



Impact of the 2022 OSTP memo: A bibliometric analysis of US federally funded publications, 2017–2021

Eric Schares 

University Library, Iowa State University, Ames, IA

an open access  journal

Citation: Schares, E. (2023). Impact of the 2022 OSTP memo: A bibliometric analysis of US federally funded publications, 2017–2021. *Quantitative Science Studies*, 4(1), 1–21. https://doi.org/10.1162/qss_a_00237

DOI:
https://doi.org/10.1162/qss_a_00237

Peer Review:
https://www.webofscience.com/api/gateway/wos/peer-review/10.1162/qss_a_00237

Supporting Information:
https://doi.org/10.1162/qss_a_00237

Received: 26 October 2022
Accepted: 18 December 2022

Corresponding Author:
Eric Schares
eschares@iastate.edu

Handling Editor:
Ludo Waltman

Copyright: © 2023 Eric Schares.
Published under a Creative Commons
Attribution 4.0 International (CC BY 4.0)
license.



Keywords: Dimensions, embargo, federally funded research, Open Access, OSTP memo, public access

ABSTRACT

On August 25, 2022, the White House Office of Science and Technology Policy (OSTP) released a memo regarding public access to scientific research. Signed by Director Alondra Nelson, this updated guidance eliminated the 12-month embargo period on publications arising from U.S. federal funding that had been allowed from a previous 2013 OSTP memo. Although reactions to this updated federal guidance have been plentiful, to date there has not been a detailed analysis of the publications that would fall under this new framework. The OSTP released a companion report along with the memo, but it only provided a broad estimate of total numbers affected per year. Therefore, this study seeks to more deeply investigate the characteristics of U.S. federally funded research over a 5-year period from 2017–2021 to better understand the updated guidance’s impact. It uses a manually created custom filter in the Dimensions database to return only publications that arise from U.S. federal funding. Results show that an average of 265,000 articles were published each year that acknowledge US federal funding agencies, and these research outputs are further examined by publisher, journal title, institutions, and Open Access status. Interactive versions of the graphs are available at <https://ostp.lib.iastate.edu/>.

1. INTRODUCTION

On August 25, 2022, the White House Office of Science and Technology Policy (OSTP), under Director Alondra Nelson, released new policy guidance entitled “Ensuring Free, Immediate, and Equitable Access to Federally Funded Research” (Nelson, 2022). The memo states that, by 2026:

“all peer-reviewed scholarly publications authored or coauthored by individuals or institutions resulting from federally funded research **are made freely available and publicly accessible by default in agency-designated repositories without any embargo or delay after publication**” [emphasis preserved from original]

This new policy framework is an update to previous guidance on public access to scientific research. A 2013 policy released by Director John Holdren allowed a 12-month embargo on publications arising from federal funding, and only applied to federal agencies that granted over \$100 million annually (Holdren, 2013). By contrast, the new 2022 Nelson memo eliminates the possibility of a 12-month embargo period for federally

funded peer-reviewed research articles that was allowed under the previous 2013 guidance. It also extends guidance to the data underlying those publications, strengthens the data-sharing plans, contains specific metadata requirements, and applies to *all* U.S. federal granting agencies regardless of their annual granting amounts (Marcum & Donohue, 2022). The Association of Research Libraries has released a summary table outlining the details of the 2013 and 2022 OSTP memos for ease of comparison (Association of Research Libraries, 2022b).

Reactions to the Nelson memo have been plentiful and varied, with the release garnering national news coverage (Patel, 2022). Libraries (Association of Research Libraries, 2022a), universities (Association of American Universities, 2022), librarians (Moore, 2022; Anderson, 2022), societies (SPARC, 2022; European Science Foundation, 2022), consultants (Clarke & Esposito, 2022; Pollock & Michael, 2022), publishers (Association of American Publishers, 2022; PLOS, 2022; IOP Publishing, 2022), funders (Tananbaum, 2022), and researchers (American Anthropological Association, 2022) have weighed in with statements or opinion pieces, some more enthusiastic about the development than others.

Chairwoman of the House Committee on Science, Space, and Technology Eddie Bernice Johnson and Ranking Member Frank Lucas sent a joint letter to the newly confirmed Director of the OSTP, Dr Arati Prabhakar asking for clarifications (Johnson & Lucas, 2022). Questions still remain about the 2026 implementation and how specific practices will result from this new guidance, how agencies will update their policies, and concerns about participation in research if article processing charges (APCs) are increased or used more widely.

2. RESEARCH QUESTIONS

In addition to the 2022 memo, the OSTP also released a companion report on the potential economic impact of the updated guidance and its effects on federal grant funding agency policies (Office of Science and Technology Policy, 2022). The report estimates “between 195,000 and 263,000 articles were federally funded in 2020” but does not provide a more granular breakdown of these articles. An additional analysis estimates 197,000 federally funded articles in 2021 (Petrou, 2022). Other than these high-level studies, there have been limited analyses to more fully detail the characteristics of publications that fall within this newly expanded scope.

Therefore, this study seeks to address the following research questions:

- RQ1** How many U.S. federally funded publications have there been over the past 5 years? What are the yearly totals, and what proportion do these represent of worldwide and U.S.-specific output?
- RQ2** Which U.S. federal funding agencies awarded these grants?
- RQ3** How do the number of federally funded articles vary by research category/discipline?
- RQ4** Which publishers tend to publish federally funded articles?
- RQ5** Which journals tend to publish federally funded articles?
- RQ6** Which research institutions are authors who tend to publish federally funded articles affiliated with?
- RQ7** In what manner were these federally funded articles published? Were they published openly or behind a paywall?

3. METHODOLOGY

The analysis was conducted using the bibliometric database Dimensions, available at <https://app.dimensions.ai>. This study used the paid version of the tool; there is also a free version available, though with limited functionality. Dimensions ingests metadata from Crossref to make connections across publications, authors, funders, institutions, and more (Herzog, Hook, & Konkiel, 2020). The database uses this as a starting point and further enriches funding information by analyzing text provided in authors' acknowledgments sections and through agreements with publishers to obtain additional funding information.

Dimensions was chosen for this study because of relevant advantages over other commonly used bibliometric databases. It indexes a wider range of journals and has more complete coverage than Web of Science, which is estimated to cover only 10–12% of journals (Clarivate, 2021). OpenAlex, a free and open bibliographic database, also uses metadata reported to Crossref by publishers, as well as data from the now-discontinued Microsoft Academic Graph, scraping publisher websites, and other sources (Priem, Piwowar, & Orr, 2022). However, OpenAlex does not include funding information in its records of works.

This study is particularly affected by funding information that is deposited to Crossref and included in Dimensions. The availability of major metadata elements in Crossref was quantified by van Eck and Waltman (2021), who found 25% of articles in 2020 reported some funding information. Kramer and de Jonge (2022) specifically analyzed funder information in several bibliometric data sources and quantified the extraction of additional funding information from acknowledgment text, going beyond what is deposited by publishers to Crossref. Web of Science, Scopus, and Dimensions all infer this additional funding information. Dimensions reported funding information on 81% of the records in a study of publications by the Dutch Research Council, compared to 67% availability in Crossref. However, the information was inconsistent, with not all publications correctly naming the funder or providing the funder ID. The performance also varied considerably by publisher.

A case study deeply analyzing one example funding statement clearly illustrates the difficulties in untangling personal, financial, and logistical acknowledgments in the same section. Different bibliometric databases and tools are also shown to have different interpretations of the same funding information (Gibson, van Honk, & Calero-Medina, 2022).

3.1. Two Possible Approaches

The most crucial part of this analysis was defining the custom Funder group, which controls which publications are included/excluded in the analysis. The Dimensions web interface offers the ability to create a custom group of any single facet type—in this case, Funders. There were two main options to consider when deciding how to construct this custom group—define what to *include*, or define what to *remove*.

Attempting to include all federal grant-funding agencies in one custom Funder group was the first attempt. In theory, this sounds like the simpler approach; only keep those agencies whose funded output would qualify under the new OSTP guidance. Additionally, the OSTP economic impact report states that just six federal agencies “account for more than 94 percent of the approximately \$150 billion” in federal research and development (Office of Science and Technology Policy, 2022). This means the filter would be very nearly complete after including only six agencies.

However, it quickly became apparent that identifying and building one custom filter that covered all possible agencies, divisions, institutes, centers, and their name variants was not

feasible. For example, the U.S. Department of Health and Human Services (HHS) is a large federal granting agency. Within it are several Operating Divisions, such as the National Institute for Health (NIH) and the Food and Drug Administration (FDA). Within each of these Operating Divisions are further institutes and centers, such as the NIH's National Institute on Aging or the Office of AIDS Research. Dimensions enriches the publisher-supplied metadata from Crossref with additional information from a publication's acknowledgments section, but authors do not consistently identify funder names. Dimensions takes what it can find and does not further correlate or gather these variants into coherent groups. Depending on what is specifically acknowledged in a publication, the funding information returned may be as granular as a specific division, or as broad as an entire agency.

With this approach considered unmanageable, work then turned to the second option, which channeled the sculptor Michelangelo: Remove everything that is not a federal funder. However, this too had a fatal flaw. Although the funder information was more consistent and the granular nature of federal data was not a concern, removing private foundations, 501(c)(3)s, corporations, nonprofits, state agencies, or other organizations had the unintended effect of also removing desired publications. If an article acknowledged funding from both a foundation and a federal agency, the fact that the foundation was being removed from the analysis meant the entire paper would be excluded, even though it did have federal funding and should rightly be included in the data set.

Therefore, the final answer turned out to be a combination of the two approaches.

3.2. Defining Custom Funder Group

With the Dimensions web interface limited to the years 2017–2021 and Country of Funder set to United States, the Analytical View for Funders was able to quickly export the top 500 funders to meet those criteria (both federal and nonfederal).

Once the funder names were exported to an Excel sheet, some funders from countries other than the United States were still present due to publications with support from multiple grants and international collaborations. Limiting the exported column Country to United States reduced the number of funders from 500 to 331. It was then a manual task to search each funder one at a time, investigate its status, and determine if it was a federal agency or not. Those that were found to be foundations, 501(c)(3)s, corporations, nonprofits, state agencies, or other organizations were flagged. Each of these nonfederal funders was then added to a second, temporary custom Funder group in Dimensions to exclude them all in one large batch.

In the first round of investigation, 193 of the 331 organizations (58%) were determined to be nonfederal funders, and the process of exporting, filtering, and manual investigation was repeated two more times. This uncovered 87 and 36 more funding agencies to remove, respectively, for 316 nonfederal funders. The process was halted there, as the nonfederal funders were showing relatively low publication activity (<156 over 5 years, or approximately 30 publications per year) and there were diminishing returns when going further.

With the 316 nonfederal funders excluded, there were 1.129 million publications remaining. However, this is not the number to be concerned with, because by definition it will be an undercount. Some papers with federal funding get excluded by the fact that they *also* have nonfederal funding. Now that nonfederal agencies are cleaned out, the list of funders exported from the Analytical view has only federal granting agencies remaining, giving a cleaner picture of the dispersed and fractured naming conventions. These 177 federal agencies were then added to a custom filter one by one, with the number of newly qualifying papers recorded after each agency was added.

The last step was to look at the existing Funder groups in Dimensions that appear to all users. These predefined groups attempt to reconcile the fractured naming conventions for a few federal agencies, including CDC, DoD, DoE, NIH, NOAA, NSF, NASA, and USDA. These existing groups were expanded and the individual centers within were compared to the 175 federal agencies that made up the new custom group thus far. Sixty-two additional funders were found and added to the list, for 239 federal funding agencies. We can be reasonably confident in the completeness of this final custom group. By the end of the list, agencies are typically contributing single digits worth of unique publications to the whole, or a thousandth of a percent.

The full list of funders that make up the custom Funder group of U.S. federal granting agencies is included in the Supplemental material, Table S1.

4. RESULTS

The Dimensions web search interface was used for the analysis. The Dimensions API was considered and investigated briefly, but it would take a nontrivial amount of work to process text fields and strings of funder information. The web interface also offers a wide range of pre-built Analytical Views, which were very useful when conducting this analysis.

Years of publication were limited to 2017, 2018, 2019, 2020, or 2021. The Publication Type “Preprint” was excluded from the analysis, as the OSTP memo states only peer-reviewed publications qualify under its guidance. Preprints are versions of publications before peer review has been conducted; however, it is worth noting that several recent initiatives have begun to offer or organize peer reviews of preprints (Peer Community In, 2022; PREreview, 2022; Review Commons, 2022). All other document types, including Article, Proceeding, Chapter, Edited Book, and Monograph, were included as the memo states that it applies to “scholarly publications.” No restrictions or distinctions were made between a corresponding author on a publication and a researcher’s name appearing anywhere in the author list.

The full search run in Dimensions on October 18, 2022 is shown in Figure 1:

- Publication Year = 2017–2021
- Publication Type = NOT Preprint
- Country of Funder = United States
- Funder = Custom US federal funders group

With the data set defined, we can now move on to answering the specific research questions outlined in Section 2.

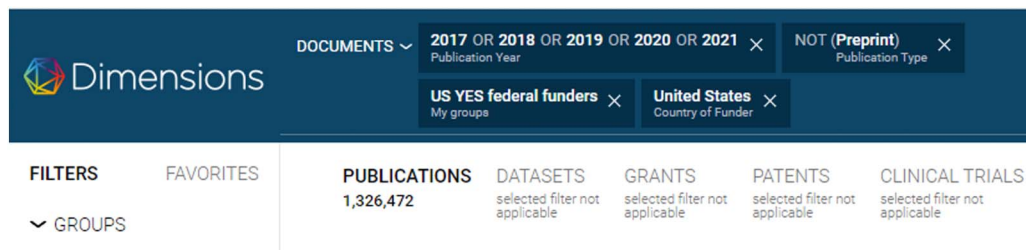


Figure 1. Dimensions search query.

4.1. RQ1: How Many Federally Funded (FF) Publications?

The search resulted in 1,326,472 publications. In general, approximately 250,000–277,000 publications per year are a result of U.S. federal funding (Table 1) and would thus be affected under the new OSTP guidance. The 5-year average is 265,294 publications. This falls on the high end of the range quoted by the OSTP economic impact memo of 195,000–263,000 annually (Office of Science and Technology Policy, 2022).

Comparing just the year 2021, the result is nearly 40% higher than the 197,000 U.S. federally funded articles estimated by an analysis using the database Web of Science (Petrou, 2022). As indicated above, Dimensions indexes more broadly and captures a wider range of journals and their articles.

The total of 1.3 million represents 33% of all U.S. domestic research output over these 5 years ($n = 4,020,840$), defined as any publication type, funded or nonfunded, with at least one author from a US institution (Location – Research Organization – Country/Territory = United States). This matches the 31% federally funded output found by Petrou (2022). It also represents 4.47% of total global research output over those 5 years ($n = 29,646,485$).

4.2. RQ2: Which Federal Granting Agencies?

Table 2 shows the top 10 U.S. federal granting agencies in terms of number of resulting publications over the 5-year period in this study. It is important to note that these funding agency names appear as they come from Dimensions, and are not synthesized or otherwise combined. This is most readily apparent when looking at the multiple NSF agencies that appear in the top 10: NSF MPS, NSF CISE, and NSF ENG. Beyond this top 10, a complete list of 319 grant making agencies, both US federal and nonfederal, and their number of related publications is provided with the data availability statement.

Exporting funder information is limited to the top 500 funders at a time in the Dimensions web interface. It would seem possible to export, then sum the number of publications from each funder to see how many of the total 1.326 million results are included in the top 500 (“Publications” column of Table 2). However, while the web interface deduplicates results and only returns the actual count of unique publications, the exported data *is not* deduplicated. Each agency is listed with the total number of papers that acknowledge funding from it. This means that a single paper that includes funding information from two or more granting agencies would be counted two or more times in the exported file, one for each named agency. As a result, adding up the counts from the Funder export file gives a sum of 2.58 million records.

Table 1. Number of US federally funded publications by year

Year	# Publications
2021	275,825
2020	277,407
2019	262,682
2018	259,518
2017	251,040
Total	1,326,472

Downloaded from http://direct.mit.edu/qss/article-pdf/4/1/1207/8449/qss_a_00237.pdf by guest on 07 September 2023

Table 2. Top 10 U.S. funding agencies by number of resulting publications

Name	# Publications
National Cancer Institute (NCI)	137,496
Directorate for Mathematical & Physical Sciences (NSF MPS)	118,881
National Institute of General Medical Sciences (NIGMS)	118,095
U.S. Department of Energy (DOE)	113,712
National Heart Lung and Blood Institute (NHLBI)	89,767
Directorate for Computer and Information Science and Engineering (NSF CISE)	80,068
Directorate for Engineering (NSF ENG)	77,784
National Center for Advancing Translational Sciences (NCATS)	77,170
National Institute of Allergy and Infectious Diseases (NIAID)	76,801

The large number of qualifying federal funders in the custom filter is one of the major changes in the 2022 OSTP memo, as the 2013 version applied only to those U.S. federal granting agencies that award more than \$100M annually. For example, the U.S. Department of Agriculture (USDA)'s Animal and Plant Health Inspection Service (APHIS) award total for FY 2020 was \$19 million (National Center for Science and Engineering Statistics (NCSES), 2022), so it was not covered under the 2013 Holdren memo. USDA APHIS was acknowledged on 1,693 publications over the 5-year period studied here, coming in at rank #149. In 2021, 107 of the 399 articles acknowledging funding from APHIS were published behind a paywall (26%). The updated Nelson memo guidance does away with the \$100 million annual funding threshold, so APHIS will start to be covered by the zero embargo policy when it goes into effect by 2026.

4.3. RQ3: Which Research Categories Are FF Publications in?

The number of publications affected by the updated OSTP memo will vary considerably depending on research field. Some disciplines are more publication intensive, and some rely more heavily on federal grant monies.

Figure 2 shows the 10 highest research categories in terms of number of publications, taken from a Dimensions Analytical view in the web interface. Dimensions uses the ANZSRC 2020 category classifications for discipline coding (Australian Bureau of Statistics, 2020). Looking at the top three categories, Biomedical and Clinical Sciences published nearly 400,000 total articles over these 5 years. Biological Sciences and Engineering were numbers 2 and 3 at around 200,000 publications each, and groups 4–6 each published around 120,000 total articles.

4.4. RQ4: Which Publishers Tend to Publish FF Articles?

Publishers vary in the subject matter they produce, so the impact of the new OSTP guidance will affect publishers to different degrees. Figure 3 shows the top 500 publishers in terms of number of U.S. federally funded outputs. It plots the number of FF publications from 2017 to 2021 v. the total number of publications over those years, broken out by publisher. The trendline

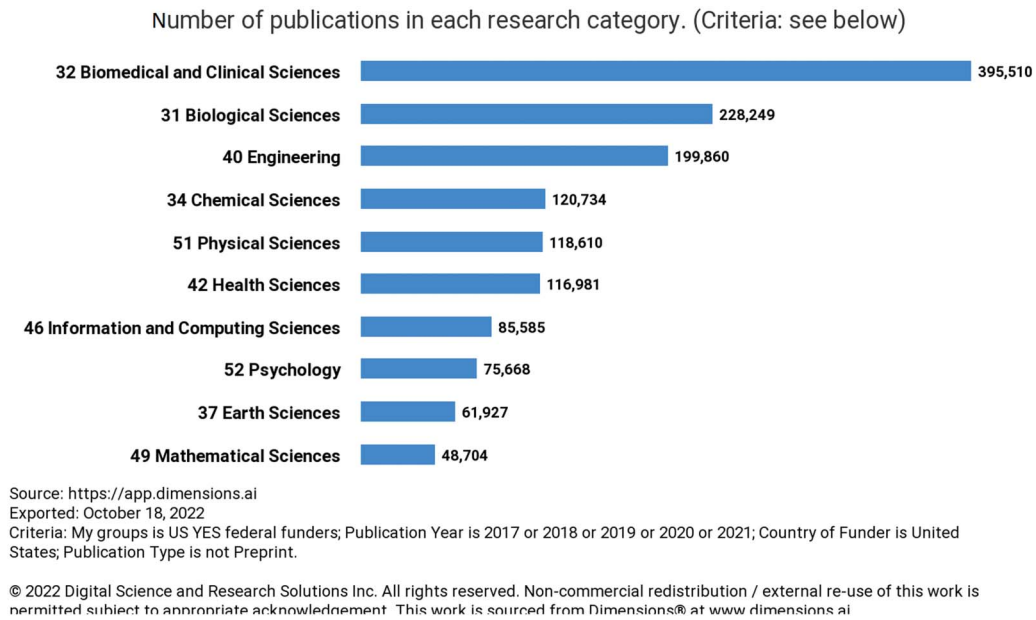


Figure 2. Research category classifications of publications, 2017–2021.

shows a strong relationship between the two types; as the total number of publications increases, so too does the number of federally funded articles tend to increase.

Some interesting data points begin to appear. Wiley, American Chemical Society (ACS), American Physical Society (APS), American Geophysical Union (AGU), the American

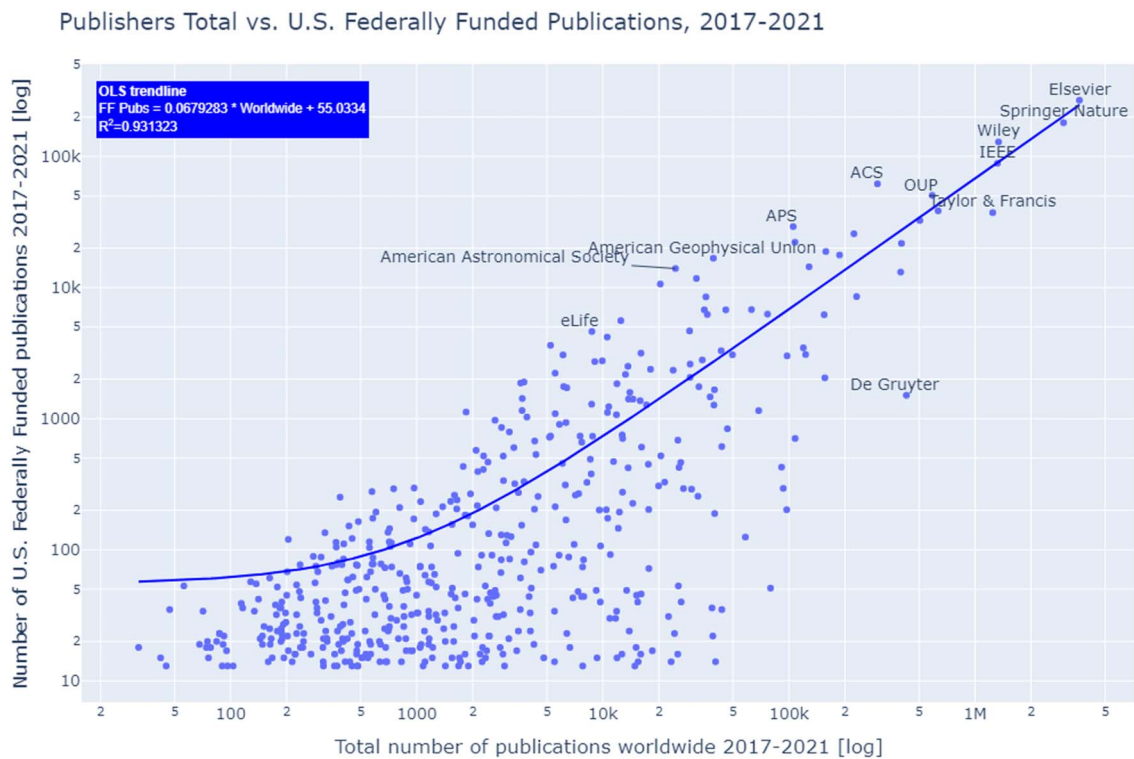


Figure 3. By publisher: US federally funded publications v. total publications. Interactive version at <https://ostp.lib.iastate.edu/>.

Astronomical Society, eLife, and others all appear as points above the trendline. This means they publish more federally funded research than might be expected from the overall trend.

Many data points are clustered in the lower left corner of Figure 3, causing data labels to overlap and preventing them all from being shown. To facilitate deeper exploration and allow the user to investigate data points of interest, interactive versions of each plot shown in this analysis are available online at <https://ostp.lib.iastate.edu/>.

It may be tempting to say that publishers above the trendline will be more affected by the new guidance, as more of their papers will fall under zero-embargo with immediate public access. However, Figure 3 is showing only the absolute numbers of publications. Further, the x- and y-axes show numbers on a log scale. Therefore, it can be difficult to determine the exact proportion of articles that are federally funded for each publisher, as the scales are not linearly increasing.

Figure 4, then, continues the investigation by showing the *percentage* of each publishers' total output that arise from U.S. federal funding. In general, most publishers cluster between 0 and 20%, with thinning numbers from 20–40% and 21 out of 500 publishers above 40% FF. The highest percentage of FF research is among publishers that have relatively low total output.

Here again we can see the American Astronomical Society and eLife having relatively high percentages of FF articles. However, additional publishers jump to the top, such as the International Ocean Discovery Program (95% of total), National Institute on Alcohol Abuse and Alcoholism (75%), and American Society for Clinical Investigation (69%). Table 3 shows the 12 highest publishers by federal funding percentage, though many publishers listed publish relatively few articles overall. The American Astronomical Society has the highest total output of those with over 50% federally funded, at 24,541 total articles in the 5 years studied here.

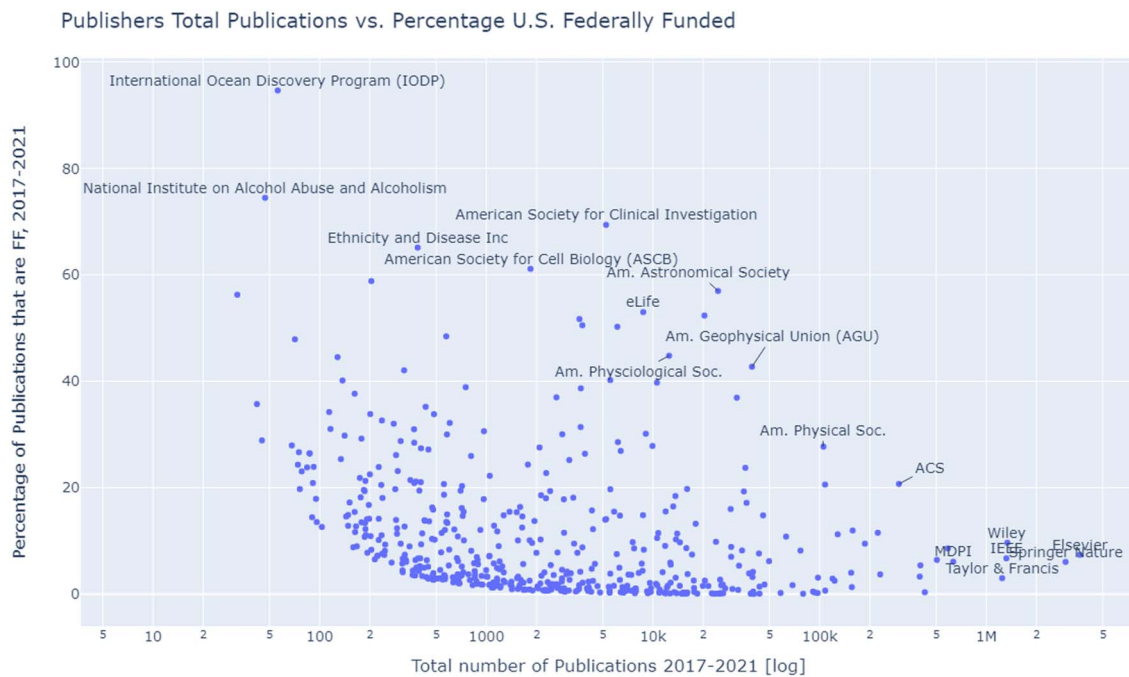


Figure 4. By publisher: percentage of federally funded publications v. total number. Interactive version at <https://ostp.lib.iastate.edu/>.

Table 3. Publishers with highest percentage of FF publications

Publisher	FF pubs	All pubs	% FF
International Ocean Discovery Program (IODP)	53	56	94.64
National Institute on Alcohol Abuse and Alcoholism	35	47	74.47
American Society for Clinical Investigation	3,627	5,228	69.38
Ethnicity and Disease Inc	252	387	65.12
American Society for Cell Biology (ASCB)	1,124	1,839	61.12
Tobacco Regulatory Science Group	120	204	58.82
American Astronomical Society	13,980	24,541	56.97
Academy of Natural Sciences of Philadelphia	18	32	56.25
eLife	4,625	8,728	52.99
Proceedings of the National Academy of Sciences	10,649	20,344	52.34
Rockefeller University Press	1,870	3,618	51.69
The American Association of Immunologists	1,903	3,767	50.52

The American Physical Society (APS) and American Chemical Society (ACS) move farther out on the *x*-axis with higher total output and lower on the *y*-axis when looking at relative percentages in Figure 4, to 27.7% and 20.7%, respectively. Wiley was also above the overall trendline in absolute numbers, but shows less than 10% of its total output as a result of federal funding.

4.5. RQ5: Which Journals Tend to Publish FF Research?

Next we look at the individual journal level, starting again with the top 500 titles in terms of the number of U.S. federally funded publications from 2017–2021, and plotting v. the total number from those years.

Scientific Reports and *PLOS ONE* both publish a very high number of federally funded research, but they also publish a very high number of articles in general (top right corner of Figure 5). *IEEE Access* also publishes a large number of total articles, but relatively few of them are the result of federal funding. *Nature Communications*, *PNAS*, and the *Astrophysical Journal* all published between 8,000 and 12,000 FF articles during the time period studied here, much higher than would be expected from the trendline. However, as with Section 4.4 on publishers, it is difficult to determine the percentage of publications with federal funding within each journal from this graph.

Figure 6 moves from absolute numbers to percentages and shows a similar trend to publishers, but with the curve more filled in. Instead of being very scattered and vertically dispersed with large gaps between data points at high percentage levels, the journal-level is more smoothly spread out as the curve moves to the top left. More titles are present in the >40% region of the chart, with 120 out of 500 data points falling in this range.

Nevertheless, the highest percentage of FF research at the journal level again occurs in journals that publish relatively few articles overall. Sixty-four journals have 50% or more of their

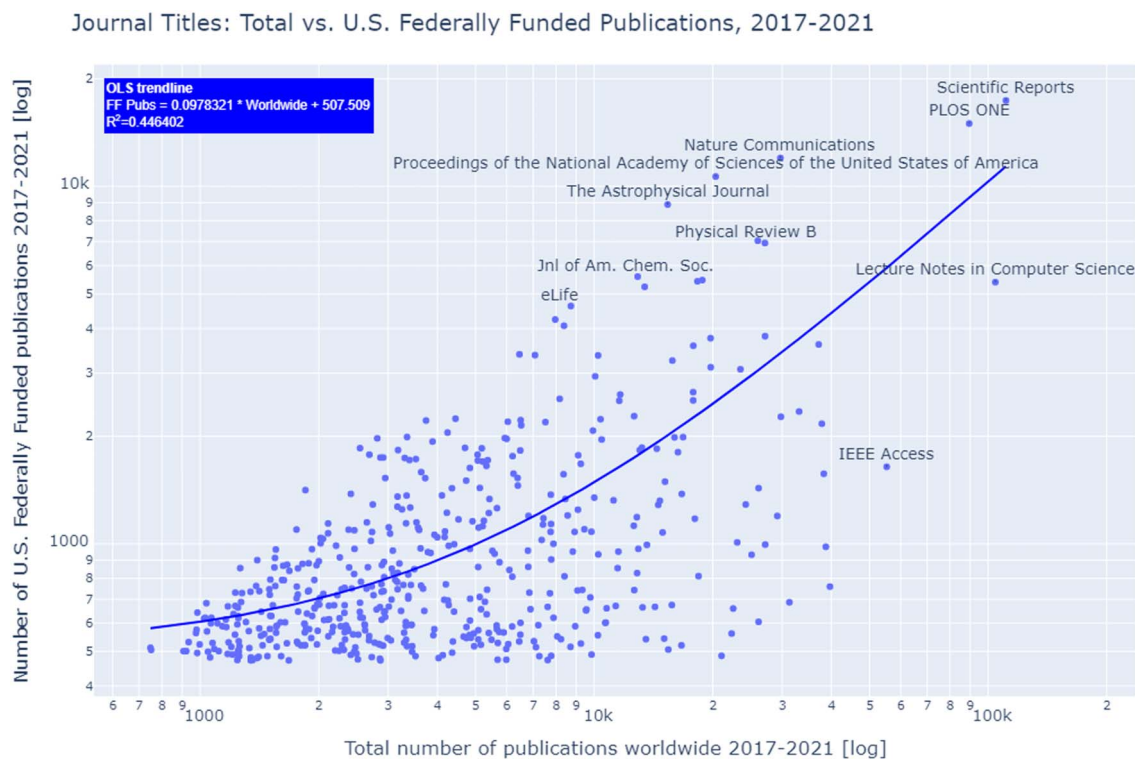


Figure 5. By journal: US federally funded publications v. total publications. Interactive version at <https://ostp.lib.iastate.edu/>.

total output from U.S. federally funded research. *AIDS and Behavior* has the highest FF percentage at 76%, followed by *JCI Insight* and the *Astronomical Journal*, with both over 70%. Table 4 lists the journals with the highest FF percentages and their corresponding numbers. When looking at percentages of FF research, the large total volume destinations *Scientific Reports* and *PLOS ONE* mentioned earlier both come in at around 16% each, with *IEEE Access* at slightly less than 3%.

Petrou (2022) looked deeply at the case of four journals in particular, some “of the most prestigious, mostly paywalled, scholarly journals”: *Nature*, *Science*, *Cell*, and *PNAS*. The finding was that more than 40% of these journals’ papers were from U.S. federally funded research. This analysis agrees in the cases of *Cell* and *PNAS* (60% and 52%), but differs for *Nature* and *Science*. In those cases, this study finds only 15% and 13% of papers, respectively, to be a result of federal funding over 5 years. One possible explanation is the extensive front matter and high level of editorial content in these journals, which is included in the Dimensions “Article” type, but is provided as a separate facet and thus able to be filtered out of results in Web of Science (see also Section 5). However, even if one were to take the Dimensions FF number as the numerator (casting the widest possible net to find FF articles) and the WoS number as the denominator (narrowing to a stricter definition of Article type), the percentages only increase modestly to 22% and 23%.

4.6. RQ6: Which Research Institutions Are Authors Who Tend to Publish FF Research Affiliated with?

Research institutions will also see effects to varying degrees from the updated OSTP memo. This is what will likely be of most interest to individual libraries and universities—how will our specific campus be affected by this new guidance?

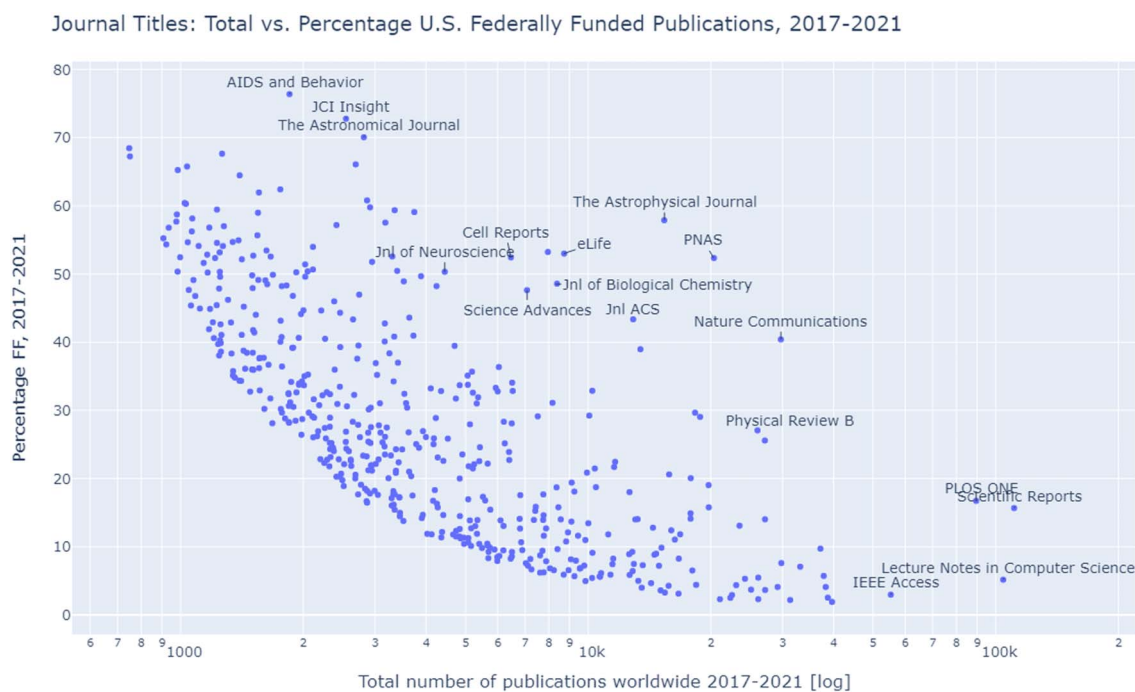


Figure 6. By journal: percentage of federally funded publications v. total number. Interactive version at <https://ostp.lib.iastate.edu/>.

The analysis in this section was filtered down to look only at research institutions in the United States (Dimensions filter: Location – Research Organization – Country/Territory = United States). Note that organizations from around the world will also be affected by the new OSTP memo, not only those in the United States. When publishing research collaboratively with a researcher that receives some federal funding from the United States, the resulting

Table 4. Journals with highest percentage of FF articles

Journal title	FF pubs	All pubs	% FF
<i>AIDS and Behavior</i>	1,413	1,850	76.38
<i>JCI Insight</i>	1,852	2,545	72.77
<i>The Astronomical Journal</i>	1,972	2,815	70.05
<i>Preventing Chronic Disease</i>	512	748	68.45
<i>Alcoholism Clinical & Experimental Research</i>	855	1,264	67.64
<i>Genes & Development</i>	505	751	67.24
<i>Journal of Clinical Investigation</i>	1,776	2,688	66.07
<i>Contemporary Clinical Trials</i>	682	1,037	65.77
<i>Journal of Substance Abuse Treatment</i>	642	984	65.24
<i>Molecular Biology of the Cell</i>	900	1,396	64.47
<i>American Journal of Preventive Medicine</i>	1,096	1,756	62.41
<i>Neuropsychopharmacology</i>	964	1,556	61.95

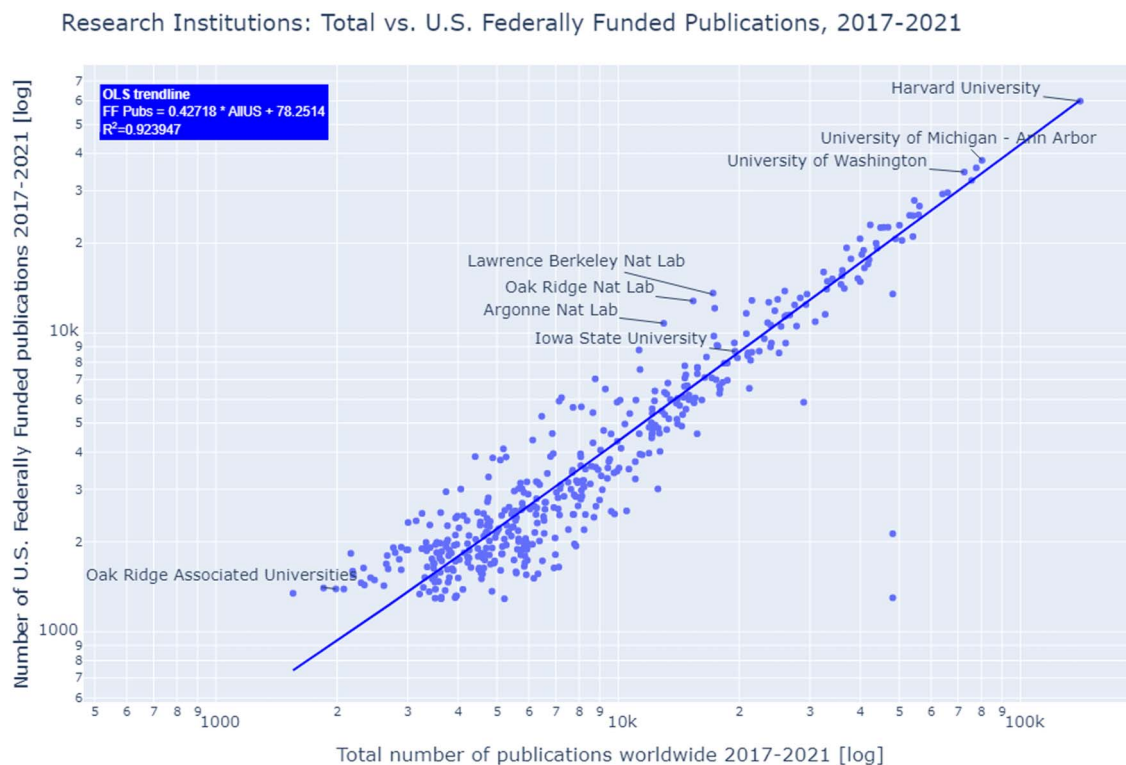


Figure 7. By research institution: US federally funded publications v. total publications. Interactive version at <https://ostp.lib.iastate.edu/>.

research output will still qualify under the updated guidance and be made publicly available immediately. Therefore, institutions from around the world do appear in this section, but with only a fraction of their total numbers represented.

Once again, the general trend holds: As the number of publications goes up, so too does the number of federally funded publications (Figure 7). Three institutions that stand out as publishing more funded research than typical are three national labs: Lawrence Berkeley, Oak Ridge, and Argonne. It makes sense that more of their research would be a result of federal funding as the labs themselves are the result of federal funding.

When looking by percentage, national laboratories again come to the top, as expected. There is such tight clustering at the top that data labels quickly become overlapped and unreadable; therefore a rectangular callout is added to Figure 8 to show the grouping of national laboratories at high FF percentages. The first nonfederal agencies to appear are Scripps Research at #21 and the Eli and Edythe L. Broad Institute of MIT and Harvard at #22, both with almost exactly 73% of their total output acknowledging federal funding. Harvard University does appear one spot before that at #20, but with their jointly administered Harvard–Smithsonian Center for Astrophysics.

Table 5 shows the top 12 institutions with high FF percentages.

4.7. RQ7: Were These FF Articles Published Open Access or Behind a Paywall?

So far, we have seen a detailed analysis of publisher, journal, and institution-level publication patterns, looking at how many U.S. federally funded articles were published over a certain time period out of a whole. However, a question still remains: In what manner were those federally

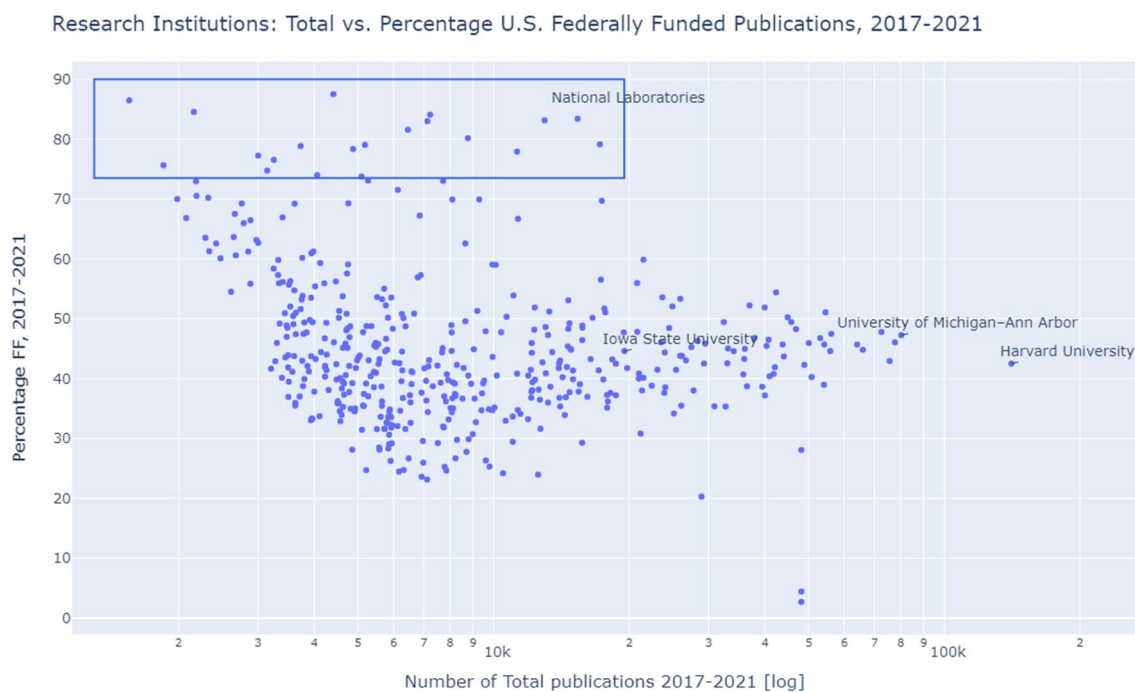


Figure 8. By research institution: percentage of federally funded publications v. total number. Interactive version at <https://ostp.lib.iastate.edu/>.

funded articles published? Are they published as some form of Open Access and thus already freely available? Or were they published behind a paywall, and if they had they been published in 2026 or later, would have represented a need to change the access mode? In other words, how many of these past publications would have required a shift in access to comply with the Nelson OSTP memo guidance?

Table 5. Research Institutions with highest percentage of FF articles

Institution	FF pubs	All US	%
SLAC National Accelerator Laboratory	3,861	4,411	87.53
National High Magnetic Field Laboratory	1,345	1,555	86.50
Frederick National Laboratory for Cancer Research	1,830	2,164	84.57
Lawrence Livermore National Laboratory	6,082	7,232	84.10
Oak Ridge National Laboratory	12,820	15,367	83.43
Argonne National Laboratory	10,792	12,978	83.16
Brookhaven National Laboratory	5,920	7,132	83.01
National Institute of Allergy and Infectious Diseases	5,268	6,457	81.59
Pacific Northwest National Laboratory	7,032	8,771	80.17
Lawrence Berkeley National Laboratory	13,623	17,215	79.13
National Renewable Energy Laboratory	4,099	5,185	79.05
National Center for Atmospheric Research	2,944	3,733	78.86

Open Access mode is provided to Dimensions by Unpaywall, an open database of free scholarly article metadata. Unpaywall determines the best OA location of a publication based on a cascading algorithm. It “prioritizes publisher-hosted content first (Hybrid or Gold), then prioritizes versions closer to the version of record (PublishedVersion over AcceptedVersion), then more authoritative repositories” (Unpaywall, 2020). Therefore, even though a publication may match multiple OA codes, each publication receives only one OA status in Dimensions. Dimensions also supplements Unpaywall’s data with a list of full OA journals for the case of Gold Open Access (Digital Science, 2021).

An important nuance to keep in mind when looking at Open Access status is that Unpaywall provides the current status of an article as it appears when a report is run (in this case, as of October 2022). There is a lack of historic OA data, so it is not possible to track the OA status at the time a publication appeared, or to find when an article qualified for a certain OA status. Therefore, it is possible an article appearing in this analysis as a certain OA mode only earned that status recently and would have returned a different result if the analysis were run earlier.

Figure 9(a) displays the breakdown of each year’s federally funded output, showing the percentage of FF articles published under each of five types of Open Access status. For comparison, Figure 9(b) shows the same plot, but for all publications worldwide from 2017–2021, roughly 5.6 million per year.

Research outputs that were published Closed access, or behind a paywall, appear at the bottom of each year’s stack in gray. These are the publications that would have been affected the most by the new OSTP memo. Approximately 26% of each year’s FF papers were published Closed access over 2017–2020, but 2021 saw an increase to nearly 32%, likely due to the 12-month embargo that is currently allowed by OSTP policy. Over time, the gray Closed access bar may return to a level more consistent with past years.

Green OA publications are self-archived by the author or a colleague by depositing the paper into a freely available university repository, disciplinary server, or a personal webpage at no charge. This is the only OA route not delivered by the publisher, and the document may

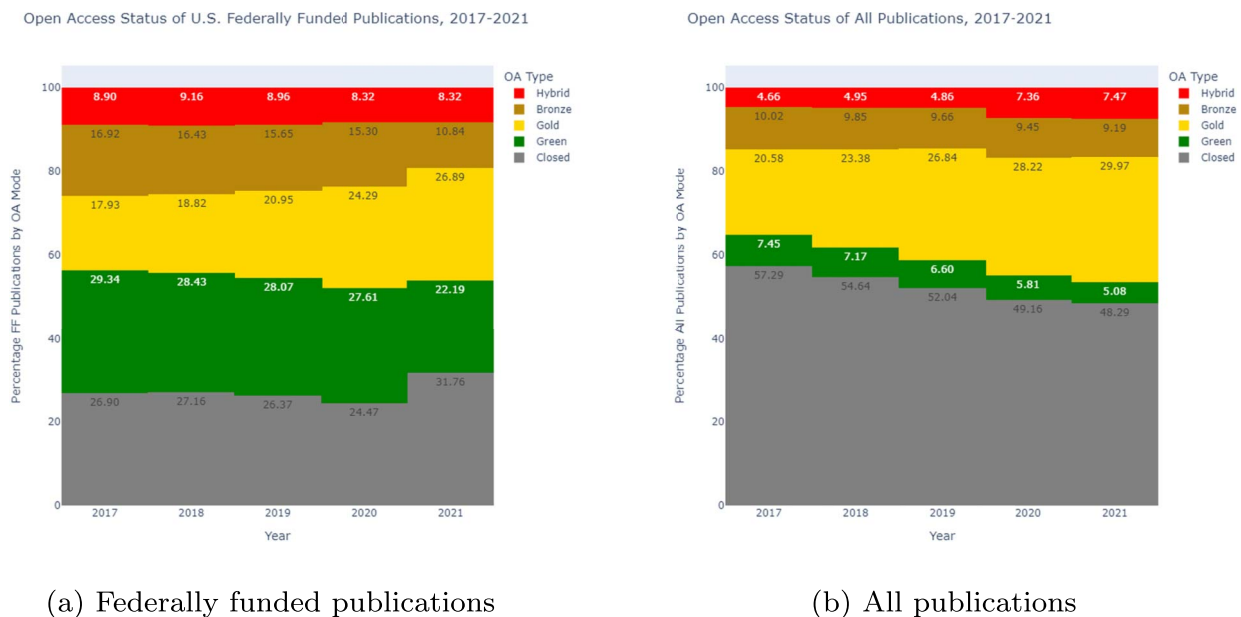


Figure 9. Open Access status of publications by year, 2017–2021.

not exactly match the final version, depending on publisher and journal restrictions. Green OA has seen its share decrease over these years, from being the most common mode in 2017 to the third most common by 2021. This may again be an artifact of the currently allowed 12-month embargo period in the 2013 OSTP memo. Authors may publish their work behind a paywall and make it publicly available through a Green OA route after 12 months. Interestingly, compliance with the new OSTP memo could increase participation in Green OA, so this mode of access may dramatically increase once the guidance takes effect by 2026.

Gold OA refers to the final version of an article published in a fully OA journal that offers all articles immediately, permanently, and freely available on the journal website. This may or may not be the result of paying an article processing charge (APC). Gold OA in federally funded publications has increased over the time period studied here, becoming the second most common mode of access in this data set, at around 27% in 2021.

Bronze OA is free access that is made temporarily available by the publisher, which can grant and remove access at any time without warning. It has seen the percentage of FF publications decrease over time.

Finally, Hybrid OA articles are published within a subscription (toll-based) journal, but made freely available on an individual, case-by-case basis by “unlocking” the article through paying an article processing charge (APC) to the publisher or journal. Hybrid contributes the smallest amount to the FF OA modes studied here, around 8%.

In terms of all publications, Figure 9(b) shows a dramatic decrease in the number of Closed publications, with Gold increasing and taking a larger share each year. Green remains relatively small at around 5%, while Bronze and Hybrid make up 7% and 9% yearly, respectively.

Papers that acknowledge U.S. federal funding are already much less likely to be published Closed access and much more likely to be deposited Green OA than a typical paper. The impact of the OSTP memo will likely accelerate this, as depositing Green is one way to achieve zero-embargo availability. Hybrid, Bronze, and Gold OA modes are roughly equivalent between US federally funded and nonfederally funded publications.

4.7.1. Open Access by publisher

Once OA status is introduced, we can combine some of the earlier aspects for further investigation. Figure 10 shows a stacked bar chart of the Open Access status of FF publications by publisher. This shows the publishers with the 16 largest amounts of FF publications in terms of absolute number.

Elsevier, Springer Nature, and Wiley are all large publishers with around 30% of FF research published as Closed access. IEEE and ACS have much higher percentages published as Closed, at around 50–60%. These publishers may be more vulnerable to the change in policy by making their previously paywalled content publicly and openly available. Pure Gold publishers appear strikingly as nearly 100% yellow, such as MDPI, Frontiers, and PLOS. Wolters Kluwer shows the highest amount of Green OA in this set of 16 publishers, at nearly 50%.

This data could be presented in many other ways. Figure 10 shows the top 16 publishers by number of FF publications, but it may also be interesting to sort by highest percentage of Closed access or most Green OA. A companion website is available at <https://ostp.lib.iastate.edu/>, which expands the data presented in Figures 10 and 11 to show the top 32 instead of only the top 16. In addition, the user is able to change the sorting method. Choices include highest total number of FF publications or highest percentage of any of the OA status (Closed, Green, Gold, Bronze, or Hybrid).

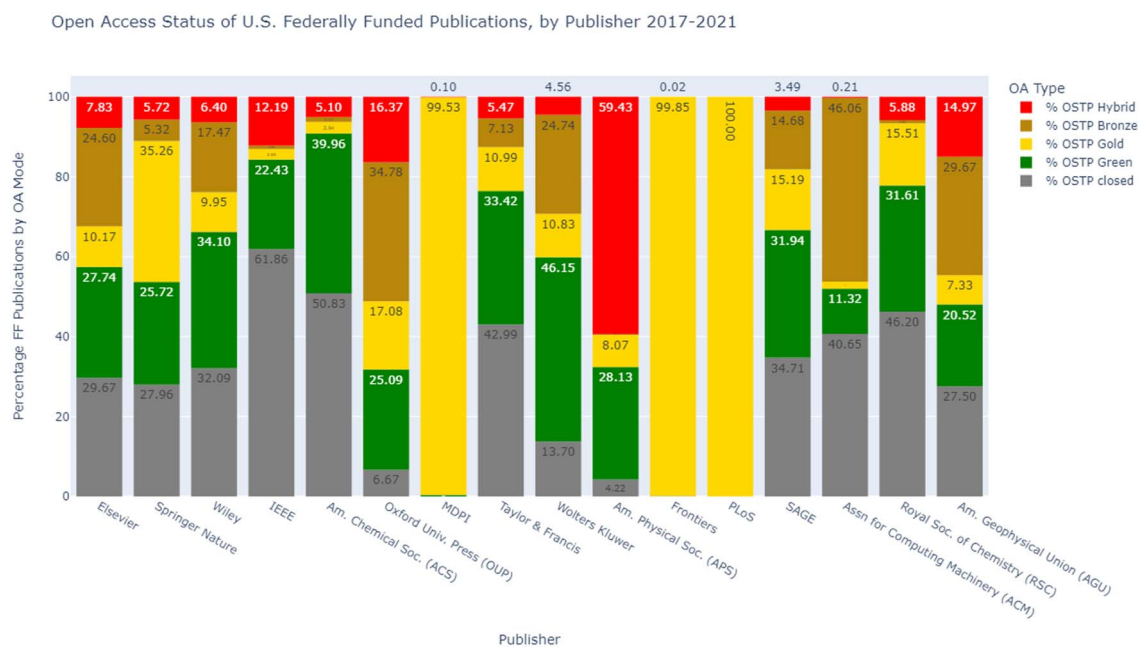


Figure 10. Open Access status of FF Publications by publisher, 2017–2021. Interactive version at <https://ostp.lib.iastate.edu/>.

4.7.2. Open Access by journal title

Moving to the OA status of individual journal titles, we can again see certain journals have the potential to be affected more heavily than others. Figure 11 shows the top 16 journal titles by absolute number of FF publications over the 5 years studied. The top three journal titles are all 100% Gold OA: *Scientific Reports*, *PLOS ONE*, and *Nature Communications*. *eLife* also

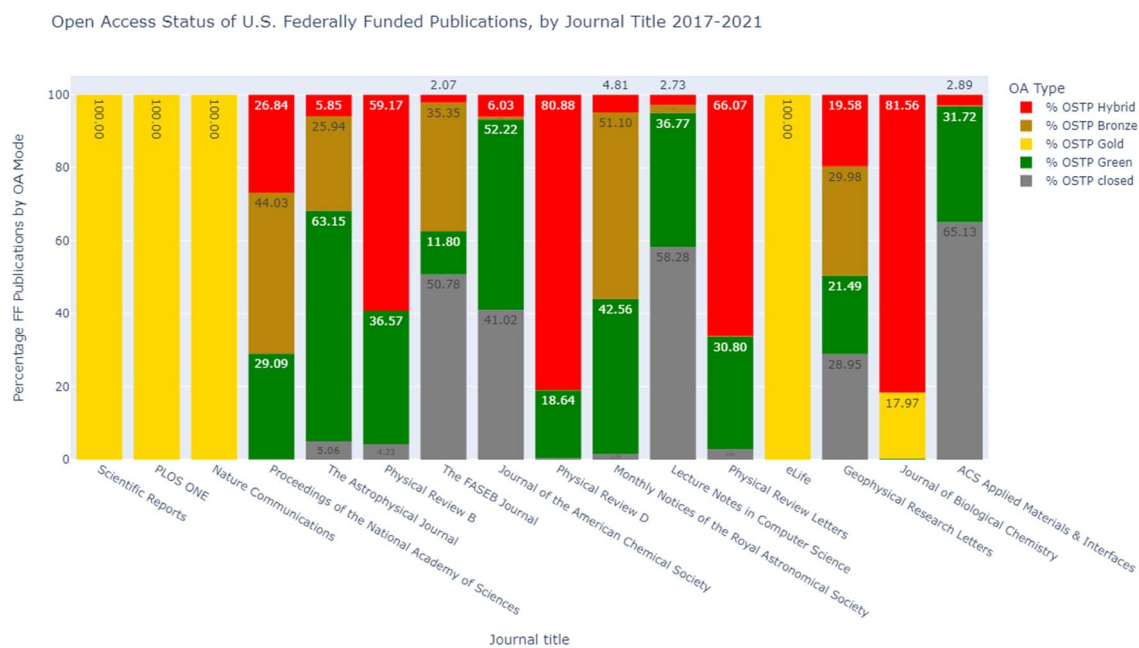


Figure 11. Open Access status of FF publications by journal title, 2017–2021. Interactive version at <https://ostp.lib.iastate.edu/>.

appears as completely Gold, at number 13. Presumably, these journals could continue operating as they are today even after the new OSTP policy framework takes effect in 2026. It is possible there is already some portion of FF output in these journals that is deposited Green OA. Unpaywall's algorithm prefers publisher-provided Open Access over repository provided OA, so the presence of Green OA within these journals would require a deeper analysis beyond what Unpaywall reports as the "best OA status."

Conversely, the *FASEB Journal*, *Journal of the ACS*, *Lecture Notes in Computer Science*, and *ACS Applied Materials & Interfaces* all publish a substantial proportion of their total federally funded output behind a paywall. These journals will need to adjust their policies and strategy to comply with the coming guidance of making FF publications immediately and publicly available.

Similar to publisher OA data, the companion website also allows a user to sort journal title OA data by highest total number of FF publications, or the highest percentage of any OA mode.

5. LIMITATIONS

The clearest limitation of this analysis is the likelihood that not all U.S. federally funded research is included in the data set. We are limited by the fact that we only know about publications that identify funding sources, and there is a possibility that some were missed. For example, the NSF's Polar Environment, Safety and Health Section (PESH) is a valid granting agency, but it returns no results when directly searched for in Dimensions. PESH is included in the list of agencies that make up the custom filter but it has no effect on the number of publications analyzed here.

As noted in Section 3, the availability of funding information from Crossref varies widely by publisher. This becomes the starting point for Dimensions, which then enriches the information with full text analysis and agreements with publishers to obtain additional funding information. Even when it is included, not all publications correctly name the funder or provide a funder ID. The new OSTP memo addresses this in section 4.a: "Agencies should ... collect and make publicly available appropriate metadata associated with scholarly publications and data resulting from federally funded research.... Such metadata should include at minimum: all author and co-author names, affiliations, and sources of funding, referencing digital persistent identifiers, as appropriate" (Nelson, 2022).

Authors will need to accurately and appropriately report their funding information, publishers need to supply that information to Crossref when registering for a DOI, and bibliographic databases must ingest that information. Enriching and enhancing funding information by analyzing the text of the acknowledgments section is also helpful, but could be improved.

It is also possible that some federal funding agencies were not explicitly included in the custom-made filter. Three rounds of refinement captured 239 individual funding agencies. One could always go further, but at some point, there are diminishing returns to continuing to export funders, manually assess them one by one, and add them to the custom Dimensions filter. A researcher with more in-depth knowledge of U.S. federal grant funding agencies and their subdivisions, institutes, and centers could investigate the custom filter that was defined and used here to identify holes or gaps.

The document type Article in Dimensions covers many types of content in journals, including editorials, letters, corrections, book reviews, news items, etc. (Digital Science, 2019). These materials were unable to be separated from the main Article type and were therefore included in this analysis.

6. CONCLUSION

The practical implications of the August 2022 OSTP memo's guidance are still being defined. Making federally funded publications immediately publicly available will involve a shift in strategy and behavior for publishers, authors, institutions, and readers. These peer-reviewed publications becoming immediately accessible to the public will expand the level of impact and reach, but it may also bring with it some ramifications that may not yet not be completely understood.

Though the OSTP released a companion impact report, it did not investigate the potential effects beyond a general estimation of the number of articles affected per year. This analysis went further but is only a first step in attempting to understand the broad-reaching implications of this updated policy. Quantifying the number, nature, and characteristics of publications from the past that would have qualified under this policy framework helps to clarify some questions and provide guidance still outstanding. It is clear that publishers, journals, and research institutions will all be affected, with some needing to adjust more than others. Once the new OSTP guidance takes effect, the equivalent of Figures 9(a), 10, and 11 will all become completely Green, Gold, or Hybrid.

Reported funder information is critical and will remain important as the OSTP guidance takes effect in 2026. Publishers, funders, and authors need to submit complete, accurate, and structured funding information, and database providers should continue to extract additional information to enhance this metadata. Dimensions and other bibliographic metadata tools will continue to define and refine funder filters to enable users to conduct a similar analysis to this on their own institution's publications.

ACKNOWLEDGMENTS

This paper was written using data obtained on October 18, 2022 from the paid version of Digital Science's Dimensions platform (Digital Science, 2018), available at <https://app.dimensions.ai>. Plots were created using Plotly version 5.10.0 (Plotly, 2022). The Dimensions Support Team was also very helpful in answering questions related to the construction of custom groups and providing further details on the data.

COMPETING INTERESTS

The author has no competing interests.

FUNDING INFORMATION

No funding was received for this research.

DATA AVAILABILITY

The data resulting from this research are made freely available in .csv format (Schaes, 2022a).

- Data set of funders
- Data set of publishers
- Data set of journal titles
- Data set of research organizations
- Data set of Open Access status by FF and worldwide

A companion website is also available at <https://ostp.lib.iastate.edu/> which includes interactive versions of each plot shown in this paper. Users may pan, zoom, and hover over data

points for more information. Additionally, they may search for a specific publisher, journal title, or research institution to enable its data label and color it red for easier identification on the graph. The source code for the website is freely available (Schaeres, 2022b).

REFERENCES

- American Anthropological Association. (2022). AAA's response to OSTP public accessibility memo. <https://www.americananthro.org/StayInformed/NewsDetail.aspx?ItemNumber=28478>. Accessed November 22, 2022.
- Anderson, R. (2022). A new OSTP memo: Some initial observations and questions. <https://scholarlykitchen.sspnet.org/2022/08/29/a-new-ostp-memo-some-initial-observations-and-questions/>. Accessed October 14, 2022.
- Association of American Publishers. (2022). Statement from Shelley Husband, Senior Vice President, Government Affairs, AAP, on decision by the White House Office of Science and Technology policy to make private sector publications freely available. <https://publishers.org/news/statement-from-shelley-husband-senior-vice-president-government-affairs-association-of-american-publishers-on-decision-by-white-house-office-of-science-and-technology-policy-to-make-private-se/>. Accessed October 14, 2022.
- Association of American Universities. (2022). AAU statement on OSTP decision to make federally funded research publicly available. <https://www.aau.edu/newsroom/press-releases/aa-u-statement-ostp-decision-make-federally-funded-research-publicly>. Accessed November 22, 2022.
- Association of Research Libraries. (2022a). ARL celebrates Biden-Harris administration's historic policy to make federally funded research immediately available. <https://www.arl.org/news/arl-celebrates-biden-harris-administrations-historic-policy-to-make-federally-funded-research-immediately-available/>. Accessed October 14, 2022.
- Association of Research Libraries. (2022b). Two-page table compares 2013 and 2022 public-access guidance from US Office of Science and Technology Policy. <https://www.arl.org/news/two-page-table-compares-2013-and-2022-public-access-guidance-from-us-office-of-science-and-technology-policy/>. Accessed December 9, 2022.
- Australian Bureau of Statistics. (2020). Australian and New Zealand Standard Research Classification (ANZSRC). <https://www.abs.gov.au/statistics/classifications/australian-and-new-zealand-standard-research-classification-anzsrc/latest-release>. Accessed October 14, 2022.
- Clarivate. (2021). Web of Science platform: Web of Science: Summary of coverage. <https://web.archive.org/web/20211214092845/https://clarivate.libguides.com/webofscienceplatform/coverage>. Accessed October 14, 2022.
- Clarke & Esposito. (2022). Zero embargo. <https://www.ce-strategy.com/the-brief/zero-embargo/>. Accessed October 14, 2022.
- Digital Science. (2018). Dimensions [software]. Available from <https://app.dimensions.ai>. Accessed October 18, 2022.
- Digital Science. (2019). What exactly is covered in the "Publications" in Dimensions? <https://dimensions.freshdesk.com/support/solutions/articles/23000018859-what-exactly-is-covered-in-the-publications-in-dimensions->. Accessed December 12, 2022.
- Digital Science. (2021). Where does the definition of "Open Access" come from in Dimensions? What does it include? <https://dimensions.freshdesk.com/support/solutions/articles/23000018863-where-does-the-definition-of-open-access-come-from-in-dimensions-what-does-it-include->. Accessed December 9, 2022.
- European Science Foundation. (2022). cOAlition S welcomes the updated Open Access policy guidance from the White House Office of Science Technology and Policy. <https://www.coalition-s.org/coalition-s-welcomes-the-updated-open-access-policy-guidance-from-the-white-house-office-of-science-technology-and-policy/>. Accessed October 14, 2022.
- Gibson, D., van Honk, J., & Calero-Medina, C. (2022). Acknowledging the difficulties: A case study of a funding text. <https://www.leidenmadrics.nl/articles/acknowledging-the-difficulties-a-case-study-of-a-funding-text>. Accessed December 7, 2022.
- Herzog, C., Hook, D., & Konkiel, S. (2020). Dimensions: Bringing down barriers between scientometricians and data. *Quantitative Science Studies*, 1(1), 387–395. https://doi.org/10.1162/qss_a_00020
- Holdren, J. (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. Accessed October 14, 2022.
- IOP Publishing. (2022). IOP Publishing's response to the OSTP memorandum on ensuring free, immediate, and equitable access to federally funded research. <https://iopublishing.org/news/iop-publishings-response-to-the-ostp-memorandum-on-ensuring-free-immediate-and-equitable-access-to-federally-funded-research/>. Accessed November 22, 2022.
- Johnson, E. B., & Lucas, F. (2022). Letter to OSTP Director Dr. Arati Prabhakar. House Committee on Science, Space, and Technology. https://science.house.gov/imo/media/doc/ostp_letter_research_access.pdf
- Kramer, B., & de Jonge, H. (2022). The availability and completeness of open funder metadata: Case study for publications funded by the Dutch Research Council. *Quantitative Science Studies*, 3(3), 583–599. https://doi.org/10.1162/qss_a_00210
- Marcum, C. S., & Donohue, R. (2022). Breakthroughs for all: Delivering equitable access to America's research. <https://www.whitehouse.gov/ostp/news-updates/2022/08/25/breakthroughs-for-all-delivering-equitable-access-to-americas-research/>. Accessed October 7, 2022.
- Moore, S. A. (2022). Thoughts on the new White House OSTP open access memo. <https://unlockingresearch-blog.lib.cam.ac.uk/?p=3271>. Accessed October 14, 2022.
- National Center for Science and Engineering Statistics (NCSES). (2022). *Federal Funds for Research and Development: Fiscal Years 2020–21*. Number NSF 22-323, Table 15. National Science Foundation, Alexandria, VA. Available at <https://ncses.nsf.gov/pubs/nsf22323/>.
- Nelson, A. (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>. Accessed October 7, 2022.
- Office of Science and Technology Policy. (2022). Economic landscape of federal public access policy. Technical report, White House Office of Science and Technology Policy. <https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Congressional-Report.pdf>.
- Patel, V. (2022). White House pushes journals to drop paywalls on publicly funded research. *New York Times*, August 25. <https://www.nytimes.com/2022/08/25/us/white-house-federally-funded-research-access.html>

- Peer Community In. (2022). Free peer review & validation of preprints of articles. <https://peercommunityin.org/>. Accessed November 28, 2022.
- Petrou, C. (2022). Guest post – quantifying the impact of the OSTP policy. <https://scholarlykitchen.sspnet.org/2022/09/13/guest-post-quantifying-the-impact-of-the-ostp-policy/>. Accessed December 9, 2022.
- PLOS. (2022). PLOS cheers the OSTP memorandum “Ensuring free, immediate, and equitable access to federally funded research”. <https://theplosblog.plos.org/2022/09/plos-cheers-the-ostp-memorandum/>. Accessed November 22, 2022.
- Plotly. (2022). Plotly open source graphing library for Python [software]. Available from <https://plotly.com/python/>. Accessed October 25, 2022.
- Pollock, D., & Michael, A. (2022). News & views: OSTP memo – Modeling market impact. <https://deltathink.com/news-views-ostp-memo-modeling-market-impact/>. Accessed October 25, 2022.
- PREreview. (2022). A site for crowdsourcing preprint reviews. <https://prereview.org/>. Accessed November 28, 2022.
- Priem, J., Piwowar, H., & Orr, R. (2022). OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. <https://arxiv.org/abs/2205.01833>
- Review Commons. (2022). The preprint peer-review platform. <https://www.reviewcommons.org/>. Accessed November 28, 2022.
- Schares, E. (2022a). OSTP impact. <https://doi.org/10.5281/zenodo.7254815>. Accessed December 9, 2022.
- Schares, E. (2022b). OSTP impact. GitHub repository. https://github.com/eschares/OSTP_impact
- SPARC. (2022). Taxpayers to get immediate access to publicly funded research. <https://sparcopen.org/news/2022/taxpayers-to-get-immediate-access-to-publicly-funded-research/>. Accessed October 14, 2022.
- Tananbaum, G. (2022). Open research funders group applauds bold OSTP action. <https://www.orfg.org/news/2022/8/29/open-research-funders-group-applauds-bold-ostp-action/>. Accessed December 9, 2022.
- Unpaywall. (2020). Data format, DOI object. <https://unpaywall.org/data-format#doi-object>. Accessed October 19, 2022.
- van Eck, N. J., & Waltman, L. (2021). Crossref as a source of open bibliographic metadata. In *Proceedings of the 18th International Conference of the International Society for Scientometrics and Informetrics (ISSI)* (pp. 1169–1174). https://www.issi-society.org/proceedings/issi_2021/Proceedings%20ISSI%202021.pdf#page=1201
- White House Press Release. (2022). OSTP issues guidance to make federally funded research freely available without delay. <https://www.whitehouse.gov/ostp/news-updates/2022/08/25/ostp-issues-guidance-to-make-federally-funded-research-freely-available-without-delay/>. Accessed October 24, 2022.