

# Public Access to Federally Funded Research

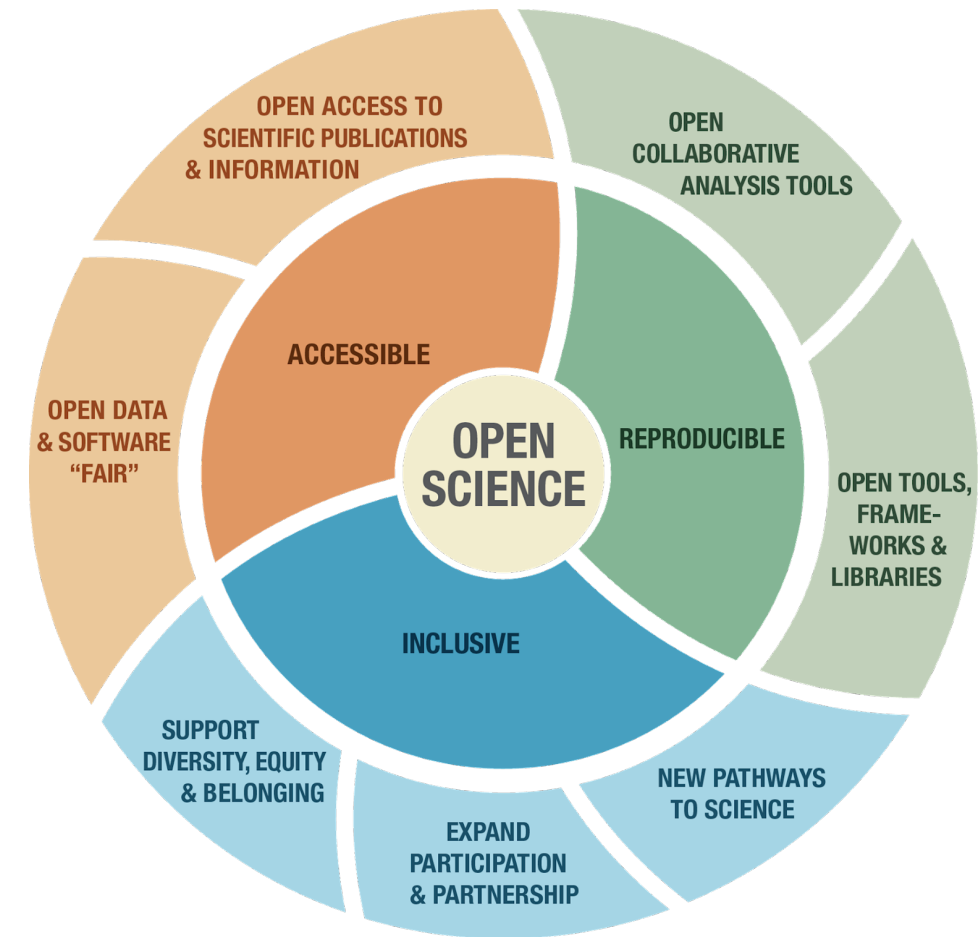
Louis Barbier, OCS

July 17, 2023



# NASA's Perspective on the 2022 OSTP Public Access Memo

- NASA strongly supports the intent of the 2022 OSTP Public Access Memo, *Ensuring Free, Immediate, and Equitable Access to Federally Funded Research*
- Promoting the full and open sharing of scientific research results with research communities, private industry, academia, and the general public is one of NASA's long-standing core values.



# NASA's Approach: Open Science Core Values

- Openness is fundamental, security is essential, and freedom and integrity are crucial
- Increase the accessibility, inclusion, and reproducibility of NASA scientific activities
- When possible, minimize the burden both within NASA and for our research communities
- NASA's Office of the Chief Scientist (OCS) is leading an Agency-wide update to the 2014 NASA plan
- The draft updated Public Access Plan was submitted to OSTP for review on Feb. 21, 2023





# A SHIFT IN THE WAY SCIENCE IS DONE



## OPEN (**TRANSPARENT**) SCIENCE

scientific process and results should be visible, accessible, and understandable



## OPEN (**INCLUSIVE**) SCIENCE

process and participants should welcome participation by and collaboration with diverse people and organizations

## OPEN (**ACCESSIBLE**) SCIENCE

data, tools, software, documentation, and publications should be accessible to all (FAIR)



## OPEN (**REPRODUCIBLE**) SCIENCE

scientific process and results should be open such that they are reproducible by members of the community



# Summary

## Agency-wide group efforts to date:

- Reviewed NASA Open Access Policy – **NPD 2230.1 Research Data and Publication Access** - no changes were needed to be compliant with the new OSTP memo of 2022
- Revised NASA's 2014 Open Access Plan for consistency with the OSTP 2022 memo
- Submitted revision on Feb. 21, 2023 to OSTP
- **Significant changes to be expected:**
  - Removing requirement for a one-year embargo on publications
  - Software (s/w) guidance for the process for releasing and maintaining s/w (consistent with multiple existing NASA policy documents covering s/w)
  - Integrating persistent identifiers for data and software with publications
  - Encouraging sharing of data that goes beyond a publication

**The following slides represent a summary of key principles and should not be relied upon without reference to the underlying policies**



# Key Points



## OSTP Open Access Memo Reference

These links are to the White House [article](#) about the [memo](#) signed August 2022 to ensure that federally funded research is freely available without delay. This memo is the next step after the [2013 memo](#) which charged federal agencies with developing and implementing a plan to increase public access to federally funded research.



## Open Access Publication

The memo requires that any federally funded research must be published in journals that allow for open access to its articles immediately upon publication.



## Open Access to Data Used in Publications

The memo also says that these publications must be made without an embargo on the data. This data is specifically that which was used to make the publication.

# Impacts to NASA funded Scientists



## Publishing Impact

NASA Science Programs will need to provide its scientists with additional budget to cover open access fees. Need to be certain we do not unwittingly preclude publication of NASA science in top journals due to the open access policies.



## Non-Published Data

The memo also calls on agencies to develop an approach on how to share non-published data.



## Future Contracts, Grants and Agreements

We need to be sure that future funding calls make clear that federally funded research must follow federal laws and OMB policies, including the open access publication and associated data rules. This is regardless of the rules of the publisher. This must be part of the budgets submitted in future proposals.



# What is subject to the Public Access Policy?

Applies to all research generating scientific data, including but not limited to:

**Research Projects funded through grants**

Small Business SBIR/STTR

Internal research at Centers

Does not apply to research projects not generating scientific data or non-research projects, including but not limited to:

Training

**Fellowships**

Conferences

Research-Related Infrastructure Programs



# Scientific Publications

**Peer-reviewed publications shall be made freely available at the time of publication. This supersedes NASA's previous guidance which allowed for a one-year embargo on publications**

**For authors who wish to publish in an open access journal, NASA allows all Article Processing Charges (APCs) to be included in the grant proposal budget**

This plan includes all peer-reviewed scientific research publications authored or co-authored by researchers receiving NASA-appropriated funds.

- This includes both civil servant and non-civil servant researchers, both intramural and extramural researchers
- Submission of a publication and metadata to a designated repository will be the responsibility of the lead author, or NASA-funded co-author if not the lead author
  - Metadata should be coded so that it can be crawled by automated search engines in order to facilitate discovery



# Sharing Publications

For articles that are published as Open Access the final published article (*i.e.*, the publisher's version of record) may be made publicly available in the STI Repository through one of the following mechanisms:

- For articles published as Open Access by journal publishers participating in the [Clearinghouse for the Open Research of the United States](#) (CHORUS), the published article will be made publicly available in the STI Repository on behalf of the authors. Authors should verify that their article is available in the STI Repository following its publication, in which case no further action is required by the author.
- For articles published as Open Access that are indexed in the [NASA Astrophysics Data System](#) (ADS), no further action is required by the researcher to comply with public access requirements for the article at this time.
- For articles published as Open Access that are not covered by CHORUS or ADS, authors must submit either the final published article or the author's copy of an accepted manuscript to the STI Repository via the [PubSpace submission page](#) no later than the article's publication date. Use this link to submit your publication as an attachment: <https://sti.nasa.gov/sti-contact-form/?RequestType=PublicAccess>
- List of CHORUS publisher members is here: <https://www.chorusaccess.org/about/our-members/>





# Open Access Models

- **Green** – Author self-archiving an early version of the article online. Once the peer-review process has been completed by a journal, the author posts an accepted manuscript to their own website, an institutional repository, a funder repository, or central open access repository (i.e., [arXiv.org](https://arxiv.org)).
- **Gold** - Publishers that make all their articles and related content (supplemental information and data) available for free on their website. The journals are funded by article processing fees (APCs) which are paid for by the authors or research sponsor (fees can range up to \$5,000 per article).
- **Hybrid** - Journals partially funded by subscriptions and only provide open access for the specific articles where the author paid a publication fee (APC). This option usually costs more than Gold OA.

# Scientific Data – definition\*

*“Research data are defined as the recorded factual material commonly accepted in the scientific community as necessary to validate research findings, but not any of the following: preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues. This 'recorded' material excludes physical objects (e.g., laboratory samples).”*

*Research data does not include:*

*(A) Trade secrets, commercial information, materials necessary to be held confidential by a researcher until they are published, or similar information which is protected under law; and*

*(B) Personal and medical information and similar information the disclosure of which would constitute a clearly unwarranted invasion of personal privacy, such as information that could be used to identify a particular person in a research study.”*



# Data and Software Management Plans

A Data Management Plan (DMP) shall provide a plan for making research data that underlie the results and findings in peer-reviewed publications digitally accessible at the time of publication

- including the data as supplementary information to the published article
- through NASA archives
- other means (specify)
  - The published article should indicate how these data can be accessed.

A Software Management Plan (SMP) that describes whether and how software generated through the course of the proposed research will be shared and preserved

- must describe how software sharing and preservation will enable validation of published results, or
- how such results could be validated if software are not shared or preserved

# Limitations on Data Sharing

Data Management and Sharing Plans should maximize appropriate sharing:

**Justifiable ethical, legal, and technical factors for limiting sharing of data include:**

- Informed consent will not permit or limits scope of sharing or use
- Privacy or safety of research participants would be compromised and available protections insufficient
- Explicit federal, state, local, or Tribal law, regulation, or policy prohibits disclosure
- Restrictions imposed by existing or anticipated agreements with other parties
- Datasets cannot practically be digitized with reasonable efforts

**Reasons not generally justifiable to limit sharing include:**

- Data are considered too small
- Researchers anticipate data will not be widely used
- Data are not thought to have a suitable repository

**Additional considerations:**

- NASA respects Tribal sovereignty and supports responsible management/sharing of AI/AN participant data
- SBIR/STTR Program Policy Directive permits withholding data for 20 years, as stipulated in agreements and consistent with program goals

# Supplemental Information:

## Elements of a Data Management Plan

- **Data type**
  - Identifying data to be preserved and shared
- **Related tools, software, code**
  - Tools and software needed to access and manipulate data
- **Standards**
  - Standards to be applied to scientific data and metadata
- **Data preservation, access, timelines**
  - Repository to be used, persistent unique identifier, and when/ how long data will be available
- **Access, distribution, reuse considerations**
  - Description of factors for data access, distribution, or reuse
- **Oversight of data management**
  - Plan compliance will be monitored/ managed and by whom

# Monitoring Compliance with DMPs

- **Approved DMP becomes a Term and Condition of Award**
- Annual progress reports
  - Include status and changes
- NASA reviews compliance annually
  - Failure to comply may result in an enforcement action, including additional special terms and conditions or termination of award, and may affect future funding decisions
- *Plans may be made publicly available in the future*



# When to share data?

- No later than publication of research results
  - For data published in peer-reviewed journals
- No later than the end of the award period
  - For data not published in peer-reviewed journals

# Open Data Portal

NASA's data portal:  
<https://data.nasa.gov/>

<https://www.nasa.gov/open/data.html>

## Want NASA data?

Check out our data catalog below or you can read about other open-government websites further down the page.

[Go to the DATA CATALOG](#)



# Allowable Costs for data management

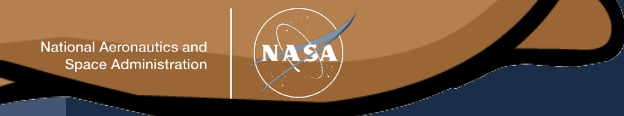
**Reasonable costs allowed in budget requests (must be incurred during the performance period)**

- Curating data/developing supporting documentation
- Preserving/sharing data through repositories
- Local data management considerations

**NOT considered data sharing costs**

- Infrastructure costs typically included in indirect costs
- Costs associated with the routine conduct of research (e.g., costs of gaining access to research data)

Over time NASA hopes to learn more about what constitutes reasonable costs for various data management and sharing activities



# Scientific Software

This plan applies to unrestricted, scientific software that is developed to support the scientific processing and analysis of data produced as part of a mission, unless restricted due to Agency or Federal policy.

Scientific software includes computer programs in source and object code that provide users some degree of scientific utility or produce a scientific result or service. This includes the software that is necessary to validate research findings, associated with scholarly publications. It may also include not only the source code, but also metadata describing the software, descriptions of how to use the software, and associated software documentation.

*Scientific Software does not include:*

- Software developed only for preliminary analysis, plans for future research, or communication with colleagues; or commercial (COTS) software*
- Software restricted from release due to law or security considerations. This would include, but not limited to, export Control, ITAR, CUI, materials necessary to be held confidential, privacy or medical information, or materials with limited release due to security concerns.*
- Software restricted from release due to intellectual property considerations. This would include, but not limited to, copyright, trade secrets, commercial information, patent protected software, or similar information which is protected under law.*



# Software as part of the scientific process

## Stages of reproducibility:

1. Software is not mentioned
2. Software is mentioned, but not described
3. Software is described, but not made available or by request only
4. Software is made openly available
5. ...with a permissive license
6. ...with good documentation and testing
7. Bonus: Contributed to a generalized package

## Challenges in reproducibility:

- Data and plots are shared, but it isn't clear how the plots are derived from the data
- Student leaves the field or the software is lost
- Algorithms are described but with insufficient detail to reproduce
- Configurations or steps are not shared.

# SOFTWARE MANAGEMENT PLANS

SMP shall include:

- The types of software to be produced in the course of the project
- The standards to be used for software development
- Policies for accessing and sharing the software, including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, and other rights or requirements
- Policies and provisions for re-use, re-distribution, and the production of derivatives
- Plans for providing access to the software used in any science publication
- Plans for archiving and preserving of the software, as appropriate (use of existing databases or public repositories will be strongly encouraged), including how long the software will be preserved or maintained
- And shall make publicly available appropriate metadata describing the software

# Advantages of Open Source Scientific Software

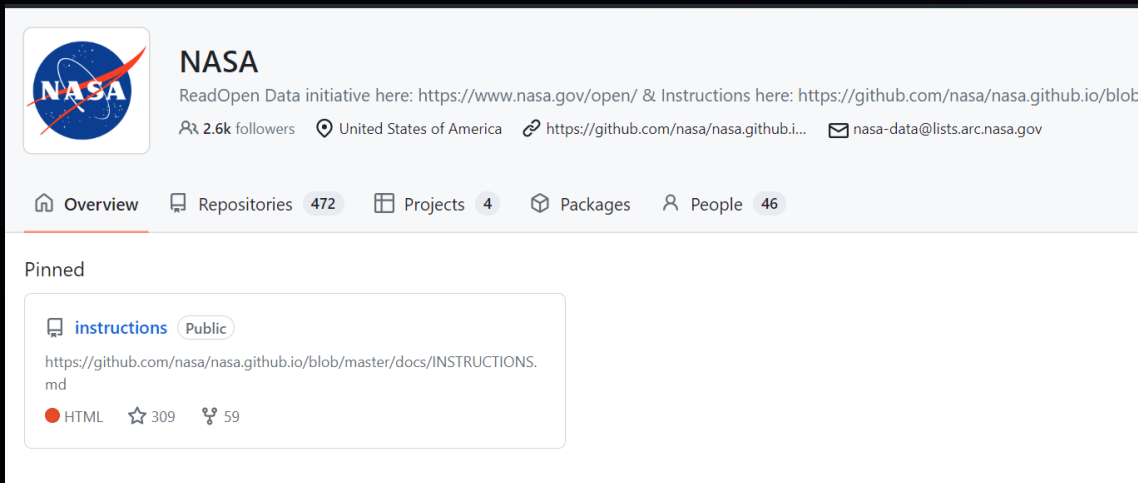


- Increased reproducibility
- Reuse
- Curation and archiving
- Better understanding of the results
- Credit

# Tools for Open Science

## Version Control Platforms

GitHub, Gitlab, bitbucket



The screenshot shows the NASA GitHub profile page. At the top left is the NASA logo. To its right, the text reads "NASA" followed by "ReadOpen Data initiative here: <https://www.nasa.gov/open/> & Instructions here: <https://github.com/nasa/nasa.github.io/blob/master/docs/INSTRUCTIONS.md>". Below this, it says "2.6k followers", "United States of America", and "nasa-data@lists.arc.nasa.gov". A navigation bar includes "Overview", "Repositories 472", "Projects 4", "Packages", and "People 46". Under the "Pinned" section, there is a card for a repository named "instructions" (Public) with the URL <https://github.com/nasa/nasa.github.io/blob/master/docs/INSTRUCTIONS.md>. It also shows "HTML", "309" stars, and "59" forks.

## Software Archiving

Software Heritage,  
Zenodo, ASCL

**Software is Fragile**

*Unlike words carved in stone, it  
can be deleted or get corrupted*

## Software Publishing and Curation

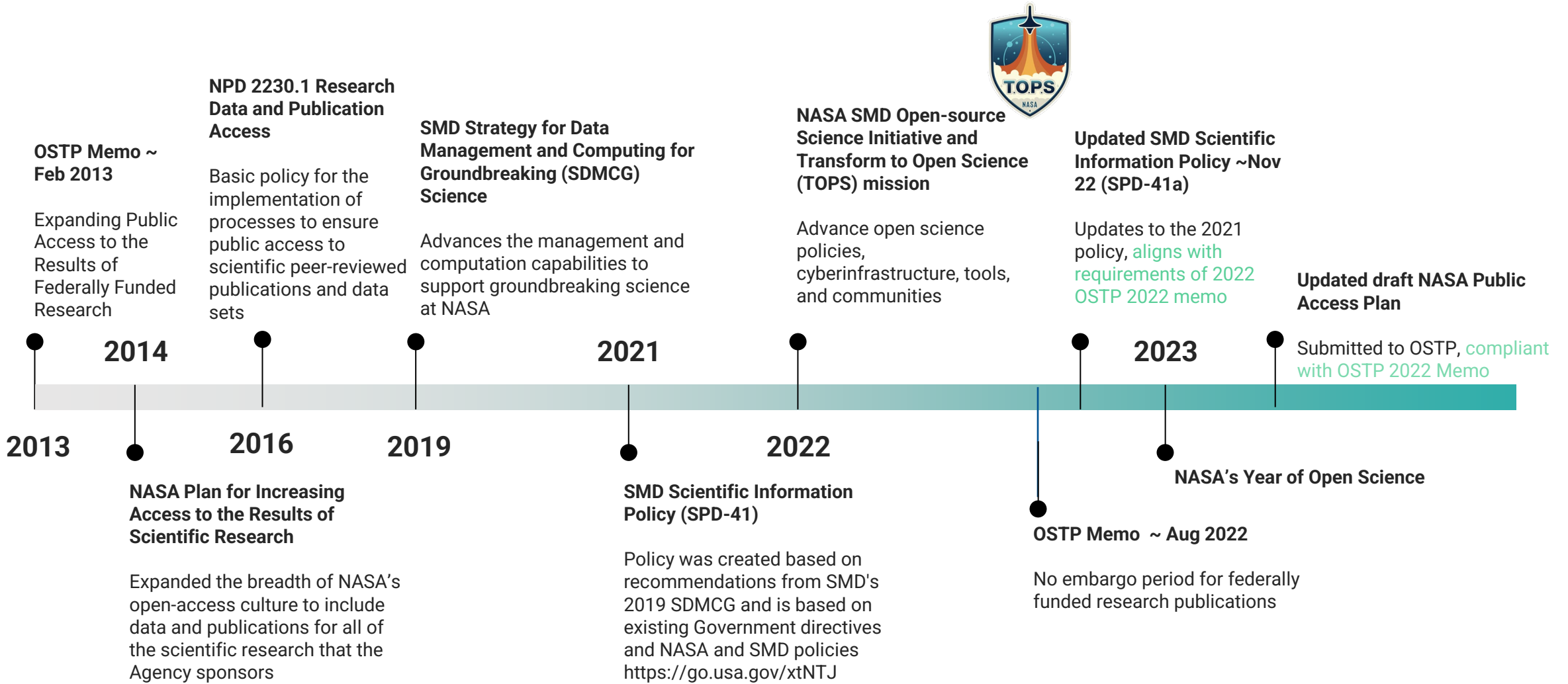
JOSS, ADS, AAS, ...



The screenshot shows the homepage of The Journal of Open Source Software (JOSS). The header includes the JOSS logo, the text "The Journal of Open Source Software", and navigation links for "About", "Papers", and "Docs". The main content area features the text: "The Journal of Open Source Software is a developer friendly, open access journal for research software packages." Below this, it states: "Committed to publishing quality research software with zero article processing charges or subscription fees." At the bottom, there are two buttons: "Submit a paper to JOSS" and "Volunteer to review".







NPD = NASA Policy Directive; SMD = NASA's Science Mission Directorate; SPD = SMD Policy Document



A surreal landscape where nature meets space. On the left, a large, detailed tree trunk stands on a grassy hillside. In the foreground on the right, a simple wooden bench sits on a grassy slope. The middle ground is a vast, hazy expanse of clouds. In the sky, a satellite with solar panels orbits a large, detailed moon. A jet flies across the sky, leaving a bright trail. The background is a vibrant, multi-colored sky with stars and nebulae. A dark blue rectangular box is overlaid on the left side of the image, containing the word 'Questions?' in white text.

Questions?





# Open Access Models

- **Diamond/Platinum** – Journals that publish open access without charging authors a fee. Publishers require funding from other sources, such as advertisements, academic institutions, societies, or government grants.
- **Bronze** – The articles are free to read on the publisher’s website, but do not have a clear license (generally not available for reuse).
- **Black** – The unauthorized digital copying of paywalled content and making it freely available (large-scale copyright infringement). Some examples are [#ICanHazPDF](#) or [SciHub](#).
- **Gratis and Libre** (used in the [Budapest Open Access initiative](#))
  - **Gratis** – free online access to read, without re-use rights
  - **Libre** – free online access to read, plus some additional re-use rights



# Open Access Resources

- CORE: <https://core.ac.uk/>
- CrossRef: <https://search.crossref.org/>
- Directory of Open Access Books (DOAB): <https://doabooks.org/>
- Directory of Open Access Journals (DOAJ): <https://doaj.org/>
- Google Scholar: <https://scholar.google.com/>
- Open Access Directory (OAD): [https://oad.simmons.edu/oadwiki/Main Page](https://oad.simmons.edu/oadwiki/Main_Page)
- SCOAP<sup>3</sup>: <https://scoap3.org/>
- Sherpa Services: <https://beta.sherpa.ac.uk/>