA). From 2014-2017, Dr. Bourne was the Associate Director for Data Science at the National Institutes of Health (NIH). In this role he led the Big Data to Knowledge Program, coordinating access to and analyzing biomedical research from across the globe and making it available to scientists and researchers. While there, he was also responsible for governance and strategic planning activities for data and knowledge management, and established multiple trainings in data science. He has done exceptional work to make biomedical research accessible, as well as to advance the field of data science. Prior to his time at the NIH, Dr. Bourne spent 20 years on the faculty at the University of California-San Diego, eventually becoming Associate Vice Chancellor of Innovation and Industrial Alliances. He is a highly respected and oft-cited scholar who brings a wealth of experience to UVA.

TATIANA BRYANT is the Research Librarian for digital humanities, History, and African American Studies at University of California Irvine Libraries. She holds a M.P.A. in International Public and Nonprofit Administration, Management, and Policy from New York University, a M.S. in Information and Library Science from Pratt Institute, and a B.A. in History from Hampton University. She has been a SPARC OpenCon Berlin fellow and a Digital Native American and Indigenous Studies Fellow through the National Endowment for the Humanities Office of Digital Humanities Institute. Her research includes studies on gender identity and performance in library work as well as perceptions of open access publishing among faculty who identify as Black, Indigenous, and/or people of color. She is a co-editor of the forthcoming volume, *Implementing Excellence in Diversity, Equity, and Inclusion: A Handbook for Academic Libraries* (ACRL Press).

MICHAEL CROW became the 16th president of Arizona State University (ASU) on July 1, 2002. He is guiding the transformation of ASU into one of the nation's leading public metropolitan research universities, an institution that combines the highest levels of academic excellence, inclusiveness to a broad demographic, and maximum societal impact—a model he terms the "New American University." During his tenure the university more than quadrupled research expenditures, completed an unprecedented infrastructure expansion, and was named the nation's most innovative school by *U.S. News and World Report* in 2016, 2017 and 2018. Dr. Crow was previously executive vice provost of Columbia University, where he also was professor of science and technology policy in the School of International and Public Affairs. He played the lead role in the creation of and served as the founding director of the Earth Institute at Columbia University, and in 1998 founded the Consortium for Science, Policy, and Outcomes, dedicated to linking science and technology to optimal social, economic, and environmental outcomes. An elected fellow of the American Association for the Advancement of Science (AAAS) and the National Academy of Public Administration, and member of the Council on Foreign Relations and U.S. Department of Commerce National Advisory Council on Innovation and Entrepreneurship, he is the author of books and articles analyzing science and technology policy and the design of knowledge enterprises and higher education institutions and systems. Dr. Crow received his Ph.D. in public administration (Science and Technology Policy) from the Maxwell School of Citizenship and Public Affairs, Syracuse University.

LAUREN COLLISTER, Chair of Committee on Scholarly Communication in Linguistics at the Linguistic Society of America. Dr. Collister is also the Director of the Office of Scholarly Communication and Publishing at the University Library System, University of Pittsburgh. She oversees all of the university's open access publishing, repository, copyright, altmetrics, and scholarly communication work. Broadly, she is an advocate for Open and her work is to advance the Open Scholarship movement. Her research interests in linguistics include discourse markers and deixis and how those intersect with linguistic innovation in online spaces. She is a member of the Linguistics Data Interest Group for the Research Data Alliance and also work as an advocate for good scholarly communication practice within the field of Linguistics. Dr. Collister has a Ph.D. in Sociolinguistics and has since worked to help linguists share their research openly, including advancing the data citation and sharing practices of our field.

JULIA STEWART LOWNDES is Senior Fellow at the National Center for Ecological Analysis and Synthesis (NCEAS) of University of California-Santa Barbara and founding director of Openscapes. Dr. Lowndes champions kinder, better science in less time through open data science and teamwork. As a marine data scientist, Mozilla Fellow '19 and Senior Fellow at NCEAS, she has over 7 years designing and leading programs to empower science teams with skillsets and mindsets for reproducible research, empowering researchers with existing open tools and communities. She has been building communities of practice in this space since 2013 with the Ocean Health Index after earning her Ph.D. at Stanford University studying drivers and impacts of Humboldt squid in a changing climate. Dr. Lowndes is a Carpentries instructor, lead creator of the Ocean Health Index's open data science training, and a co-founder of Eco-Data-Science and R-Ladies Santa Barbara.

EKEMINI RILEY is the managing director of Aligning Science Across Parkinson's (ASAP), a research funding initiative that coordinates targeted basic research and resources to uncover the roots of Parkinson's disease. Prior to ASAP, Dr. Riley was a director at the Milken Institute Center for Strategic Philanthropy where she helped shape and co-direct their medical research practice. She designed and facilitated several multi-sector think tank sessions to inform the strategic deployment of philanthropic capital, crafted research programs, and seeded multi-funder collaboration. Dr. Riley led the development and launch of ASAP, as well as the Gilbert Family Foundation's Gene Therapy and Vision Restoration Initiatives. Her work also laid the foundation for Play It Forward Pittsburgh, an organ donation awareness campaign. Dr. Riley completed her B.A. in natural sciences from the Johns Hopkins University and Ph.D. in molecular medicine from the University of Maryland School of Medicine. Her doctoral research focused on gene regulation of an endogenous protease inhibitor and its role in innate immunity and tumor suppression.

SANJAY SRIVASTAVA is a professor in the Department of Psychology at the University of Oregon and director of the Personality and Social Dynamics Lab. He teaches courses on a number of topics, including introductory psychology, motivation and emotion, social and personality psychology, and advanced statistics. Prior to coming to the University of Oregon, Dr. Srivastava was a postdoctoral research scientist at Stanford University. His research focuses on how personality affects and is affected by the social environment. This includes research on interpersonal perception, emotions, personality dynamics and development, and the psychology of online societies. He received his B.A. in psychology from Northwestern University and his Ph.D. in psychology from the University of California, Berkeley.

CAMILLE THOMAS is the Scholarly Communications Librarian at Florida State University. She currently leads initiatives to support with students, faculty, and staff to engage with new modes of research and teaching, including open access and open education. She has worked as a SPARC Fellow, on public interest partners and enhancing discovery for Open Educational Resources. Ms. Thomas received her Masters of Library and Information Studies from Florida State University in 2015 and a B.A. in Creative Writing and Journalism from the University of Central Florida in 2012. Her research includes data in libraries, early career leadership and management, user experience and open access, open education.

Developing a Toolkit for Fostering Open Science Practices: Overview[1]

Greg Tananbaum, Open Research Funders Group

Launched in 2019, the National Academies of Sciences, Engineering, and Medicine's Roundtable on Aligning Incentives for Open Science convenes critical stakeholders to fundamentally improve the correlation between open science practices, credit/reward systems, and research values. The Roundtable brings together senior leaders from universities, funding agencies, societies, philanthropies, and industry whose organizations have ambitious missions—tackling seemingly intractable societal challenges, pursuing the mysteries of science and nature, attempting to gain a better understanding of the human condition and our place in the universe—in an attempt to properly incentivize a more rigorous, transparent, and effective research culture.

An important focus of the Roundtable's work to date has been defining resources that can help key stakeholders discuss, develop, and deploy open science incentivization plans that are both consistent with common norms and appropriate for their specific communities. Given the points of leverage that these stakeholders manage (e.g., hiring, review, tenure and promotion, funding), how they can be activated to create better alignment across research values, practices, and incentives? Put succinctly, incentives are the tools we use to ensure that research practices are consistent with the organizational values we espouse (see Figure 1).

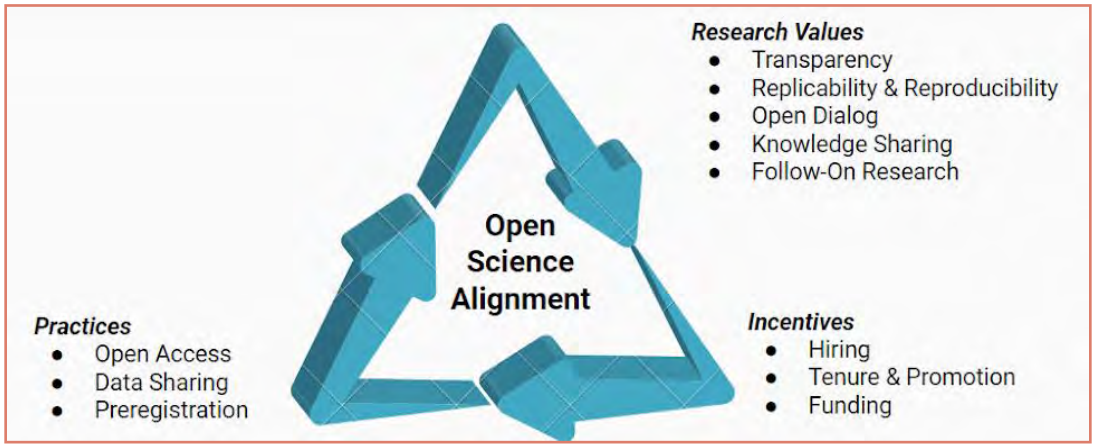

Figure 1. Open Science Alignment

[1] This document was prepared to stimulate discussion at the November 5, 2020 National Academies workshop on Developing a Toolkit for Fostering Open Science Practices. The views expressed are those of the author and do not necessarily reflect the official policies or positions of his employing organizations. This document is not a report of the National Academies of Sciences, Engineering, and Medicine and has not been subjected to its review procedures.

At the November 5, 2020 public workshop on Developing a Toolkit for Fostering Open Science Practices, a broad group of thought leaders and researchers will share perspectives on adopting and utilizing information resources compiled in the form of a toolkit. As background for the discussion, a number of possible toolkit elements that have been developed by expert authors in recent months are being circulated in advance. The draft toolkit elements are intended to stimulate discussion among the community about how such a toolkit might be used, what additional materials are needed, and how such a toolkit should be disseminated for broad adoption. The toolkit elements will be revised following the workshop and included in the published workshop proceedings.

The toolkit is primarily intended to assist university leadership, academic department chairs, research funders, learned societies, and government agencies. The draft toolkit elements to be discussed are:

The Open Science Imperative. This essay communicates the benefits of open science using approachable language.

Open Science by the Numbers Infographic. Intended to communicate the benefits of open science in a graphic form.

Signaling Language Templates 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).

Good Practices Primers. These are concise guides intended to offer policymakers a high level overview of open sharing.

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

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.

Suggested Open Science Signaling Language & Insertion Targets[1]

By
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
Lisa Nichols, OSTP
Greg Tananbaum, ORFG
Roger Wakimoto, UCLA

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 & 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), 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.

2. For each of the categories below, please provide *representative examples* demonstrating how you have made research outputs resulting from other projects openly accessible. If possible, please 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

[1] This document was prepared to stimulate discussion at the November 5, 2020 National Academies workshop on Developing a Toolkit for Fostering Open Science Practices. The views expressed are those of the authors and do not necessarily reflect the official policies or positions of their employing organizations. This document is not a report of the National Academies of Sciences, Engineering, and Medicine and has not been subjected to its review procedures.

- Other outputs (please describe)

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, please provide *representative examples* demonstrating how you have made research outputs resulting from this project openly accessible. If possible, please 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)

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, please provide *representative examples* demonstrating how you have made research outputs resulting from this project openly accessible. If possible, please 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)

Universities

Faculty Annual Report

1. For each of the categories below, please provide representative examples demonstrating how (where appropriate) you have made outputs resulting from your research openly accessible. If possible, please 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)

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. Please 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 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). Please describe the impact that your openly available research outputs have had from the research, public policy, pedagogic, and/or societal perspectives.

Sending Signals Rubric[1]

By
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
Lisa Nichols, OSTP
Greg Tananbaum, ORFG
Roger Wakimoto, UCLA

This rubric is intended to complement 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 & 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. The fourth sheet 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.

[1] This document was prepared to stimulate discussion at the November 5, 2020 National Academies workshop on Developing a Toolkit for Fostering Open Science Practices. The views expressed are those of the authors and do not necessarily reflect the official policies or positions of their employing organizations. This document is not a report of the National Academies of Sciences, Engineering, and Medicine and has not been subjected to its review procedures.

Amalgamated Version – Application Stage

Application Stage (e.g., jobs, grants)	**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/code according to FAIR 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, ORCIDs, 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 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, ORCIDs, 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, ORCIDs, Creative Commons licenses).
Provide examples of	The researcher cannot provide	The researcher can provide qualitative	The researcher can provide qualitative	The researcher can provide qualitative

Application Stage (e.g., jobs, grants)	**Beginning 1**	**Developing 2**	**Accomplished 3**	**Exemplary 4**
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)	qualitative and/or quantitative evidence that any of their recent (<5 years) open research outputs have been used by others.	and/or quantitative evidence that at least one of their recent (<5 years) open research outputs has been used by others.	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.	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 plans to engage in open activities in the future	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.	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	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	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.

Application Stage (e.g., jobs, grants)	**Beginning 1**	**Developing 2**	**Accomplished 3**	**Exemplary 4**
		researcher has articulated a plan that demonstrates an awareness of at least some aspects of good open science hygiene (e.g., use of DOIs, ORCIDs, Creative Commons licenses).	articulated a plan that demonstrates an intent to engage in good open science hygiene in most instances (e.g., use of DOIs, ORCIDs, Creative Commons licenses).	Additionally, the researcher has articulated a clear and consistent plan to engage in good open science hygiene (e.g., use of DOIs, ORCIDs, Creative Commons licenses).

Amalgamated Version - Reporting Stage

Reporting Stage (e.g., faculty tenure & promotion reviews, interim and final grant reports)	**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, ORCIDs, 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 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 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	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

Reporting Stage (e.g., faculty tenure & promotion reviews, interim and final grant reports)	**Beginning 1**	**Developing 2**	**Accomplished 3**	**Exemplary 4**
			(e.g., use of DOIs, ORCIDs, Creative Commons licenses).	DOIs, ORCIDs, Creative Commons licenses).
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)	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.	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.	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.	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.
For your work [related to this grant/during this time period], enumerate plans to engage in open activities in the future	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.	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	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	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

Reporting Stage (e.g., faculty tenure & promotion reviews, interim and final grant reports)	**Beginning 1**	**Developing 2**	**Accomplished 3**	**Exemplary 4**
		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, ORCIDs, Creative Commons licenses).	(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, ORCIDs, Creative Commons licenses).	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, ORCIDs, Creative Commons licenses).

Articles Version – Application Stage

Application Stage (e.g., jobs, grants)	**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, ORCIDs, 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, ORCIDs, 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 articles available in open access journals or repositories; and (b) consistently demonstrating good open science hygiene (e.g., use of DOIs, ORCIDs, Creative Commons licenses).
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 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)	The researcher has articulated a clear plan to make most research articles available openly for access and reuse. Specific activities include (a) making	The researcher has articulated a clear plan to make all appropriate research articles available openly for access and reuse. Specific activities

Application Stage (e.g., jobs, grants)	**Beginning 1**	**Developing 2**	**Accomplished 3**	**Exemplary 4**
		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, ORCIDs, Creative Commons licenses).	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, ORCIDs, Creative Commons licenses).	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, ORCIDs, Creative Commons licenses).

Articles Version – Reporting Stage

Reporting Stage (e.g., faculty tenure & promotion reviews, interim and final grant reports)	**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 access activities, including representative examples	The researcher has not, in their research [for this project/period], demonstrably engaged in making research articles openly available for access and reuse.	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, ORCIDs, Creative Commons licenses).	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, ORCIDs, Creative Commons licenses).	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, ORCIDs, Creative Commons licenses).
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)	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.	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.	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.	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.

Reporting Stage (e.g., faculty tenure & promotion reviews, interim and final grant reports)	**Beginning 1**	**Developing 2**	**Accomplished 3**	**Exemplary 4**
For your work [related to this grant/during this time period], enumerate 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 (including, but not limited to, articles and data) 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, ORCIDs, 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, ORCIDs, 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, ORCIDs, Creative Commons licenses).

Data Version – Application Stage

Application Stage (e.g., jobs, grants)	**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, ORCIDs, 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, ORCIDs, 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, ORCIDs, Creative Commons licenses).
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)	The researcher cannot provide qualitative and/or quantitative evidence that any of their recent (<5 years) open datasets have been used by others.	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.	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.	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.
Enumerate plans to engage in open data activities in the future	The researcher has not articulated a clear plan to make at least some research data available for	The researcher has articulated a clear plan to make at least some research data available for access	The researcher has articulated a clear plan to make most research data available for access	The researcher has articulated a clear plan to make all appropriate research data

Application Stage (e.g., jobs, grants)	**Beginning 1**	**Developing 2**	**Accomplished 3**	**Exemplary 4**
	access and reuse according to FAIR principles.	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, ORCIDs, Creative Commons licenses).	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, ORCIDs, Creative Commons licenses).	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, ORCIDs, Creative Commons licenses).

Data Version – Reporting Stage

Reporting Stage (e.g., faculty tenure & promotion reviews, interim and final grant reports)	**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, ORCIDs, 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, ORCIDs, 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, ORCIDs, Creative Commons licenses).
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	The researcher cannot provide qualitative and/or quantitative evidence that any of their open datasets [for this project/period] have been used by others.	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.	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.	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.

Reporting Stage (e.g., faculty tenure & promotion reviews, interim and final grant reports)	**Beginning 1**	**Developing 2**	**Accomplished 3**	**Exemplary 4**
(include DOIs if possible)				
For your work [related to this grant/during this time period], enumerate plans to engage in open activities in the future	The researcher has not articulated a clear plan to make at least some research data available for access and reuse according to FAIR principles.	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, ORCIDs, Creative Commons licenses).	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, ORCIDs, Creative Commons licenses).	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, ORCIDs, Creative Commons licenses).

Other Outputs Version – Application Stage

Application Stage (e.g., jobs, grants)	**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, ORCIDs, Creative Commons licenses).	The researcher has (a) 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, ORCIDs, 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, ORCIDs, Creative Commons licenses).
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)	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.	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.	The researcher can provide qualitative and/or quantitative evidence that (a) some 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.	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.
Enumerate plans to engage in open activities in the future, beyond the open sharing of	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.	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;	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	The researcher has articulated a clear plan to (a) make all appropriate research outputs (beyond articles and data) available openly for access

Application Stage (e.g., jobs, grants)	**Beginning 1**	**Developing 2**	**Accomplished 3**	**Exemplary 4**
articles and data.		and (b) engage in at least some aspects of good open science hygiene (e.g., use of DOIs, ORCIDs, Creative Commons licenses).	in good open science hygiene in most instances (e.g., use of DOIs, ORCIDs, Creative Commons licenses).	and reuse; and (b) engage in consistent good open science hygiene (e.g., use of DOIs, ORCIDs, Creative Commons licenses).

Other Outputs Version – Reporting Stage

Reporting Stage (e.g., faculty tenure & promotion reviews, interim and final grant reports)	**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 (beyond sharing articles and data), including representative examples	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.	The researcher has (a) occasionally made research outputs research [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, ORCIDs, Creative Commons licenses).	The researcher has (a) 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, ORCIDs, Creative Commons licenses).	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, ORCIDs, Creative Commons licenses).
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)	The researcher cannot provide qualitative and/or quantitative evidence that any of their open research outputs (beyond articles and data) [for this project/period] have been used by others.	The researcher can provide qualitative and/or quantitative evidence that at least one of their open research outputs (beyond articles and data) [for this project/period] has been used by others.	The researcher can provide qualitative and/or quantitative evidence that (a) some of their open research outputs (beyond articles and data) [for this project/period] have been used by others; and/or (b) a narrower range of their open research outputs (beyond articles and data) [for this project/period] have been used deeply within a specific community.	The researcher can provide qualitative and/or quantitative evidence that (a) a wide range of their open research outputs (beyond articles and data) [for this project/period] have been used by others; and/or (b) a narrower range of their open research outputs (beyond articles and data) [for this project/period] have been used deeply within a specific community.

Reporting Stage (e.g., faculty tenure & promotion reviews, interim and final grant reports)	**Beginning 1**	**Developing 2**	**Accomplished 3**	**Exemplary 4**
For your work [related to this grant/during this time period], enumerate plans to engage in open activities (beyond sharing articles and data) in the future	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.	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, ORCIDs, Creative Commons licenses).	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, ORCIDs, Creative Commons licenses).	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, ORCIDs, Creative Commons licenses).

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.
- "Other Outputs" include a range of research products such as the NASEM Roundtable list enumerated here.

Good Practices Primers[1]

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

Many organizations are in the information-gathering stage with respect to their open science investigation. These concise primers are intended to provide decision makers with a high-level overview of the *what's* and *how's* of open sharing. Each primer (1-2 pages) addresses a different output type, delving into exemplars, dependencies, resourcing, and a range of other considerations. The following ***drafts*** are intended to provide a sense of what the primers will encompass. Feedback from workshop participants is actively encouraged.

[1] This document was prepared to stimulate discussion at the November 5, 2020 National Academies workshop on Developing a Toolkit for Fostering Open Science Practices. The views expressed are those of the authors and do not necessarily reflect the official policies or positions of their employing organizations. This document is not a report of the National Academies of Sciences, Engineering, and Medicine and has not been subjected to its review procedures.

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 will inform an organization's approach to making articles open are the following:

- ***Fulfillment***. Can researchers adhere to the policy by publishing in a fully open access journal, a "hybrid" journal, 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., six months) permissible?
- ***Financial Support.*** Will the policy maker provide funding to defray costs of open access (e.g., article processing charges [APCs])? 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 & Reuse.*** What type of licensing requirements will the policy include to facilitate reuse? 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.

Approaches

The practical implementation of a policy requiring access to published articles can take a number of different forms. In general, policies that permit delays, exceptions, and restrictions are considered less open that policies that encourage immediate, permissive sharing. However, certain organizations (e.g., signatories to the Open Research Funders Group Incentivization Blueprint) have taken the approach that an initial policy can be built upon and rendered more progressive over time.

Some policies require publication in an open access journal or a hybrid journal (a subscription-based journal that allows authors to make individual articles gold open access immediately on payment of an article publication charge). This can introduce a modest restriction on researchers' choice of publication venue, although thousands of journals are open access or offer a hybrid option.

Examples of Open Access Policies Requiring Publication in Open Access Journals

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

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. It does require the author 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 the author.

Examples of Self-Archiving Open Access Policies

- All US 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.
- 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.
- 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.

SPARC (the Scholarly Publishing and Academic Resources Coalition) maintains a succinct resource for tracking, comparing, and understanding U.S. federal funder article sharing policies. ROARMAP (the Registry of Open Access Repository Mandates and Policies) provides similar information about funders and universities. The Federal interagency group, CENDI, posts information about Federal agency public access policies. These sites can be used to compare and contrast different approaches that stakeholders are taking to open access policies.

Resourcing

Administrators may be concerned about how policy changes can create additional operational work to already busy staff. Policies can require discussion within the organization, integration into current workflows, and the engagement of operational staff (e.g., for the disbursement of article processing charges, compliance checking, etc.).

Once implemented, there is a range of activities that organizations can take to manage open policies. 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 they are able to devote to these activities.

Next Steps

The Open Research Funders Group (ORFG) can provide support and insight into best practices and available resources. The ORFG Incentivization Blueprint provides model language that can be adopted and adopted by

funders and other organizations. It offers a stepwise approach to deploying a policy that can grow to encompass not only open access articles, but data, code, and other research outputs.

Code and Software

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:

- ***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., six months) permissible? In the event that research findings are not published or posted, should code and software be made publically 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 publically available after the conclusion of the grant, does the grantee have a mechanism to request additional financial support?
- ***Future Proofing.*** What are the expectations for the duration and extent to which code should be kept up to date?
- ***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?
- ***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

In general, policies that permit delays, exceptions, and restrictions are considered less open than policies that encourage immediate, permissive sharing. 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". More funder-specific TOP guidance may be found here.

Some agencies within the U.S federal government, including the Consumer Financial Protection Bureau, unequivocally state, "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." Other agencies, such as the Department of Education, make the source code for their prominent public-facing initiatives (in DOE's case, the College Scorecard) openly available. Both of these organizations deposit these research outputs (software as a product, not a byproduct, of the grant) on GitHub. In cases where code is developed to interpret or analyze research findings (code as a secondary output of the grant), organizations such as the Wellcome Trust typically require the code to be shared at the time the primary research is published.

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. All funding proposals must include software development plans that are vetted as part of the application process.
- The United States 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 (OSS) for three years".
- A number of 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.

Resourcing

Administrators may be concerned about how policy changes can create additional operational work to already busy staff. Policies can require discussion within the organization, integration into current workflows, and some form of compliance checking. 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 implemented, there is a range of activities that organizations can take to manage open policies. 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. Organizations without open policies may view administration and compliance as daunting tasks. However, each organization can make its own appropriate determination about the resources they are able to devote to these activities.

Next Steps

The TOP Guidelines provide sample language for three levels of open code policies. 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

Best Practices for a Future Open Code Policy for NASA Space Science Report. 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.

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 shed light on negative results and unproductive research directions, has the potential to improve the efficiency of the research process, and can lead to novel lines of inquiry. In particular, shared data can be re-used 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 is 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***. Should data related to reported results be made available concurrent with the posting of research findings? Should researchers be given a window of exclusivity (e.g., 6-12 months) to analyze their research data before sharing it with the community? If data are not reported in a publication, what is an appropriate timeline 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), Family Educational Rights and Privacy Act (FERPA), General Data Protection Regulation (EU GDPR), or other privacy restrictions. Such datasets, however, can often be shared after anonymization or deidentification 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, e.g., by funders, other institutions, or individuals?

Approaches

The generally accepted approach to open data is to ensure that policies should require data to be findable, accessible, interoperable, and reusable, i.e., to meet the FAIR data principles. 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. 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 "as open as possible and as closed as necessary". This approach recognizes that data should be shared in a manner that promotes reuse and transparency, while at the same time recognizing that certain safeguards may be required to protect sensitive information that could compromise subject privacy or other norms and regulations.

Many US 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 timelines and other conditions of access. For certain types of research studies, Federal science agencies have developed more specific guidance or requirements (see NIH example below).

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 circumstances. Others, such as the National Cancer Institute's Moonshot initiative, provide much more explicit guidance as to the timing, licensing, and dissemination of data sharing activities.

Examples of Open Data Policies

- The National Institutes of Health has policies that apply to genomic data and clinical trial data, as well as other specific research programs and data types. It recently issued for public comment a draft data management and sharing policy that would cover all awards that generate data.

- The American Heart Association requires grant applicants to include a data sharing plan as part of the application process. Any research data that is 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 principles.
- 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.

Resourcing

Administrators may be concerned about how policy changes can create additional operational work to already busy staff. Policies can require discussion within the organization, integration into current workflows, and some form of compliance checking. For data specifically, it will be 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 implemented, there is a range of activities that organizations can take to manage open policies. 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 they are 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.

- The TOP Guidelines provide sample language for three levels of open data policies. This wording can be adapted and adopted to suit the specific circumstances of a range of organizations.

- The Open Research Funders Group Incentivization Blueprint offers sample open data policy language that can be adapted for a range of use cases.

- The American Heart Association's website contains a detailed FAQ page that articulates questions commonly asked by researcher subject to an open data policy.

- 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.

Protocols & 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 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 analysis plan (how the collected data will be organized and evaluated), or both? Of primary interest in ensuring the integrity of the research outcome is information about the prespecified outcome measures/endpoints. However, decisions made during analysis can also impact 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 also allows for structured data entry to facilitate searching and cross-study comparison. If a summary, 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 personnel). 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. Within the social, behavioral, and pre-clinical sciences, the Open Science Framework is becoming a default registry. 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 IRB approval been received). The timing of an update is also linked to the degree to which a change has implications on the full preregistration (for example, challenges in recruiting a full sample may necessitate moving from a single cohort to a multi-cohort design). Protocols shared at study initiation can more clearly establish a project's aims and plan. Does the registry support timestamped 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?

Approaches

There are a range of different preregistration locations available, primarily driven by discipline. All NIH-funded clinical trials and most clinical trials of 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 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.

Examples of Preregistration and Protocols Policies

- The Chan Zuckerberg Initiative requires grantees to make experimental protocols publicly available and has nurtured dedicated protocol communities of CZI-funded investigators.
- The American Economic Association encourages researchers to register their randomized controlled trials (including research designs and analysis plans) in the AEA RCT Registry.
- 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.
- Arnold Ventures requires all funded empirical studies that involve statistical inference to be preregistered before the start of intervention or data collection on OSF.

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. Researchers in the social, behavioral, and cognitive sciences often use the Open Science Framework platform. The Registry for International Development Impact Evaluations (RIDIE) hosts impact evaluations related to development in low and middle income countries.

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. To be most effective, preregistrations and protocols should be closely linked to associated publications and other study information so 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 pregistration. This wording can be adapted and adopted to suit the specific circumstances of a range of organizations. The TOP recommendations include (1) disclosing whether or not work was preregistered, (2) verifying that any preregistered work adheres to the pre-specified plans, and (3) requiring preregistration for relevant research studies (typically inferential and hypothesis-testing work).

The Center for Open Science provides multiple resources on how to preregister studies and analytic plans, including templates.

The 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.

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 or not all or part of the work would be 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 pre-collection hypotheses and data analysis plans have been reviewed and registered.

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. In the event that the study does not receive an “in principle acceptance” (IPA) offer from a journal, the plan can still be preregistered by the authors 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, 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*.

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.

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.

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, 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.

The Open Science Framework provides a searchable database of registered reports across a range of disciplines. These may offer useful guidance to better understand the core elements of a well-constructed registered report.

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.

$3.2 TRILLION

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]

37 MILLION
BIOLOGICAL RECORDS

766
UNIVERSITIES, MUSEUMS, RESEARCH ORGANIZATIONS

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

700

Global Open Data for Agriculture and Nutrition (GODAN) is an open data sharing initiative drawing on the participation of **over 700 private and public sector, nonprofit, and academic organizations** with the goal of developing solutions to global hunger.[3]

1 month from first reported COVID-19 case to genetic sequencing, rapidly expedited by open science and data sharing

NIH The Human Genome Project timeline

1988

$965 BILLION
ECONOMIC OUTPUT

$293 BILLION
IN PERSONAL INCOME

4 MILLION JOBS

2000

The Human Genome Project, an international research collaboration which mapped the human genome and made the findings publicly available, generated **$965 billion** in economic output between **1988 and 2012**, creating more than **$293 billion** in personal income through wages and benefits, and nearly **4 million jobs**.[4]

2012

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, NIST; Greg Tananbaum, ORFG; Alan Tomkins, NSF; Thomas Wang, American Heart Association/UT Southwestern Medical Center

Sources:
1. https://www.omidyar.com/sites/default/files/file_archive/insights/ON%20Report_061114_FNL.pdf
2. http://symbiota.org/docs/
3. https://www.usda.gov/media/blog/2018/07/13/open-data-enabling-fact-based-data-driven-decisions
4. https://sparcopen.org/impact-story/human-genome-project/
5. https://www.biorxiv.org/content/10.1101/795310v1

The Open Science Imperative[1]

By
Derrick Anderson, Arizona State University
Rachel Bruce, UK Research and Innovation
Ashley Farley, Bill & Melinda Gates Foundation
Robert Hanisch, NIST
Greg Tananbaum, ORFG
Alan Tomkins, NSF
Thomas Wang, American Heart Association/UT Southwestern Medical Center

Draft version 3.0

October 31, 2019

This narrative is intended to communicate 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 around scientific literacy and education, and improving public policy and democracy.[2] 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, non-inclusive 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 Aligning Incentives for Open Science, briefly sketches the current state of Open Science, contrasts the diminishing returns of the traditional scientific model with

[1] This document was prepared to stimulate discussion at the November 5, 2020 National Academies workshop on Developing a Toolkit for Fostering Open Science Practices. The views expressed are those of the authors and do not necessarily reflect the official policies or positions of their employing organizations. This document is not a report of the National Academies of Sciences, Engineering, and Medicine and has not been subjected to its review procedures.

[2] Tennant et al., 2016; Zuccala, 2010

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[3] but 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."[4] 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 pre-registration, open access to data and materials, and development of open source software.[5]

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 is shared. In one survey, the number of researchers who reported making their data openly available increased from just over 55% to 64% between 2016 and 2018. From before 1990 through the 2010s, the percentage of researchers who were unaware of what license under which they made their data openly available decreased from 71% to 54%.[6] 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% to 27%.[7]

The rise of open access (OA) 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% of all scholarly publications are available as open access and that the proportion is growing. The same analysis indicated that, given existing trends, 70% of all article views will be to open access papers by 2025.[8] This trend appears to be driven by the values held by researchers: "over 90% of OA authors published this way because of the principle of free access,"[9] 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."[10] 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% year-on-year.[11]

[3] Gold, 2016
[4] Nielsen, 2011
[5] Berg and Niemeyer, 2018; Gold, 2016; Gold et al., 2019
[6] Science, Digital, et al. 2018, p. 8
[7] Ibid., p. 13.
[8] Piwowar, et al. 2019
[9] Swan and Brown, 2004, p. 5
[10] Swan and Brown, 2005, Executive Summary p. 1
[11] Abdill and Blekhman, 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.[12]

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 non-scientists 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 policymakers 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 four of these dimensions.

[12] Spellman, Gilbert and Corker, 2018; Tennant et al., 2016

1. **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[13]) 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.[14]
2. **The Group on Earth Observations (GEO)** is a global network of 100+ national governments and 100+ participating 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 Observing 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, "[d]ata 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."[15]
3. **The Lab @ DC** is a unit within the Washington DC Mayor's administration that works to design public policy and program interventions for the District. 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. [16]
4. **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% of projects funded by the National Science Foundation Advancing Digitization of Biodiversity Collections (NSF-ADBC) 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.[17] Importantly, Symbiota's software design philosophy and implementation

[13] Tripp and Grueber, 2011.
[14] SPARC, n.d.
[15] Zittrain, 2006; Benkler, 2006, NRC, 2009, and Mayo and Steinberg, 2007, cited in Uhlir, 2015, p. 13.
[16] The Lab @ DC, n.d.
[17] Symbiota, n.d.

was driven by its "*user community* – e.g., collections managers, taxonomists, ecologists, data entry personnel, programmers, informaticians, and students."[18] Symbiota is freely available to researchers and the public.

5. **Global Open Data for Agriculture and Nutrition (GODAN)** is an initiative of the US Department of Agriculture (USDA) and US Agency for International Development (USAID) that promotes open data sharing to increase global access to information about agriculture and nutrition. Leveraging data input from a partner network of over 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.[19]
6. **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 their 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.[20]
7. **California Policy Lab** is a nonprofit based at UCLA and the University of California, Berkley, that partners with state and local governments to solve social issues, including homelessness, poverty, crime, and educational inequality.[21] The California Policy Lab utilizes the Open Science Framework and has established data-sharing agreements with over a dozen county agencies in Los Angeles, Sonoma, and San Francisco covering "medical, mental health, criminal justice, social service, and homeless management information systems."[22] 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.
8. **The International Virtual Observatory** is an open platform enabling astronomers, educators, and the general public to discover, access, and integrate open data from world-wide (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 has powered more than 2,300 scholarly papers[23], covering the entire electromagnetic spectrum, from gamma-rays to radio waves.

Open Science and the Status Quo

[18] Gries, Gilbert and Franz, 2014.
[19] Adams, 2018.
[20] Wellcome Trust Sanger Institute, 2016
[21] California Policy Lab, n.d.
[22] California Policy Lab, 2018
[23] Data accessed from SAO/NASA Astrophysics Data System, October 16, 2020.

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 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.[24] Researchers become understandably hesitant to make their data and findings openly available out of fear of being "scooped" by other researchers.[25] 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 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.[26] 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 (for example, by better informing and directly involving patients in biosciences).[27] By making data openly accessible between researchers and the public, Open Science can provide greater opportunities for interdisciplinary, collaborative research across institutions worldwide.[28] Heightened collaboration can also lead to dynamic new knowledge hubs and remove barriers to upstream research and tech transfer.[29]

Greater efficiency and speed. Open data practices also drive efficiency by enabling real-time, data-driven decision-making.[30] The sharing of data reduces transaction costs, increases

[24] Heenan and Williams, 2018
[25] Berg and Niemeyer, 2018
[26] Brooks, 2010
[27] Gold, 2016
[28] Uhlir, 2015
[29] Gold, 2016
[30] Adams, 2018; SPARC, n.d.

reproducibility and re-use of data, decreases redundancy, and drives greater transparency, heightened efficiency, and accelerated sustainable innovation.[31]

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.[32]

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 towards facilitating inter-institutional 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:

1. *Creating supportive environments:* the Tannenbaum 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 neuro-informatics platforms, and an early-stage drug discovery unit that collaborates with academia and industry partners.[33]
2. *Incentivizing open access publishing.* The Bill and 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

[31] Gold 2016; Gold et al., 2019; Tennant et al., 2016
[32] Popkin, 2019; Uhlir, 2015
[33] Gold, 2016; Neuro, The, n.d.

requires scientists to publish papers in open access journals rather than those that charge subscriptions.[34]

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.[35]

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, pre-registration, and collaboration. The National Academies of Sciences, Engineering, and Medicine's 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., & Blekhman, R. (201, April 24). Meta-Research: Tracking the popularity and outcomes of all bioRxiv preprints. *eLife,* 2019;8:e45133 DOI: 10.7554/eLife.45133.

Adams, J. (2018, July 13). Open Data: Enabling Fact-Based, Data-Driven Decisions. Retrieved July 17, 2019, from https://www.usda.gov/media/blog/2018/07/13/open-data-enabling-fact-based-data-driven-decisions

Antelman, K. (2004). Do open-access articles have a greater research impact?. *College & research libraries*, *65*(5), 372-382.

Berg, D. R., & Niemeyer, K. E. (2018). The case for openness in engineering research. F1000Research, 7.

Brooks Jr, F. P. (2010). *The design of design: Essays from a computer scientist*. Pearson Education.

California Policy Lab. (2018, December 13). California policy lab awarded $1.2M UC Multicampus research grant. Retrieved from https://www.capolicylab.org/wp-content/uploads/2018/12/CPL-Press-Release-re-MRPI-12-12-18-final.pdf

California Policy Lab. (n.d.). What we do. Retrieved October 16, 2019 from https://www.capolicylab.org/what-we-do/

[34] Stokstad, 2018.
[35] National Institutes of Health, 2017.

Fecher, B., & Friesike, S. (2014). Open science: one term, five schools of thought. In *Opening science* (pp. 17-47). Springer, Cham.

Friesike, S., Widenmayer, B., Gassmann, O., & Schildhauer, T. (2015). Opening science: towards an agenda of open science in academia and industry. *The Journal of Technology Transfer*, *40*(4), 581-601.

Gold, E. R. (2016). Accelerating translational research through open science: The Neuro Experiment. *PLOS biology*, *14*(12), e2001259.

Gold ER, Ali-Khan SE, Allen L et al. (2019). An open toolkit for tracking open science partnership implementation and impact [version 1; peer review: awaiting peer review]. Gates Open Res 2019, 3:1442 (https://doi.org/10.12688/gatesopenres.12958.1)

Gries C, Gilbert E, Franz N (2014). Symbiota – A virtual platform for creating voucher-based biodiversity information communities. Biodiversity Data Journal 2: e1114. doi: 10.3897/BDJ.2.e1114

Hajjem, C., Harnad, S., & Gingras, Y. (2006). Ten-year cross-disciplinary comparison of the growth of open access and how it increases research citation impact. *arXiv preprint cs/0606079*.

Heenan, A. & Williams, I.D. (2018). How open data can help the world better manage coral reefs [HTML]. Retrieved from https://theconversation.com/how-open-data-can-help-the-world-better-manage-coral-reefs-88805

Neuro, The. (n.d.). Open science, to accelerate discovery and deliver cures. Retrieved September 3, 2019 from https://www.mcgill.ca/neuro/open-science-0

National Institutes of Health. (2018). Open science prize announces epidemic tracking tool as grand prize winner. Retrieved September 3, 2019 from https://www.nih.gov/news-events/news-releases/open-science-prize-announces-epidemic-tracking-tool-grand-prize-winner

Nielsen, M. (2011). An informal definition of open science. The OpenScience Project. http://openscience.org/an-informal-definition-of-openscience/

Piwowar, H., Priem, J., & Orr, R. (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, May 13). Data sharing and how it can benefit your scientific career. Retrieved June 13, 2019, from https://www.nature.com/articles/d41586-019-01506-x

Scheliga, K., & Friesike, S. (2014). Putting open science into practice: A social dilemma?. *First Monday*, *19*(9).

Science, Digital; Hahnel, Mark; Fane, Briony; Treadway, Jon; Baynes, Grace; Wilkinson, Ross; et al. (2018). The State of Open Data Report 2018. figshare. Journal contribution.

Seyednasrollah, F., Koestler, D. C., Wang, T., Piccolo, S. R., Vega, R., Greiner, R., ... & Kanigel Winner, K. (2017). A DREAM Challenge to build prediction models for short-term discontinuation of docetaxel in metastatic castration-resistant prostate cancer. *JCO clinical cancer informatics*, *1*, 1-15.

Smith, E., Haustein, S., Mongeon, P., Shu, F., Ridde, V., & Larivière, V. (2017). Knowledge sharing in global health research–the impact, uptake and cost of open access to scholarly literature. *Health research policy and systems, 15*(1), 73.

SPARC. (n.d.). From Ideas to Industries: Human Genome Project. Retrieved December 8, 2019, from https://sparcopen.org/impact-story/human-genome-project/

Spellman, B. A., Gilbert, E. A., & Corker, K. S. (2018). Open science. *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, *5*, 1-47.

Stokstad, Erik. (2018, Nov. 5). 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 Brown, S. (2004), Authors and open access publishing. Learned Publishing, 17: 219-224. doi:10.1087/095315104323159649

Swan, A. and Brown, S. (2005). Open access self-archiving: an author study. Key Perspectives Limited. Retrieved September 3, 2019 from http://cogprints.org/4385/1/jisc2.pdf

Symbiota. (n.d.). Symbiota Introduction. Retrieved September 3, 2019 from http://symbiota.org/docs/

Tripp, S and Grueber, M. (2001, May). Economic Impact of the Human Genome Project. Technology Partnership Practice. Retrieved September 5, 2019 from https://www.battelle.org/docs/default-source/misc/battelle-2011-misc-economic-impact-human-genome-project.pdf?sfvrsn=6

Tennant, J. P., Waldner, F., Jacques, D. C., Masuzzo, P., Collister, L. B., & Hartgerink, C. H. (2016). The academic, economic and societal impacts of Open Access: an evidence-based review. F1000Research, 5.

The Lab @ DC. (n.d.). Home. Retrieved October 10, 2019 from https://thelab.dc.gov/

Uhlir, P. (2015). The Value of Open Data Sharing. *Group on Earth Observations, GEO-XII–11-12 November*.

Wellcome Trust Sanger Institute. (2016, November). Online epidemic tracking tool embraces open data and collective intelligence to understand outbreaks. Retrieved September 4, 2019 from https://www.sanger.ac.uk/news/view/online-epidemic-tracking-tool-embraces-open-data-and-collective-intelligence-understand

Zuccala, A. (2010). Open access and civic scientific information literacy. *Information Research: An International Electronic Journal*, *15*(1).

Open Science Success Stories Database[1]

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

The database is available at: https://projectopen.io/

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.

Scientists, scholars, librarians, department chairs, university administrators, philanthropic program officers, government agency representatives, policymakers, publishers, journalists and other stakeholders can use the curated resources to understand how open science is positively impacting 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.

[1] This document was prepared to stimulate discussion at the November 5, 2020 National Academies workshop on Developing a Toolkit for Fostering Open Science Practices. The views expressed are those of the authors and do not necessarily reflect the official policies or positions of their employing organizations. This document is not a report of the National Academies of Sciences, Engineering, and Medicine and has not been subjected to its review procedures.

Reimagining Outputs Table[1]

By
Boyana Konforti, Howard Hughes Medical Institute
Elizabeth Albro, U.S. Dept. 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 which should be accounted for as we think about the behaviors and activities we should be rewarding. What are the outputs that are consistent with the values we collectively espouse? 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 we enumerate these research products, what considerations must we contemplate and address to create appropriate alignment between values and activities? The authors believe it will be crucial to ensure that we take an expansive view of the types of research products that should be "open" - available for access and reuse without gatekeeping or payment.

[1] This document was prepared to stimulate discussion at the November 5, 2020 National Academies workshop on Developing a Toolkit for Fostering Open Science Practices. The views expressed are those of the authors and do not necessarily reflect the official policies or positions of their employing organizations. This document is not a report of the National Academies of Sciences, Engineering, and Medicine and has not been subjected to its review procedures."

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.	• 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
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. In order 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.	• Stewardship/ownership of repositories -- ensuring these are open and sustainable

Research Output Type	Exemplar Open Practices	Importance to Open Ecosystem	Concerns/Considerations
Commentaries & Analyses	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.	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.	• Commentaries and summaries are an important way for learned societies to add value and continue to earn some subscription income
Data	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.	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.	• Stewardship/ownership of repositories -- ensuring these are open and sustainable • Timing of data release • Restrictions on data reuse (e.g., text and data mining)

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 a number of 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.	• 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.	• 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.	• Grantees/faculty members may require additional guidance as to what constitutes an appropriate research output

Research Output Type	Exemplar Open Practices	Importance to Open Ecosystem	Concerns/Considerations
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 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% 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 collaboration) • A number of 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.

Research Output Type	Exemplar Open Practices	Importance to Open Ecosystem	Concerns/Considerations
Preprints	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. Preprints should be posted in a timely manner, ideally at the time of first submission to a journal.	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.	• 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 • A number of initiatives are emerging to support preprints. For example, ASAPbio.org is a comprehensive resource for information on preprints, peer reviews, transparency, etc. Transpose is a directory of journal policies, co-reviewing, and preprints.

Research Output Type	Exemplar Open Practices	Importance to Open Ecosystem	Concerns/Considerations
Preregistration Analysis Plans	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 pre-registration 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.	Unreported flexibility in data analysis decrease scientific credibility and invalidate 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. pre-registering 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.	• 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
Protocols	Guidelines detailing 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.	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.	• Protocols can be shared prior to conducting work, which provides insights into research that does not ultimately get published; this is uncommon at present

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 timeline documentation, add appropriate time (e.g. 2 to 4 months) for peer review process at the beginning of the relevant project phases. Communicate with 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.	• 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

Research Output Type	Exemplar Open Practices	Importance to Open Ecosystem	Concerns/Considerations
Research Materials	Biological and other physical samples (in particular starting materials), research tools (including reagents, animal models, and the like) and other materials (including metadata) 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.	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.	• Cost of maintaining/sharing certain types of samples • Stewardship/ownership of repositories -- ensuring these are open and sustainable

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 offer a particularly unique insight into the research perspective of the emerging generation of scholars.	• Because students often try to publish portions of their theses and dissertations as articles, and because some journals still consider posted ETDs 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

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Roundtable on Aligning Incentives for Open Science

Overview

In order to increase the contribution of open science to producing better science, the National Academies of Sciences, Engineering, and Medicine's Roundtable on Aligning Incentives for Open Science will convene critical stakeholders to discuss the effectiveness of current incentives for adopting open science practices, current barriers of all types, and ways to move forward to optimally align reward structures and institutional values.

The Roundtable will convene two times per year and create a venue for exchange of ideas and a mechanism for joint strategic planning among key stakeholders. Each Roundtable meeting will have a theme. The diverse themes will target slightly different audiences but the core audience will be university, government, foundations, and society groups doing work related to open science. The Roundtable aims to improve coordination among stakeholders and increase awareness of current and future efforts in the broader open science community. The Roundtable will also convene one symposium per year, which may produce proceedings in brief.

The first meeting of the Roundtable took place on Monday, February 25, 2019 in Washington, DC. This initial meeting identified key challenges and opportunities relating to aligning incentives for open science, and develop the initial work plan or set of priorities for the Roundtable. On September 20, 2019, the Roundtable organized a public symposium, *Advancing Open Science Practices: Stakeholder Perspectives on Incentives and Disincentives*, in conjunction with its second meeting on September 19-20, 2019 in Washington, DC. A Proceedings of a Workshop—in Brief, issued in February 2020, synthesizes the discussions held during the event and is available at https://www.nap.edu/catalog/25725. The third meeting of the Roundtable took place on February 27, 2020. The Roundtable will convene a virtual public workshop, *Developing a Toolkit for Fostering Open Science Practices* on Thursday, November 5, 2020, as part of its fourth meeting on Friday, November 6 2020.

The Roundtable will be overseen by the Board on Research Data and Information (BRDI). Additional information about board activities can be found at www.nas.edu/brdi. The Roundtable discussions will be informed by the 2018 BRDI consensus study *Open Science by Design: Realizing a Vision of 21st Century Research*, which assesses the status of Open Science overall and across several disciplines, identifies motivations and barriers, and develops recommendations and implementation actions for research enterprise stakeholders.

Leadership, Participation, and Support

Dr. Keith Yamamoto, Vice Chancellor for Science Policy and Strategy at University of California, San Francisco and a member of both the National Academy of Sciences and National Academy of Medicine, and Mr. Thomas Kalil, Chief Innovation Officer of Schmidt Futures, are the co-chairs of the Roundtable. In addition to individual members, mainly from academia, the Roundtable includes representatives from government, foundations, stakeholder associations, and international organizations as ex-officio members. The roster of current Roundtable members is available at https://www.nationalacademies.org/our-work/roundtable-on-aligning-incentives-for-open-science.

The Roundtable on Aligning Incentives for Open Science is being supported by Arcadia, Arnold Ventures, Eric & Wendy Schmidt Fund for Strategic Innovation, Leona M. and Harry B. Helmsley Charitable Trust, National Library of Medicine, Open Research Funders Group, Open Society Foundations, Robert Wood Johnson Foundation, and the Wellcome Trust.

Roundtables at the National Academies of Sciences, Engineering, and Medicine

Roundtables are a type of continuing activity at the National Academies that provide a means for representatives of government, industry, private businesses, academia and other stakeholder groups to gather periodically in a neutral setting to identify and discuss issues of mutual concern on a continuing basis.

In contrast to the institutional requirements for members appointed to ad hoc committees or to other types of standing bodies of the National Academies, roundtable/forum members are not subject to any institutional restrictions with respect to conflicts of interest.

Because roundtables are not subject to institutional conflict of interest requirements, a roundtable may not be used to provide advice or recommendations nor may it be used in the development of a report.

Subject to the availability of funds, roundtables may commission the preparation of individually authored documents to explore and stimulate discussion of a topic or issue for use by the roundtable and dissemination to interested members of the public. Subject to the availability of funds, roundtables may use workshops to inform their meetings and discussions. These activities are subject to separate National Academies approval, and must be workshops that result in either a rapporteur-authored proceedings or a collection of individually-authored papers. Roundtable discussions may lead to proposals for specific studies to be undertaken by units of the National Academies resulting in institutional reports. Such studies, if undertaken, will be conducted independently from the roundtable and in accordance with all institutional policies and procedures governing such study activities.