National Science Foundation

NSF Public Access Plan 2.0

Ensuring Open, Immediate and Equitable Access to National Science Foundation Funded Research

NSF 23-104

February 21, 2023

Table of Contents

Foreword by NSF Director Sethuraman Panchanathan 3
Executive Summary 4
Section 1: Background and Policy Principles 5
General Introduction 5
Public Access at the National Science Foundation 5
NSF Strategic Plan and Public Access Objectives 7
Section 2: Open Science Responses – Opportunities and Equity 8
2A. COVID-19 and Other Threats to National Health, Prosperity and Security 8
2B. Equity Opportunities and Concerns 8
Section 3: Increasing Equitable Access to Federally Funded Research Results 10
3.A. Peer-Reviewed Scholarly Publications 10
3.A.i. How Peer-Reviewed Scholarly Publications Will be Made Publicly Accessible 10
3.A.ii. Maximizing Public Access to Scholarly Publications With Machine-Readability 11
3.A.iii. Re-use Rights for Publications Made Freely and Publicly Available 11
3.B. Scientific Data 12
3.B.i. Scientific Data Underlying Peer-Reviewed Scholarly Publications 12
3.B.ii. Scientific Data Not Associated With Peer-Reviewed Scholarly Publications 13
3.B.iii. Digital Repositories for Scientific Data 13
3.B.iv. Federal Agency Research 14
3.C. Researcher Responsibilities for Managing and Sharing Data Arising From NSF Awards 14
3.C.i. Legal, Privacy, Ethical, Technical, Intellectual Property, and Security Limitations on Access to Scientific Data Arising From NSF Awards 14
3.C.ii. Plans for Maximizing Appropriate Data and Providing Risk-Mitigated Opportunities for Limited Data Access 15
3.C.iii. Specific Online Digital Repositories 15
3.D. Allowable Publication and Data Costs 15
3.E. Ongoing Reporting to OSTP on Request Regarding Plan Status 16
Section 4: Ensuring Scientific and Research Integrity 16
4.A. Collect and Make Publicly Available Appropriate Publication Metadata 16
4.A.i. Names, Organizational Affiliations and Funding 16
4.A.ii. Publication Date(s) 17
4.A.iii. Identifiers for Research Outputs 17
4.B. Obtaining an NSPM-33 Digital Persistent Identifier for Award Senior Personnel 17
4.C. Assigning PIDs to Research and Development Awards 17
Section 5: Public Access Plan Coordination Among Federal Agencies 18
Section 6: General Provisions 18
Information 18
Authority 19
Roles and Responsibilities 19
Section 7: Taking Next Steps Together 19
Glossary 21
Endnotes 23

FOREWORD

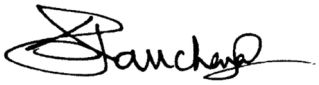
To celebrate our unwavering commitment to open science and public access to the results of federally funded research, the U.S. National Science Foundation (NSF) is proud to present this ***Public Access Plan 2.0 – Ensuring Open, Immediate and Equitable Access to National Science Foundation Funded Research***. This plan broadens the impacts of the vast range of research and education activities that NSF supports to keep the United States at the leading edge of science and engineering internationally.

NSF's updated public access plan integrates new agency guidance issued by the White House Office of Science and Technology Policy in August of 2022. This guidance, which includes zero-embargo public access for research publications and their supporting data, was developed with leadership from the National Science and Technology Council Subcommittee on Open Science, in which NSF has always been actively engaged. NSF developed its plan while considering issues of importance to our many partners across academia and industry, and in alignment with all other U.S. agencies which fund scientific research.

This plan is a first step, and we look forward to its further evolution as we address changes in technology and in the needs of members across our communities. Promoting immediate public access to federally funded research results and data is a critically important aspect of achieving the NSF mission of promoting the progress of science, securing the national defense, and advancing the national health, prosperity, and welfare. Indeed, scientific openness, academic freedom, scientific integrity, equity in science, and fairness are American values that rest on the pillar of public access to federally funded research and data.

I look forward to joining you as we continue to expand the frontiers of science and discovery at speed and scale. Together, our collective impact will create a stronger, more equitable scientific enterprise and drive positive change for decades to come.

Sethuraman Panchanathan



Director, National Science Foundation

Executive Summary

This ***U.S. National Science Foundation Public Access Plan 2.0*** has been prepared in response to the memorandum dated August 25, 2022, from the White House Office of Science and Technology Policy, or OSTP, titled ***Ensuring Free, Immediate, and Equitable Access to Federally Funded Research***, and signed by Alondra Nelson (referred to in this document simply as the 2022 OSTP memo).

This updated public access plan recognizes the trend towards openness inherent in today's scientific enterprise; the complex ecosystem of researchers, organizations, publishers and funders that comprise that enterprise; and the rapid pace of technological change continually altering the landscape of data sharing and publication. Public access to federally funded research and data is a pillar on which American values of scientific openness, academic freedom, scientific integrity, equity in science and fairness rest. Limiting access to scientific discovery limits the ability of all Americans to benefit from the returns on America's research and development investments. NSF is therefore committed to working with communities to make our shared federal scientific ecosystem stronger and more equitable for all.

The NSF public access plan recognizes that open and equitable communication of research results and data is inextricably linked to NSF's mission of promoting scientific progress. As described in the agency's current strategic plan, NSF is dedicated to funding research, developing STEM talent, fostering innovation and cultivating a science and engineering community based on shared values and strategic cooperation. This plan represents activities supporting these aims.

A primary consideration during the development of NSF's plan has been the potential equity impacts of public access implementation. NSF's goal is to improve equity throughout the research life cycle, making data and opportunities available to a broader range of researchers, including those from marginalized communities and under-resourced institutions of higher education in the U.S. NSF is considering potential impacts and unintended consequences and challenges to equity that the plan and its implementation may produce.

The sections of this plan describe how NSF will ensure:

- That all peer-reviewed scholarly publications resulting from NSF-funded research will be made freely available and publicly accessible by default in the NSF Public Access Repository, or NSF-PAR, without embargo or delay.
- That scientific data associated with peer-reviewed publications resulting from NSF awards will be made available in appropriate scientific disciplinary repositories.
- That exceptions to the data-sharing requirements will be made based on legal, privacy, ethical, intellectual property and national security considerations.
- That persistent identifiers, or PIDs, and other critical metadata associated with peer-reviewed publications and data resulting from NSF-funded research will be collected and made publicly available in NSF-PAR.
- That the agency coordinates with other federal funders of scientific research in implementing new public access requirements.

NSF looks forward to continuous improvement in the implementation of this plan in the coming months to ensure that publicly funded research is made freely and equitably available.

Section 1: Background and Policy Principles

General Introduction

NSF supports research and education in all branches of science and engineering, making it unique among federal science agencies. Public access to federally funded research and data is a pillar on which American values of scientific openness, academic freedom, scientific integrity, equity in science and fairness rest. NSF is committed to working with its partners to make the federal scientific ecosystem stronger and more equitable. ***NSF Public Access Plan 2.0*** has been prepared in response to the White House Office of Science and Technology Policy memorandum titled ***Ensuring Free, Immediate, and Equitable Access to Federally Funded Research*** (referred to in this document as the 2022 OSTP memo), dated Aug. 25, 2022 and signed by Alondra Nelson. For clarity and convenience, the section headings and numbering of this plan match the structure of the 2022 OSTP memo.

Public Access at the National Science Foundation

NSF's mission is "to promote the progress of science; to advance the national health, prosperity, and welfare; to secure the national defense; and for other purposes."[1] This mission is met by enabling scientific and technological discovery and innovation and by training the nation's scientific and engineering workforce. Public access to scientific discovery is a critical component of this mission: It reduces communication barriers, fosters new research and enables innovation that contributes to the economic prosperity and well-being of American citizens.

Since the release of the 2013 ***Memorandum on Increasing Access to the Results of Federally Funded Research*** by OSTP (2013 OSTP memo),[2] public access policy at NSF has been governed by ***Today's Data, Tomorrow's Discoveries*** (NSF report NSF 15-52).[3] That public access plan laid out an approach to:

- Create an extensible repository to serve as a single point of public access to research products resulting from NSF funding, starting with juried articles.
- Leverage available resources from federal agencies, universities, research institutes and the private sector.
- Build on existing policies and practices in ways that minimize burdens on awardees, while also broadening access to NSF-funded research findings with necessary and appropriate safeguards.

The development and implementation of NSF's Public Access Repository — the NSF-PAR system,[4] created and managed in partnership with the Department of Energy — met these goals for scholarly publications. Beginning with new awards starting January 2016, NSF-funded researchers have been required to deposit publications stemming from funded research in NSF-PAR as PDFs.[5] Either an official publisher's version of record (referred to as a VOR) or an author's accepted manuscript (referred to as an AAM, the final publication after peer review but before publisher's final formatting)[6] may be used to meet the requirement. Submitted publications have an automatic 12-month embargo as mandated by the 2013 OSTP memo. NSF-PAR is freely available to researchers and the public worldwide,[7] providing access to more than 200,000 articles arising from agency-funded research across the disciplines NSF serves.

In the 2022 OSTP memo, it was noted that the efforts made by federal agencies, including NSF, have had the intended benefits for the American public, contributing to a research culture shift away from data residing solely with principal investigators (known as PIs) toward data-sharing and collaboration. This shift has made over 8 million scholarly publications publicly accessible.[8] NSF is proud to be part of this shift toward a more equitable, responsive and transformative research culture.

The 2022 OSTP memo furthers this trend by calling for all federal agencies, to the extent consistent with applicable law, to:[9]

1. Update their public access policies as soon as possible, and no later than Dec. 31, 2025, to make publications and their supporting data resulting from federally funded research publicly accessible **without an embargo on their free and public release.**
2. Establish transparent procedures that ensure scientific and research integrity is maintained in public access policies.
3. Coordinate with OSTP to ensure equitable delivery of federally funded research results and data.

In considering how to advance these goals, NSF has reviewed internal administrative resources and cyberinfrastructure assets; solicited input from cognizant program officers, internal working groups, and a variety of publishers, scholarly societies and external researchers; and received additional public input via official NSF communication channels (e.g., webinars, email correspondence to publicaccess@nsf.gov and NSF officials, and in-person and virtual meetings).

This plan sets forth a framework for further improving public access to NSF-funded research. It builds on existing NSF policies and practices, primarily annual reporting requirements for award recipients. NSF will continue to identify additional approaches to increasing equitable public access to research products, including peer-reviewed scholarly publications and the data supporting such publications. In cooperation with other federal agencies, NSF will continue deliberating how the storage, preservation, discoverability and reusability of research products can be improved and streamlined, as well as how publications stemming from these activities can best be equitably shared among researchers and the public.

NSF Strategic Plan and Public Access Objectives

Improving public access to scientific products aligns with the goals set forth in the current NSF strategic plan:[10]

1. **Empower**: Empower STEM talent to fully participate in science and engineering.
2. **Discover**: Create new knowledge about our universe, the world and ourselves.
3. **Impact**: Benefit society by translating knowledge into solutions.
4. **Excel**: Excel at NSF operations and management.

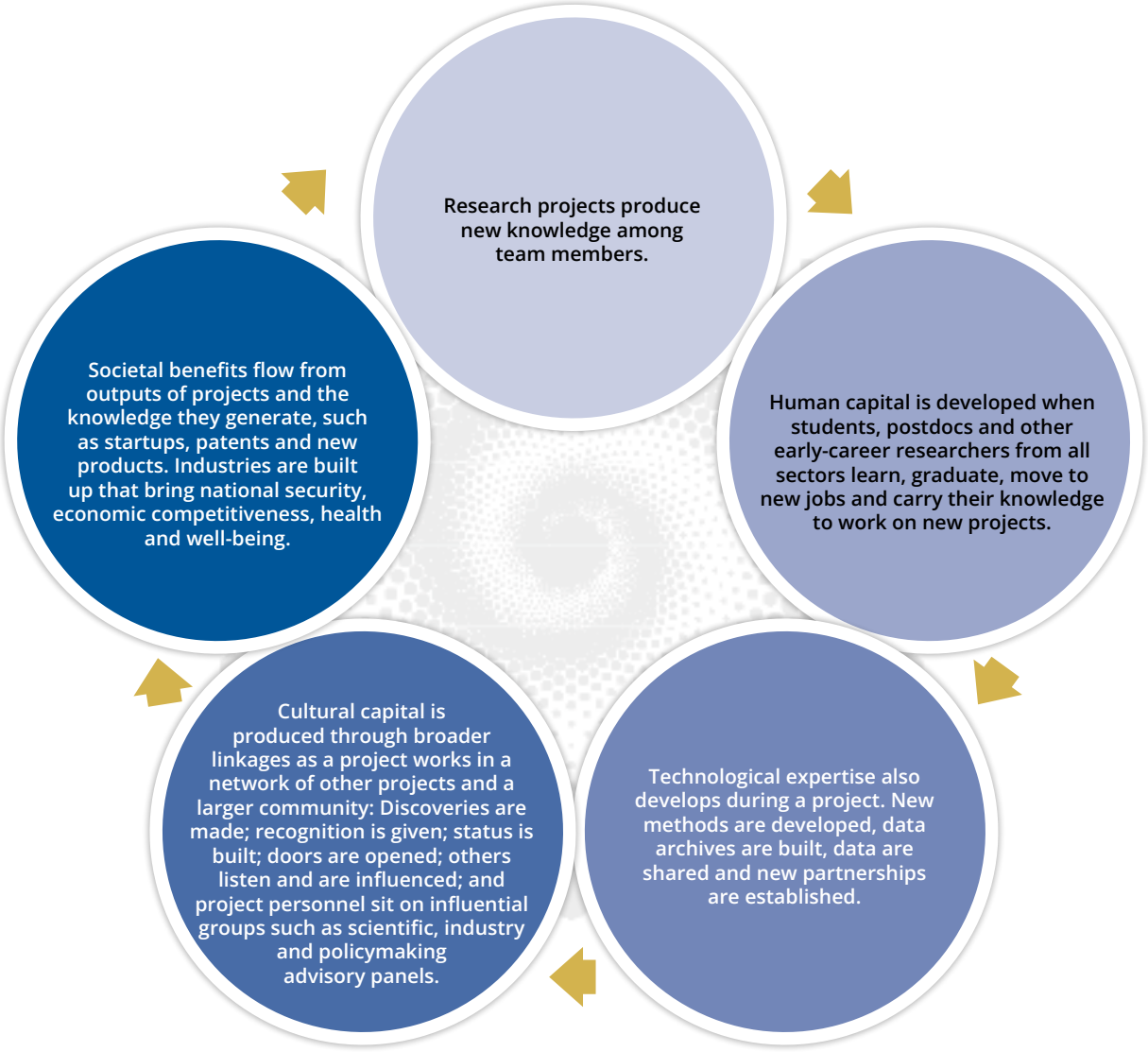


Figure 1: *Spiral of Knowledge Mobilization, from NSF Strategic Plan for Fiscal Years 2022 – 2026 (p. 20).*

Efforts that provide greater access to publications and associated data can lower barriers for groups currently under-represented and/or underserved in STEM and help empower a diverse, talented science and technology workforce. Scientific discoveries can be more rapidly translated into innovative solutions and novel technologies through public access to research products. Promulgation of this culture of sharing also advances the global research ecosystem and the data-sharing cyberinfrastructure found within it, with NSF-funded researchers advancing research practices and impacts in scholarly communities.

Section 2: Open Science Responses – Opportunities and Equity

2A. COVID-19 and Other Threats to National Health, Prosperity and Security

As described in the 2022 OSTP memo:

> Immediate public access to COVID-19 research is a powerful case study on the benefits of delivering research results and data rapidly to the people. The insights of new and cutting-edge research stemming from the support of federal agencies should be immediately available—not just in moments of crisis, but in every moment. Not only to fight a pandemic, but to advance all areas of study, including urgent issues such as cancer, clean energy, economic disparities, and climate change. American investment in such research is essential to the health, economic prosperity, and well-being of the Nation. There should be no delay between taxpayers and the returns on their investments in research.

NSF-funded research was a critical part of the national response to the COVID-19 pandemic.[11] Further, NSF participated in national initiatives to provide immediate public access to related research such as the creation of the public CORD-19 Database[12] and the COVID-19 High Performance Computing Consortium[13] and funded the creation of the COVID Information Commons,[14] a broad database of the hundreds of NSF RAPID[15] awards relating to the pandemic.

The national crisis response to the COVID-19 pandemic is indeed a case study to be considered with regard to many other urgent challenges which demand attention, such as cancer, clean energy, economic disparities and environmental sustainability. NSF is well-aligned to address the wide disciplinary range of these challenges, since it provides support for research and education in all branches of science and engineering. Improved public access to the products of NSF-funded work accelerates the dissemination of fundamental science, assisting researchers in advancing knowledge, innovation and technology across disciplines. It has the potential to be transformative for the nation's prosperity. For example, NSF funding of cyberinfrastructure development to advance access to research data has recently been highlighted as impactful for the geosciences, where NSF's investment "motivated communities to move towards cyberinfrastructure and open science in general."[16]

This updated public access plan recognizes the trend toward openness inherent in today's scientific enterprise; the complex ecosystem of researchers, organizations, publishers and funders that comprise that enterprise; and the rapid pace of technological change that is continually altering the landscape of data-sharing and publication. Fundamentally, it recognizes that open and equitable communication of research results and data is inextricably linked to NSF's mission of promoting scientific progress. As described in the agency's current strategic plan, NSF is dedicated to funding research, developing STEM talent, fostering innovation and cultivating a science and engineering community based on shared values and strategic cooperation. This plan represents some of the activities supporting these aims.

2B. Equity Opportunities and Concerns

A primary consideration during the development of NSF's plan has been the potential equity impacts of public access implementation. It is the explicit goal of NSF to improve equity throughout the research life cycle, making data and opportunities available to a broader range of researchers, including those from marginalized communities and under-resourced institutions of higher education in the U.S. As such, the agency has considered potential unintended consequences and challenges to equity that the plan and its implementation may produce.[17] NSF is committed to examining unintended equity effects or impacts of new public access requirements such as those highlighted in the National Academies of Science, Engineering, and Medicine, or NASEM, 2018 report ***Open Science by Design***:[18]

> Certain approaches to implementing open publication have the potential to affect the research ecosystem in significant ways, with differential impacts on different stakeholders. For example, a system that strongly favors publication approaches based on the payment of article processing charges would favor established researchers and wealthy institutions over early career researchers and institutions with fewer resources. In planning new policies and transitions, it will be necessary to anticipate differential impacts to the extent possible, consider ways of avoiding these, and build in evaluative and corrective mechanisms to address unanticipated consequences.

NSF believes that all researchers should have equitable opportunities to publish their work and access research findings regardless of funding availability, discipline, organizational affiliation or institutionalized barrier. To inform the development of this plan, NSF has initiated both in-reach and outreach activities to understand the equity challenges that may arise and the concerns NSF communities have about the public access policy. Primary areas of interest that will shape NSF policy as implementation approaches are formulated include:

- **Minimizing the equity impact of over-reliance on article processing charges, or APCs,** also known as the "Gold Open Access" publication model, including inequity for fields, organizations or researchers lacking access to funding; consequences of possible citation bias; the impact on ability to fund research and training activities; and potential negative impacts with respect to public trust.

- **Promoting use of author's accepted manuscripts, or AAMs,** as a no-cost option to comply with public access requirements.

- **Minimizing the consequences of changing publishing ecosystems,** including impacts for organizations least able to weather dramatic changes to subscription policies, which can increase precarity for those affiliated with these organizations.

- **Ameliorating the possible impacts of large APCs** on small awards.

- **Involving affected communities** regarding issues associated with data collection, data governance, verifying permitted data access, and data destruction, particularly for groups that have previously suffered from the appropriation or misuse of data.[19]

- **Ensure accessibility of data and results,** including access to data cyberinfrastructures for under-resourced and underserved institutions/researchers, as well as considerations for persons with visual disabilities.[20]

- **Maximizing the reach and impact of U.S. research** while seeking to minimize both access barriers in under-resourced and underserved communities and challenges related to the language or interpretability of data.

- **Identifying the full range of costs (tangible and intangible)** associated with data provision and addressing any inequities introduced by these costs.

- **Developing processes for addressing inequities** identified in sharing and accessing research findings.

NSF acknowledges that ensuring equity — for those wishing to access information as well as for researchers required to deposit and share it — will require significant investments in data cyberinfrastructure and training, particularly for the benefits of public access to reach local, public or marginalized communities.[21] Indeed, as the previously cited 2018 NASEM ***Open Science by Design*** report stated:

> Research institutions should work to create a culture that actively supports Open Science by Design by better rewarding and supporting researchers engaged in open science practices. Research funders should provide explicit and consistent support for practices and approaches that facilitate this shift in culture and incentives.

In support of such sentiments, NSF expects to advance

- findable, accessible, interoperable, reusable — or FAIR[22] — data principles;
- open science practices and research;[23]
- training opportunities, workforce development, and workshops around data cyberinfrastructure skills and literacy;[24]
- transparency of costs and impacts of public access practices to identify and remedy equity-related barriers.

NSF will enhance reporting and review procedures for NSF awards around data sharing. NSF is committed to examining and addressing unintended inequities that may result from new public access requirements.

Section 3: Increasing Equitable Access to Federally Funded Research Results

NSF will implement the new OSTP guidance regarding federally funded research results prior to the OSTP goal of Dec. 31, 2025. The agency's policy development for implementation of provisions in Section 3 will be completed and published before Dec. 31, 2024. The effective date of NSF's implementation will occur no later than Jan. 31, 2025.[25] These efforts will include activities identified in Section 5 of this plan, including interagency collaboration, public engagement and OSTP feedback to the agency as the plan is developed.

3.A. Peer-Reviewed Scholarly Publications

As required by OSTP, and as appropriate and consistent with applicable law, NSF will ensure that all peer-reviewed scholarly publications resulting from agency-funded research are made freely available and publicly accessible by default in NSF-PAR without any embargo or delay.

3.A.i. How Peer-Reviewed Scholarly Publications Will be Made Publicly Accessible

Building on longstanding practice and expectations, NSF will continue to require that award recipients make peer-reviewed scholarly publications publicly accessible through annual reporting requirements and procedures in Research.gov. These publications are integrated with NSF-PAR. NSF's revised award terms and conditions will require NSF awardees to report journal articles, as well as papers in juried conference proceedings, and make this content freely accessible through NSF-PAR with no embargo as part of the updated reporting requirements. Submitting either the publisher's version of record or the author's approved manuscript to NSF-PAR will satisfy the public access requirement. NSF award recipients will be obligated to make deposits upon submission of the next annual project report covering when the article was accepted for publication.

As has been the case since the implementation of the previous public access plan, publications submitted to NSF-PAR are publicly accessible and searchable via a web-based interface, and the metadata for deposited material are available without embargo. Award recipients will also still be able to make deposits to NSF-PAR after an award has been closed, though NSF has no authority to require actions after the period of performance of an award. NSF will encourage proposers to include relevant NSF-PAR products, including any submitted following the close of prior awards, in the "Results from Prior NSF Support" and "Biographical Sketch(es)" sections of future proposals as allowable, which are considered in the merit review process.

NSF will change the current system to implement the zero-embargo requirement by removing the default one-year embargo period, making publications freely and immediately accessible to the public. NSF will engage prospective proposers, sponsored project administrators and other partners through enhanced outreach to help them adapt to these changes and to relevant new features of the NSF-PAR system as they are introduced. NSF also will pursue in-reach with program and administrative staff to ensure appropriate management of the proposal and award process, in alignment with updated requirements and NSF-PAR workflows.

The 2022 OSTP memo defines "Peer Reviewed Scholarly Publications" with the following footnote regarding additional research output formats on page 4:

> Such scholarly publications always include peer-reviewed research articles or final manuscripts published in scholarly journals, and may include peer-reviewed book chapters, editorials, and peer-reviewed conference proceedings published in other scholarly outlets that result from federally funded research.

The category described in the NSF Research Performance Progress Report specification[26] as "non-periodical, one-time publications" is already broadly inclusive of many such document types. One publication type relevant to NSF awards is the final product of conference and workshop awards. To date, these reports have not been made systematically accessible and may not be peer-reviewed. NSF has already made plans to add such conference/workshop proceedings and reports as a research publication format available for public access and dissemination in the NSF-PAR 3.0 system. Following the guidance of the 2022 OSTP memo, awardees will be required to make ***juried*** conference proceedings publicly accessible with no cost or embargo period. This includes both individual peer-reviewed papers resulting from an NSF award to a conference and, in the case of a conference award, the peer-reviewed proceedings or workshop report in its entirety. NSF also is considering extending its public access requirements to publication types such as book chapters and ***non-juried*** conference proceedings, as well as other document formats and additional research outputs (e.g. software) which award recipients regularly report in annual and final annual project reports. Future versions of NSF reporting processes may be extended to cover these and additional research output types in NSF-PAR and Research.gov, for purposes of public access.

3.A.ii. Maximizing Public Access to Scholarly Publications With Machine-Readability

Text-mining and other analytical work utilizing text derived from the underlying XML content of scientific articles have been an active and growing area of interest in the research community.[27] The 2022 OSTP memo requires agencies to provide free public access to articles in machine-readable formats, with details concerning this format in a footnote on page 4:

> "Machine readability" refers to a format that can be easily processed by a computer without human intervention while ensuring no semantic meaning is lost (such as the NISO Z39.96-2015 JATS XML standard currently used by PubMed Central)."

Accordingly, NSF plans to provide free public access to NSF-PAR article content and article metadata expressed as machine-readable XML, beginning with the effective date of implementation, as previously described. NSF will explore the specific mechanisms for providing this access in the intervening period, and prior to this effective date, report to OSTP the specifics of how this XML content will be made available. NSF is working with the National Institutes of Health, Department of Energy, and other agencies to better understand possible XML conversion mechanisms to accomplish this goal.

3.A.iii. Re-use Rights for Publications Made Freely and Publicly Available

The 2022 OSTP memo requires agencies to clarify the "circumstances or prerequisites needed to make the publications freely and publicly available by default, including any use and re-use rights, and which restrictions, including attribution, may apply." NSF strongly agrees with the broad aims of scientific transparency, reproducibility and replicability, and other aspects of open science encompassed in Section 3.A.iii of the 2022 OSTP memo.[28] To this end, NSF has developed the following clarifications and suggestions for interagency alignment on circumstances and prerequisites going forward.

To make publications freely and publicly available under financial assistance agreements subject to 2 CFR Part 200, ***Uniform Administrative Requirements, Cost Principles, and Audit Requirements for Federal Awards*** ("Uniform Guidance"),[29] there needs to be federal-wide consensus that republishing federally funded works falls within the existing rights and requirements in 2 CFR 200.315(b), specifically that the republication would be made "for federal purposes." Lacking that ability, changes would need to be made to the Uniform Guidance to ensure that agencies have legal authority, as well as to minimize the risk of legal actions from award recipients and publishers.

With that agreement, federal funding agencies would then need to agree on a common licensing scheme, such as Creative Commons Attribution License (e.g. CC-BY)[30] or other open licensing to allow for reuse.

NSF stands ready to engage with the development of such a shared understanding of reuse rights and welcomes interagency alignment processes recommended in Section 5 of the 2022 OSTP memo.

3.B. Scientific Data

NSF acknowledges the central role of scientific data in the research ecosystem.[31] Freely sharing data arising from awards has been an important public access goal for NSF for years, as reflected in the agency having added the capability to make datasets publicly accessible in NSF-PAR version 2.0 (implemented in December 2021). Notably, improved data-sharing advances core elements of NSF's strategic plan, reducing barriers to participation in science, adding directly to available scientific knowledge and increasing the transparency and impact of NSF-funded research. Ideally, it is desired that datasets from all NSF-funded research be readily accessible, well-curated and structured such that they are useable by a wide variety of researchers and experts. Achieving this ideal will require ongoing attention to create data-sharing mechanisms that provide necessary protections for privacy, security (national and economic), and confidentiality as appropriate and consistent with applicable law and policy. These mechanisms must also limit unintended consequences,[32] while accounting and allowing for the diversity, complexity and sizes of datasets collected across the breadth of NSF-funded research.

NSF currently has in place agency-wide requirements for the management and sharing of scientific data resulting from awards that are specified in the ***NSF Proposal & Award Policies & Procedures Guide***, or PAPPG.[33] These requirements include broad expectations for the dissemination and sharing of research results, as well as instructions for preparation of required plans for data management and research product sharing found within all proposals. As described in the PAPPG, NSF directorates, offices, divisions, programs or other units may communicate more specific data management requirements via guidance provided on the NSF website. As described below, this approach accommodates the various circumstances of different research domains, including the use of discipline-specific data repositories and expectations for appropriate data sharing.

NSF's annual and final annual project reporting processes currently provide researchers with the optional capability to make datasets publicly accessible and discoverable by recording persistent identifiers (e.g., digital object identifiers, or DOIs) for such datasets, wherever they may be archived. A key distinguishing feature of NSF's approach to making such research datasets publicly accessible is leveraging the significant investments that NSF has made in disciplinary data repositories. In the 2021 implementation of public access for datasets in NSF-PAR, the agency was intent on neither undermining nor replacing long-established disciplinary repositories. Rather, NSF-PAR's approach to indexing metadata — including the use of persistent identifiers, or PIDs, assigned to research datasets — leverages a federated approach whereby research data reside in appropriate repositories and NSF-PAR serves as a point of discovery.[34] This approach is efficient and powerful, and allows the datasets to be managed by or in close conjunction with the disciplinary communities that create and use such data while simultaneously making it freely available to additional user groups. However, this approach also relies on NSF continuing to provide adequate resources to repositories to ensure appropriate data preservation and access over time, per Sec. 3.B.iii of this public access plan, as well as the availability of domain-specific expertise to curate meaningful datasets within scholarly communities.

3.B.i. Scientific Data Underlying Peer-Reviewed Scholarly Publications

In compliance with the 2022 OSTP memo, annual and final annual project reporting for scientific data associated with peer-reviewed scholarly publications which result from NSF awards is mandatory. However, also as stipulated by the 2022 OSTP memo, this requirement will be waived if such data are subject to limitations as described in Section 3.C.ii, and will further be subject to NSF guidelines for researcher responsibilities regarding data management and sharing plans.

Deposited datasets must also include acknowledgement of NSF support, the award number and appropriate attribution to facilitate reporting, and will be required to be submitted via a DOI when deposited into NSF-PAR. NSF-PAR harvests metadata associated with DOIs for deposited datasets, which uniformly include a minimum set of data elements. As research communities continue to refine metadata standards for datatypes and datasets, NSF will take steps to identify common and informative elements and incorporate these in NSF-PAR as technologically feasible and as resources allow.

These new requirements will entail evaluation of several criteria on the part of NSF award recipients during their annual and final annual reporting activities to discern whether they must provide access to a particular dataset (e.g., was it used in the research underlying an article, was it developed using funds from the award, is it subject to federal access restrictions, etc.). Therefore, preparing and providing careful guidance to award recipients in making such determinations will be essential to the success of this new requirement prior to implementation in 2025. NSF will develop procedures to request reasonable accommodations to delay the public release of data which have access limitations associated with them, devoting careful attention to minimize confusion and maximize positive public access outcomes.

3.B.ii. Scientific Data Not Associated With Peer-Reviewed Scholarly Publications

As required in the 2022 OSTP memo, NSF will develop approaches and timelines for sharing other types of federally funded scientific data not associated with peer-reviewed scholarly publications. There is now a broad acknowledgement of the critical importance of sharing scientific datasets that are well-documented and reusable to both confirm the reproducibility and replicability of results and make new discoveries. Scientific datasets are now considered a primary research output in their own right, at least co-equal if not more important than scientific publications.[35]

However, scientific datasets are far more varied than publications in terms of size, complexity, potential ephemerality and possible access restrictions. NSF is aware that many excellent repositories already exist for some subsets of data, and that using existing repositories will help many researchers meet public access requirements. Therefore, award recipients will be directed to continue using disciplinary repositories for storing and maintaining access to data, along with PIDs and appropriate metadata.[36] As the scientific community further evolves its understanding and practices of sharing scientific datasets, it will be incumbent on NSF to support such efforts as a process of discovery.

NSF will continue to develop additional guidance for other types of datasets, for example those not related to a particular publication but associated with work funded by the NSF award and those not funded by NSF but which a researcher would like to deposit in conjunction with NSF-funded datasets.

3.B.iii. Digital Repositories for Scientific Data

Managing data is complex and will require further exploration and development, given the heterogeneity in form, variability in size (including the existence of potentially very large datasets) and extent to which retaining or sharing data is practical. As public access to research datasets continues to grow, NSF must continue to foster best practices in the creation and maintenance of digital repositories. NSF was an active participant in the development of the National Science and Technology Council document ***Desirable Characteristics of Data Repositories for Federally Funded Research***[37] and fully endorses this starting point for understanding basic properties of such repositories, though development of additional federal guidance in this area is necessary.

Preservation of digital content is an area of active research which will require additional support as federal agencies address the requirements of the 2022 OSTP memo. NSF fosters further development of broad guidelines and disciplinary-specific norms to shape communities of practice around data description and management,[38] including data appraisal, curation, retention and disposal or deaccessioning of data. These priorities are expected to be reflected both in repositories' policies and in the individual data management plans required for all NSF proposals. NSF will, therefore, continue to encourage new approaches to achieving long-term storage and preservation for data. NSF will also consider strategies for facilitating long-term data preservation and storage.

3.B.iv. Federal Agency Research

NSF does not operate laboratories of its own or directly employ researchers, and most scientific data arising from NSF funds originate in extramural awards. NSF does, however, house one of the federal government's statistical agencies, the National Center for Science and Engineering Statistics.[39] Data collected by NCSES responds to statutory requirements, and, unless constrained by privacy or reliability concerns as described in Sections 3.C.i and 3.C.ii, are made available for secondary use by researchers as public use files.

3.C. Researcher Responsibilities for Managing and Sharing Data Arising From NSF Awards

National security interests and restrictions in law or regulation that will place boundaries on the NSF-supported data that can be shared publicly are described in Sections 3.B.iv and 3.C.i. Within these limits, beginning in 2025, NSF will require that researchers deposit scientific data — as defined in the 2022 OSTP memo and consistent with the norms in the scientific community of the funded researcher[40] — undergirding peer-reviewed publications deposited in NSF-PAR in data repositories, as noted in Section 3.B.i. NSF will additionally provide guidance on optionally sharing other data.

3.C.i. Legal, Privacy, Ethical, Technical, Intellectual Property, and Security Limitations on Access to Scientific Data Arising From NSF Awards

NSF understands that practices surrounding data depositing — and the legal, privacy, ethical, technical and security limitations at play — will vary substantially based on the discipline and type of data under consideration (simulation, modeling, observational, instrument-driven, laboratory, survey-based, personally identifiable, commercially owned, culturally sensitive, etc.). This presents a particular challenge for producing generalized NSF guidance for funded researchers, given the breadth of research disciplines supported by the agency. NSF expects to meet this challenge by continuing the approach of articulating general agency-wide principles for data management and sharing in the PAPPG and allowing NSF directorates, offices, divisions, programs or other units to articulate more specific data management requirements as needed.

Limitations to the release of the types of data specific to NSF include, but are not limited to:

- Programs that allow for exceptions to release of proprietary or otherwise restricted data, including but not limited to personally identifiable information, confidential business information and security information. Examples include Small Business Innovation Research and Small Business Technology Transfer awards.
- Restrictions on data access, use or disclosure as stipulated by an Institutional Review Board.
- Data collected by NCSES subject to terms of a restricted use license or other confidentiality agreement, including data catalogued in the Standard Application Process portal.[41]
- Data that was used to undergird a publication but which was not funded by NSF, including commercially purchased data resources which may have been used in the course of experiments or analysis.

Equity and Unintended Consequences

The data sharing and security measures required by this policy have the potential to impose extensive time, effort, and monetary costs to appropriately safeguard confidential or sensitive data. This requirement may disadvantage researchers lacking access to institutional support in this area, and could discourage already disadvantaged researchers from pursuing research involving confidential data or otherwise restricted data.

NSF will continue to develop guidance concerning management of and access to sensitive information, building on internal NSF expertise.

3.C.ii. Plans for Maximizing Appropriate Data and Providing Risk-Mitigated Opportunities for Limited Data Access

NSF will explore approaches for risk-mitigated data access as suggested in the 2022 OSTP memo, including "secure research data centers, data use agreements, perturbing identifiable information, or excluding sensitive variables."[42] Notably, NCSES grants licenses for secure access to restricted use data under an existing framework for protecting statistical data and disseminating the data securely and equitably. Licenses are granted in accordance with the "Confidential Information Protection and Statistical Efficiency Act of 2018" and other provisions within the "Foundations for Evidence-Based Policymaking Act of 2018." NSF will also consult with America's DataHub Consortium, a collaborative project sponsored by NCSES,[43] to determine how federal statistical data can be shared and utilized by the public and research communities. NSF funds the Inter-university Consortium for Political and Social Research at the University of Michigan for the Research Data Ecosystem, a software platform that will enable access to and analysis and dissemination of social science data in an integrated fashion, increasing data access while providing appropriate confidentiality and control for sensitive data.[44]

3.C.iii. Specific Online Digital Repositories

Data management plans, or DMPs, are currently a required component of an NSF proposal.[45] To ensure that funded researchers have considered data deposition and appropriate repositories, NSF plans to more explicitly and systematically require researchers to include, as part of these DMPs, details about how the data produced will be shared publicly and with other researchers and anticipated repositories for the data, and to ensure that costs associated with the chosen repository are reflected in the proposal budget request, as described in the PAPPG, per Section 3.D. To better align with other federal agencies, NSF will retitle such plans data management and sharing plans, or DMSPs. New requirements associated with the 2022 OSTP memo for making datasets publicly accessible will be conveyed to researchers via the standard NSF communication tools, such as the NSF Electronic Research Administrators Forum series, which NSF regularly hosts nationally with university research administrators.

NSF will additionally consider whether DMSP requirements can be updated. Areas under consideration include: whether plan requirements can or should be harmonized across the agency while still providing for program-level specifications; possible additional guidance to researchers regarding data repositories and costs; and whether additional tools are needed for developing and reviewing DMSPs by review panels and NSF program directors.[46] This may reduce the burden incurred by DMSP procedures and standardize DMSP evaluation.

NSF acknowledges that determining researcher compliance with existing DMPs remains a challenge, and the agency will consider compliance during the development of implementation plans. Specifically, NSF will explore the possibility of instituting or allowing the use of machine-actionable DMSPs to assist in agency compliance checking.

3.D. Allowable Publication and Data Costs

To promote equity in terms of researchers' ability to comply with data-sharing requirements, NSF will continue its current policy that reasonable data curation and publication costs can be proposed. The PAPPG currently allows proposers to request funds for:[47]

> The costs of documenting, preparing, publishing or otherwise making available to others the findings and products of the work conducted under the award. This generally includes the following types of activities: reports, reprints, page charges or other journal costs (except costs for prior or early publication); necessary illustrations; **cleanup, documentation, storage and indexing of data and databases**; development, documentation and debugging of software; and storage, preservation, documentation, indexing, etc., of physical specimens, collections or fabricated items. **Line G.2. of the proposal budget also may be used to request funding for data deposit and data curation costs.**

NSF will participate in discussions with other federal agencies on the topic of reasonable data curation costs. NSF recognizes the many equity concerns associated with large costs of article page charges and data deposition costs and will seek to address such concerns to the extent possible.

3.E. Ongoing Reporting to OSTP on Request Regarding Plan Status

Following Office of Management and Budget guidance in the 2022 OSTP memo, when requested, NSF will report to OSTP on the status of public access plans and policy implementation, including the number of scholarly publications funded by NSF and other relevant statistics collected. NSF currently co-chairs the National Science and Technology Council Subcommittee on Open Science and will continue to utilize this interagency forum to discuss and harmonize, where appropriate, approaches to the public access plan, as it is implemented.

NSF is currently exploring data gathering mechanisms to track compliance with the existing public access plan to better shape implementation, as well as expansion to meet new requirements. Specific areas under examination include the degree to which NSF-PAR-deposited publications are linked to specific NSF awards via machine-readable metadata, the proportion of NSF-PAR submissions that lack critical metadata, and mechanisms to assess the impact and scope of NSF-PAR (e.g. the number of publications funded by the agency). The results of such analyses are intended to inform additional updates to NSF policy and NSF-PAR technical requirements.

Section 4: Ensuring Scientific and Research Integrity

NSF will implement the 2022 OSTP memo requirements for scientific and research integrity in accordance with the specified timeline. As required by the 2022 OSTP memo, agency policy development for plans implementing provisions in Section 4 will be completed and published before December 31, 2026, with an effective implementation date of by January 2027. These efforts will also include activities identified in Section 5 of this plan, including interagency collaboration, public engagement, and OSTP feedback to the agency.

4.A. Collect and Make Publicly Available Appropriate Publication Metadata

4.A.i. Names, Organizational Affiliations and Funding

Currently, metadata for juried articles is collected and made publicly accessible through NSF-PAR. The system also provides entry fields for relevant PIDs, notably DOIs, although these are not currently required. The additional metadata elements required in Section 4.A.i of the 2022 OSTP memo are excellent citational elements and will be added to the feature specifications for the relevant version of NSF-PAR slated for deployment before the end of 2027 to comply with these requirements.

NSF has already taken steps to require additional disclosure of researcher support from foreign governments, beginning January 2023:[48]

> Consistent with NSPM-33, senior personnel are required to disclose contracts associated with participation in programs sponsored by foreign governments, instrumentalities, or entities, including foreign government-sponsored talent recruitment programs. Further, if an individual receives direct or indirect support that is funded by a foreign government-sponsored talent recruitment program, even where the support is provided through an intermediary and does not require membership in the foreign government-sponsored talent recruitment program, that support must be disclosed. Senior personnel must also report other foreign government sponsored or affiliated activity. Note that non-disclosure clauses associated with these contracts are not acceptable exemptions from this disclosure requirement.

4.A.ii. Publication Date(s)

NSF award recipients must specify the date of publication for articles in peer-reviewed scholarly journals and papers in juried conference proceedings as part of their annual and final annual project reports, as well as the associated deposition of publications in NSF-PAR. Like other metadata, these dates are visible in search results. If the published paper is versioned, for any reason, and a new version subsequently submitted, the new date is likewise recorded in a new entry.

4.A.iii. Identifiers for Research Outputs

NSF-PAR provides entry fields for PIDs, such as DOIs, although these are not currently required. Requiring PIDs for all research outputs has benefits for research integrity in the long term, but is not yet a uniform practice across all research communities. NSF will seek to implement emerging best practices in this area,[49] especially as identified in joint interagency discussions as described in Section 5 of this plan.

4.B. Obtaining an NSPM-33 Digital Persistent Identifier for Award Senior Personnel

NSF is currently in the process of developing compliance plans for federal requirements as set forth in NSPM-33 instructions[50] in accordance with implementation guidance issued in 2022 by the National Science and Technology Council.[51]

As part of this compliance planning, a timeline is being developed for gathering personal PIDs from proposal senior personnel. Senior personnel will be instructed to include PIDs as part of the biographical sketch developed in SciENcv.[52] Biographical sketches are currently required for every NSF award, and researchers already have the option to use the SciENcv format to meet this requirement. The optional SciENcv format will be required after Oct. 23, 2023. The SciENcv format is currently being updated to serve as the NSF/National Institutes of Health-compliant response to NSPM-33 guidance. NSF is working with NIH and the Department of Energy to explore linking SciENcv to NSF-PAR to allow these important metadata to be shared across platforms.

4.C. Assigning PIDs to Research and Development Awards

PIDs are long-lasting URLs that uniquely resolve to specific informational resources. Currently, NSF proposals and awards are assigned a federal award identification number, also known as a FAIN, that allows them to be tracked in internal and federal-wide systems. These award numbers are currently made publicly accessible via the NSF award search interface.[53] However, the proposal and award numbers are not generally considered PIDs because the numbers by themselves are not actionable URLs. NSF will implement this requirement before the end of 2027 by adding DOI functionality to the current process for assigning proposal numbers. NSF will also provide an update as to these specifics to OSTP before the end of 2024.

The move to PIDs for NSF awards complements the "Federal Funding Accountability and Transparency Act" and will allow transparency around research funded by NSF on NSF-PAR and other federal government aggregation services. While this transition to unique PIDs for awards is accomplished, NSF will continue to make NSF award numbers visible as metadata in NSF-PAR to provide similar transparency and access to appropriate metadata.

Section 5: Public Access Plan Coordination Among Federal Agencies

NSF enjoys strong representation on the OSTP Subcommittee on Open Science, which is the locus for interagency coordination around the new public access guidance, as laid out in the 2022 OSTP memo.

Additional topics to be discussed in this forum include:

- developing guidance for data not associated with a peer-reviewed publication,
- determining reasonable costs for data curation and management,
- further leveraging SciENcv, and
- mechanisms to engage researchers and impacted communities.

Equity and Unintended Consequences
All federal funding agencies may face complications for international collaborations if data collected or compiled by international collaborators cannot be shared publicly. Additionally, many international collaborations utilizing federal support center marginalized communities. NSF believes that interagency cooperation is imperative for determining the impact of public access policy on international research and mitigating any undesired impact.

NSF will continue to work with OSTP and OMB to determine the cost implications of the plans outlined in Sections 3 and 4. Strategies such as converting documents to machine-readable formats and making a persistent identification service for NSF award numbers may have substantial associated costs. Costs for building robust cyberinfrastructures to advance equitable sharing within specific NSF-supported research areas — including those unique to NSF due to the breadth of the agency mandate — are also expected to be significant.

NSF will work to ensure that open access policies are sustainable and understood by the public in coordination with other federal agencies through participation in the 2023 Year of Open Science, a commemorative designation initiated by NASA and expanded government-wide by the White House.

Furthermore, NSF will carefully consider how implementing OSTP guidelines will impact the scientific community. NSF recognizes that some journals, such as smaller professional society publications, will find agency compliance with the 2022 OSTP memo challenging to their existing business models. NSF is committed to obtaining input from the community and will continue its current outreach program of webinars and town halls to provide impacted individuals and institutions a venue to express the needs of the communities they represent and support.

Section 6: General Provisions

Information

As noted in the 2022 OSTP memo, the new provisions regarding public access are not intended to fully replace provisions put in place following the 2013 OSTP memo. Any requirement currently contained in NSF's guidelines for researchers, award terms and conditions, and reporting requirements not directly identified in this plan will remain in place, though efforts have been made to include salient details in this document.

The requirements laid out in Section 3 will apply to new awards resulting from proposals submitted or due on or after the **January 2025** effective date; the requirements of Section 4 will apply to awards for which proposals were submitted or due after **January 2027**.

NSF requires award recipients to adhere to public access requirements resulting from this plan for all research that it supports, even when the research underlying the publication and generation of data is supported by multiple funders, such as other federal and state agencies, private funders and international sponsors.

Authority

NSF is an independent federal agency created by the "National Science Foundation Act of 1950, 42 USC 1861, et seq.", also known as the "NSF Act." The "NSF Act" authorizes the agency to make discretionary awards to support basic scientific and engineering research and education programs and to impose appropriate terms and conditions on those awards. Other relevant legal and policy authorities include, but are not limited to:

- 2 CFR Part 200.
- 45 CFR Part 602.
- 42 USC 19053 (Title 3, Section 10344 of the "CHIPS and Science Act of 2022").
- The America COMPETES Reauthorization Act of 2010.
- OSTP Memo, February 22, 2013: ***Increasing Access to the Results of Federally Funded Scientific Research***.
- OSTP Memo, August 25, 2022: ***Ensuring Free, Immediate, and Equitable Access to Federally Funded Research***.
- The ***Proposal and Award Policies and Procedures Guide***.
- General Award Terms and Conditions.

Roles and Responsibilities

National Science Board — Approves all NSF policies necessary to implement this plan.

Public Access and Open Science Working Group — Provides oversight and guidance on the implementation of this plan and informs senior leadership on progress, resource needs and other issues that may influence the plan's implementation.

Program staff — Serve as the primary interface with principal investigators and assist in overseeing compliance with the plan's requirements.

Office of Budget, Finance and Award Management — Develops and implements NSF-specific public access policy.

Office of Information and Resource Management — Ensures IT systems support public access policies and enable compliance and compliance checking.

Principal investigators and authors — Work with the publisher before any publication rights are transferred to ensure that all of NSF's public access requirements can be met; ensure that any agreement signed with a publisher preserves the author's ability to comply with NSF's public access requirements.

Recipient organizations — Ensure that authors and data generators are aware of and comply with NSF's public access requirements and ensure compliance with all award terms and conditions of awards, including the submission of final peer-reviewed manuscripts that arise directly from the organization's awards, release of any associated data and overall compliance with DMSPs.

Section 7: Taking Next Steps Together

NSF is committed to a process of continuous improvement in the implementation of this plan to ensure that publicly funded research is freely and equitably available. This plan will be adapted or modified as the impacts of implementation become clearer. As this plan is implemented, NSF will continue to solicit feedback from all relevant parties, including OSTP, OMB, other federal agencies, NSF staff, NSF-funded researchers, publishers, scholarly societies, research communities, students and members of the general public and make necessary changes.

Additionally, NSF will continue to fund research into policies and networks that make public data actionable for communities of researchers and U.S. citizens broadly, determining what works to advance equitable public use of data and what barriers exist. NSF believes that improved public access will help in communicating with Americans about

the work NSF funds and its impacts, and acknowledges that resources will need to be committed to evaluating the progression of public access policies to enable this.

Timing of any amendments and future implementations will be based on ongoing implementation experiences and feedback and on the recommendations of relevant NSF internal working groups. Changes and updates to this plan will be made available via the NSF Public Access Initiative website,[54] and NSF will continue to provide avenues for public inquiry about the initiative. Expansion of scope, related requirements, and relevant changes to the PAPPG and award terms and conditions will be communicated using long-established annual update mechanisms that include a period for public comment.

Glossary

This glossary provides working definitions of terms used in this document that may be helpful. Unless otherwise noted, the following are from, or lightly adapted from, a relevant NASA policy.[55] These definitions are not meant to be determinative and are provided only as a courtesy for interested readers.

Accessible: Data that are retrievable by their identifier using standardized communications protocols, as per the definition in the FAIR principles.

Author Accepted Manuscript (AAM):[56] Author's final manuscript of a peer-reviewed paper accepted for journal publication, including all modifications from the peer-review process.

Scientific Data: Any electronically stored information, including:

- Observations, calibrations, coefficients, documentation, algorithms and any ancillary information.
- Information needed to validate the scientific conclusions of peer-reviewed publications. This includes data underlying figures, maps and tables.
- Material captured under the heading "research data" – see below.

Data Deposit Cost:[57] A one-time charge paid at the time a data set is deposited into a data repository.

Data Curation Costs:[58] Expenses associated with preparing data into a form that others can use.

Data Management:[59] The process of validating, organizing, protecting, maintaining and processing scientific data to ensure the accessibility, reliability and quality of the scientific data for its users.

Data Management and Sharing Plan (DMSP): Document that describes whether and how data will be shared and preserved. A DMSP should be compliant with NSF policies as set forth in the PAPPG, as well as 42 USC 19053. These plans have previously been called data management plans (DMP) by NSF.

Findable: Metadata and data should be easy to find for both humans and computers, as per the definition in the FAIR principles.

Interoperable: Data can work with other applications or workflows for analysis, storage and processing; often achieved by data and metadata formats and standards, as per the definition in the FAIR principles.

Metadata:[60] Data that provide additional information intended to make scientific data interpretable and reusable (e.g., date, independent sample and variable construction and description, methodology, data provenance, data transformations, and any intermediate or descriptive observational variables).

Open Science:[61] The principle and practice of making research products and processes available to all, while respecting diverse cultures, maintaining security and privacy, and fostering collaborations, reproducibility, and equity.

Persistent identifier, or PID: A long-lasting reference to a digital source. The digital object identifier, or DOI, system is an example of a persistent identifier.

Public access:[62] The 2013 OSTP memo characterized public access with the following statement: "To the extent feasible and consistent with law; agency mission; resource constraints; U.S. national, homeland, and economic security; and the objectives listed below, the results of unclassified research that are published in peer-reviewed publications directly arising from Federal funding should be stored for long-term preservation and publicly accessible to search, retrieve, and analyze in ways that maximize the impact and accountability of the Federal research investment."

Publication: Document released through print, electronic or alternative media, including peer-reviewed manuscripts, technical reports, conference materials and books. This does not include laboratory notebooks, preliminary analyses, drafts of scientific papers or preprints, plans for future research, peer review reports or communications with colleagues.

Report: Document produced through print, electronic or alternative media containing scientific and technical information. These documents are typically not peer reviewed. Examples includes technical publications, technical memorandums, contractor reports, conference publications, special publications and technical translation. It does not include annual or final annual performance reports.

Repository: An organized storage location that makes data findable and accessible.

Research data:[63] "[T]he recorded factual material commonly accepted in the scientific community as necessary to validate research findings," as defined in 2 CFR 200.315(e)(3), "Uniform Administrative Requirements, Cost Principles, and Audit Requirements for Federal Awards." While the focus is on unclassified digital data as defined in the regulation, NSF understands that additional DMP guidance at the directorate, division or program levels may include more nuanced criteria appropriate to the scientific discipline and proposed activity.

Reusable: Metadata and data should be well-described so that they can be replicated and/or combined in different settings, as per the definition in the FAIR principles. This includes releasing the data with a clear and accessible data usage policy.

Scientific utility: Information that is necessary to validate research findings or beneficial for future research activities.

Statistical purposes:[64] The description, estimation or analysis of the characteristics of groups without identifying the individuals or organizations that comprise such groups.

Version of record, or VOR:[65] Publisher's authoritative copy of a paper, including all modifications from the publishing peer review process, copyediting, stylistic edits and formatting changes.

Endnotes

[1] NSF is governed by the "National Science Foundation Act of 1950," as amended (Public Law 81-507, the "Organic Act" or the "NSF Act"), which is codified at 42 USC 1861 et seq.

[2] OSTP. 2013. *Increasing Access to the Results of Federally Funded Scientific Research*. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf.

[3] NSF. 2015. *NSF's Public Access Plan: Today's Data, Tomorrow's Discoveries. Increasing Access to the Results of Research Funded by the National Science Foundation*. https://www.nsf.gov/pubs/2015/nsf15052/nsf15052.pdf

[4] As of 2023, NSF-PAR version 2.5 is in operation.

[5] Deposits are made using the non-proprietary PDF/A format (ISO 19005-1:2005).

[6] For a comprehensive discussion of article versioning terminology, see NISO/ALPSP Journal Article Versions (JAV) Technical Working Group. "Journal Article Versions (JAV): Recommendations of the NISO/ALPSP JAV Technical Working Group." (2008), in which the AAM is described as AM (Accepted Manuscript). Report number RP-8-2008, available online: https://www.niso.org/publications/niso-rp-8-2008-jav. Select open science/public access terms are also presented in the glossary at the end of this public access plan as a convenience to readers.

[7] NSF. n.d. https://par.nsf.gov/.

[8] OSTP. 2022. *Ensuring Free, Immediate, and Equitable Access to Federally Funded Research*. https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf.

[9] Opening of the 2022 OSTP memo; see link ibid.

[10] NSF. 2022. *Leading the World in Discovery and Innovation, STEM Talent Development and the Delivery of Benefits from Research - NSF Strategic Plan for Fiscal Years (FY) 2022 – 2026*. https://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf22068. A list of strategic goals and objectives is found on p. 28.

[11] NSF. n.d. *Fact Sheet: Utilizing NSF-Funded Research in the Fight against COVID-19.* Accessed November 27, 2022. https://www.nsf.gov/news/special_reports/coronavirus/NSF-Funded_Research_COVID-19_Fact_Sheet.pdf

[12] For more information on this effort, see Wang, Lucy Lu, Kyle Lo, Yoganand Chandrasekhar, Russell Reas, Jiangjiang Yang, Doug Burdick, Darrin Eide, Kathryn Funk, Yannis Katsis, Rodney Michael Kinney, Yunyao Li, Ziyang Liu, William Merrill, Paul Mooney, Dewey A. Murdick, Devvret Rishi, Jerry Sheehan, Zhihong Shen, Brandon Stilson, et al. 2020. "CORD-19: The COVID-19 Open Research Dataset. " *In Proceedings of the 1st Workshop on NLP for COVID-19 at ACL 2020,* Online. Association for Computational Linguistics. See also allenai. 2022. "The COVID-19 Open Research Dataset (CORD-19). https://github.com/allenai/cord19.

[13] NSF. 2020. *Fact Sheet: Harnessing Computing Power to Fight COVID-19*. Accessed December 20, 2022. https://www.nsf.gov/news/special_reports/coronavirus/NSF_Computing_Power_COVID-19_4-10-2020.pdf; COVID-19 HPC Consortium. Nd. *Projects*. Accessed December 20, 2022. https://covid19-hpc-consortium.org/projects.

[14] Columbia University in the City of New York. "Covid Information Commons." https://covidinfocommons.datascience.columbia.edu/.

[15] RAPID (Rapid Response Research) awards differ from standard awards and fund research related to unanticipated events – including on topics relating to natural or anthropogenic disasters – where there is "urgency with regard to availability of or access to, data, facilities or specialized equipment;" see NSF. 2023. "Proposal & Award Policies & Procedures Guide (PAPPG) NSF 23-1." Chapter II.F.2. https://beta.nsf.gov/policies/pappg/23-1/ch-2-proposal-preparation#2F2

[16] Karen Stocks, and Evans, Helen. 2022. "Interviews with EarthCube Funded Projects: Perspectives on EarthCube's Impact and Sustainability Priorities." *In EarthCube Organization Materials*. UC San Diego Library Digital Collections. https://doi.org/10.6075/J0513ZC0. p. 4.

[17] Frank, John, Rosemary Foster, and Claudia Pagliari. 2023. Social Science & Medicine. 317. https://doi.org/10.1016/j.socscimed.2022.115592; Helen Santoro. 2021. "Addressing equity gaps in open science." *News Feature of the American Phycological Association*. 52(6), 22. https://www.apa.org/monitor/2021/09/news-equity-gaps.

[18] National Academies of Sciences, Engineering, and Medicine. 2018. *Open Science by Design: Realizing a Vision for 21st Century Research*. Washington, DC: The national Academies Press. https://doi.org.10.17226/25116. pp. 12-13.

[19] See, for example: Carroll, Stephanie Russo, Ibrahim Garba, Oscar L. Figueroa-Rodríguez, Jarita Holbrook, Raymond Lovett, Simeon Materechera, Mark Parsons et al. 2020. "The CARE principles for indigenous data governance." *Data Science Journal*, 19(1), 43. http://doi.org/10.5334/dsj-2020-043; Global Indigenous Data Alliance, n.d. https://www.gida-global.org/care; Gaddy, Marcus and Kassie Scott. 2020. "Principles for Advancing Equitable Data Practice." Wahington, DC: *Urban Institute.* https://www.urban.org/sites/default/files/publication/102346/principles-for-advancing-equitable-data-practice_0.pdf; and Nelson, Amy L. Hawn and Sharon Zanti. 2020. "A

Framework for Centering Racial Equity Throughout the Administrative Data Life Cycle." *International Journal of Population Data Science,* 5(3), Article 3. https://doi.org/10.23889/ijpds.v5i3.1367

[20] Including compliance with Section 508 of the Rehabilitation Act - 29 U.S.C. § 798.

[21] Yoon, Ayoung and Andrea Copeland. 2020. "Toward community-inclusive data ecosystems: Challenges and opportunities of open data for community-based organizations." *Journal of the Association for Information Science and Technology,* 71(12), 1439–1454. https://doi.org/10.1002/asi.24346.

[22] Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg et al. 2016. "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific Data*. 3(1), 1-9. https://doi.org/10.1038/sdata.2016.18; and Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg et al. 2019. "Addendum: The FAIR Guiding Principles for scientific data management and stewardship." *Scientific Data* 6(1), 1-2. https://doi.org/10.1038/s41597-019-0009-6.

[23] Such support is also in alignment with the statutory requirements of 42 USC 19053 (Title 3, Section 10344 of the "CHIPS and Science Act of 2022").

[24] For example, the recent award *Learning Data Science Through Civic Engagement With Open Data* (https://www.nsf.gov/awardsearch/showAward?AWD_ID=2005890)

[25] NSF's customer service policy requires 90 days advance notice to permit sufficient time for organizational changes to systems and policy development and associated training.

[26] NSF. n.d. Research Performance Progress Report (RPPR). https://www.nsf.gov/bfa/dias/policy/rppr/.

[27] Comeau, Donald C., Chih-Hsuan Wei, Rezarta Islamaj Doğan, and Zhiyong Lu. 2019. "PMC text mining subset in BioC: about three million full-text articles and growing," *Bioinformatics* 35(18): 3533–3535. doi: 10.1093/bioinformatics/btz070.

[28] NSF. 2022. "Dear Colleague Letter: Reproducibility and Replicability in Science. NSF 23-018." https://www.nsf.gov/pubs/2023/nsf23018/nsf23018.jsp?org=NSF.

[29] National Archives and Records Administration. 2022. *Code of Federal Regulations: A point in time eCRF system*. Accessed 20 December 2022. https://www.ecfr.gov/current/title-2/subtitle-A/chapter-II/part-200.

[30] Creative Commons. n.d. About CC Licenses. Accessed 20 December 2022. https://creativecommons.org/about/cclicenses/.

[31] Data – and in this usage, scientific data – is defined in Footnote 6 of the 2022 OSTP memo as "the recorded factual material commonly accepted in the scientific community as of sufficient quality to validate and replicate research findings. Such scientific data do not include laboratory notebooks, preliminary analyses, case report forms, drafts of scientific papers, plans for future research, peer-reviews, communications with colleagues, or physical objects and materials, such as laboratory specimens, artifacts, or field notes. The definition of 'scientific data' is similar to but broader than the term 'research data' defined by 2 CFR 200.315 (e) and 45 CFR 75.322 (e)."

[32] This concern was also noted in a letter from the House Science Committee to OSTP: Johnson, Eddie Bernice and Frank Lucas, 18 October 2022.

[33] NSF. 2023. "Proposal & Award Policies & Procedures Guide (PAPPG). NSF 23-1." Chapter II.D.i.ii https://beta.nsf.gov/policies/pappg/23-1/ch-2-proposal-preparation#2D2i-ii

[34] See, for example, NSF. 2020. "*NSF 20-068. Dear Colleague Letter: Open Science for Research Data*." *https://www.nsf.gov/pubs/2020/nsf20068/nsf20068.jsp*

[35] For example, useful peer-reviewed data descriptors are published in *Scientific Data*, an offering of the Nature publication group. Springer Nature. 2022 "Journal Information." https://www.nature.com/sdata/journal-information.

[36] NSF. 2019. "Dear Colleague Letter: Effective Practices for Data. NSF 19-069." https://www.nsf.gov/pubs/2019/nsf19069/nsf19069.jsp.

[37] White House Office of Science and Technology Policy, National Science and Technology Council Subcommittee on Open Science. 2022. *Desirable Characteristics of Data Repositories for Federally Funded Research*. https://doi.org/10.5479/10088/113528

[38] Such as via the awards related to the FAIR OS RCN Solicitation (NSF. n.d. Award search, element code 7414. https://www.nsf.gov/awardsearch/advancedSearchResult?ProgEleCode=7414&BooleanElement=Any&BooleanRef=Any&ActiveAwards=true#results); awards related to the EarthCube initiative (NSF. n.d. Award search, element code 8074. https://www.nsf.gov/awardsearch/advancedSearchResult?ProgEleCode=8074&BooleanElement=Any&BooleanRef=Any&ActiveAwards=true&#results); the University of Michigan's Research Data Ecosystem (RDE) for Social Science Research (NSF. 2022. "Mid-scale RI-2: The Research Data Ecosystem (RDE), a National Resource for Reproducible, Robust, and Transparent Social Science Research in the 21st Century." https://www.nsf.gov/awardsearch/showAward?AWD_ID=1946932&HistoricalAwards=false), and the Center for Open Science's STEM education resource hub (NSF. 2020. "AN EHR Core Research (ECT) Data Resource Hub to catalyze culture change and community building for

improving rigor and reproducibility in STEM education research." https://www.nsf.gov/awardsearch/showAward?AWD_ID=1937698&HistoricalAwards=false) and trial of registered reports (NSF. 2022. "NSF PD 19-125Y – Science of Science: Discovery, Communication, and Impact (SoS:DCI)-Randomized Control Trial of the Registered Reports Publishing Format") awards, among others, as well as the on-going efforts of the National Center for Science and Engineering Statistics (NCSES, https://ncses.nsf.gov/) and NSF-sponsored Major Facilities. (https://www.nsf.gov/bfa/lfo/docs/major-facilities-list.pdf).

[39] NCSES was established within NSF by Section 505 of the America COMPETES Reauthorization Act of 2010, 42 USC 1862p.

[40] See note 28.

[41] The Standard Application Process (SAP) portal is a product of the federal statistical agencies and units and managed by the Institute for Social Research, University of Michigan on behalf of the Federal government. 2022. "ResearchDataGov." https://www.researchdatagov.org/.

[42] Footnote 12 on page 5 of the 2022 OSTP memorandum.

[43] America's DataHub Consortium. 2022. "America's DataHub Consortium: A groundbreaking initiative to collaborate with the Federal Statistical System." https://www.americasdatahub.org/

[44] NSF. 2022. "Mid-scale RI-2: The Research Data Ecosystem (RDE), a National Resource for Reproducible, Robust, and Transparent Social Science Research in the 21st Century." https://www.nsf.gov/awardsearch/showAward?AWD_ID=1946932&HistoricalAwards=false

[45] Data management plans may address: (1) types of data, software, and materials to be produced; (2) data and metadata formats and standards; (3) data access/sharing policies and provisions for protection of sensitive data; (4) policies and rights for data use and re-use; (5) plans for archiving research products including data; (6) plans for providing access to physical material. Directorate-specific DMP guidance has been collected for easy access in NSF. n.d. "About the NSF Public Access Initiative." https://beta.nsf.gov/public-access. Individual solicitations may have additional specific requirements and/or guidance.

[46] Changes to NSF requirements will be reflective of requirements of 42 USC 19053 (also known as Title 3, Section 10344 of the "CHIPS and Science Act") as well as those expressed in the 2022 OSTP memo.

[47] NSF. 2023. "Proposal & Award Policies & Procedures Guide (PAPPG) NSF 23-1." Chapter II.D.2.f.vi.b https://beta.nsf.gov/policies/pappg/23-1/ch-2-proposal-preparation#2D2fvib. Emphasis (bold) is added here to indicate current policy relevant to the requirements of the 2022 OSTP memo.

[48] NSF. 2023. "Proposal & Award Policies & Procedures Guide (PAPPG) NSF 23-1." Chapter II.D.2.h.ii. https://beta.nsf.gov/policies/pappg/23-1/ch-2-proposal-preparation#2D2hii. Note that disclosures of foreign support are made to NSF as part of submitted proposals, and in annual performance reports, but are not part of public documents.

[49] See note 34.

[50] U.S. Government. 2021. *National Security Presidential Memorandum 33* ("NSPM-33"): *Presidential Memorandum on United States Government-Supported Research and Development National Security Policy*. https://trumpwhitehouse.archives.gov/presidential-actions/presidential-memorandum-united-states-government-supported-research-development-national-security-policy/.

[51] White House Office of Science and Technology Policy, National Science and Technology Council Subcommittee on Research Security. 2022. *Guidance for Implementing National Security Presidential Memorandum 33 (NSPM-33) on National Security Strategy for United States Government-Supported Research and Development*. https://www.whitehouse.gov/wp-content/uploads/2022/01/010422-NSPM-33-Implementation-Guidance.pdf.

[52] The *Science Experts Network Curriculum Vitae* (SciENcv) is maintained by NIH. See: National Library of Medicine, National Center for Biotechnology Information. n.d. "SciENcv Background." https://www.ncbi.nlm.nih.gov/sciencv/background/.

[53] NSF. n.d. "Awards Simple Search." https://www.nsf.gov/awardsearch/.

[54] NSF. n.d. "Public Access Initiative." https://beta.nsf.gov/public-access.

[55] NASA Science Mission Directorate. 2022. *NASA SCIENCE MISSION DIRECTORATE POLICY: Scientific Information policy for the Science Mission Directorate. SMD Policy Document SPD-41a*. https://science.nasa.gov/science-red/s3fs-public/atoms/files/SMD-information-policy-SPD-41a.pdf.

[56] NSF. 2015. *NSF's Public Access Plan: Today's Data, Tomorrow's Discoveries. Increasing Access to the Results of Research Funded by the National Science Foundation*. https://www.nsf.gov/pubs/2015/nsf15052/nsf15052.pdf, p 3.

[57] NSF. 2023. "Proposal & Award Policies & Procedures Guide (PAPPG) NSF 23-1" https://beta.nsf.gov/policies/pappg/23-1/ch-2-proposal-preparation#_ftn18. Definitions are found in note 18 to Part I, Chapter II (Proposal Preparation Instructions).

[58] Ibid.

[59] NIH. 2020. *Final NIH Policy for Data Management and Sharing Notice Number: NOT-OD-21-013*. https://grants.nih.gov/grants/guide/

notice-files/NOT-OD-21-013.html.

[60] Ibid.

[61] National Science and Technology Council (NSTC), Subcommittee on Open Science (SOS), Subgroup for the Year of Open Science (SYOS), 2023. https://open.science.gov/

[62] OSTP. 2013. *Increasing Access to the Results of Federally Funded Scientific Research*. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf, p. 3.

[63] See note 56.

[64] 44 USC 3561. United States Government. "Chapter 35-Coordination of Federal Information Policy Subchapter III-Confidential Information Protection and Statistical Efficiency Part A-General" 2023. https://usc-cdn.house.gov/view.xhtml?req=granuleid:USC-prelim-title44-section3561&num=0&edition=prelim.

[65] See note 56.

NSF