2023: Year of Open Science

2023: A Year of Open Science!

On 11 January 2023 the US White House — joined by 10 federal agencies, a coalition of more than 85 universities, and other organizations — declared 2023 to be a Year of Open Science.

- White House factsheet
- NASA announces a Year of Open Science in Nature
- Federal website
- NSF releases a Year of Open Science Toolbox including powerpoint templates, letterhead, poster, and virtual backgrounds!

This year will celebrate the benefits and successes of open science and inspire more scientists to adopt open science. Ultimately, the success of the Year of Open Science will be driven by collaborations with individuals, teams, and organizations who are ready to transform the culture of science into one that celebrates openness, and inclusion!

A Year of Open Science has set out 4 goals to work towards:

1. Establish strategic approaches for advancing open science
2. Promote equitable participation in open science through transparency, integrity and equity of reviews.
3. Account for open science activities in evaluations and incentives
4. Engage underrepresented communities in the advancement of open science and research

Participating organizations

- Aligning Science Across Parkinson's (ASAP)
- ASAPbio
- Center for Open Science
- Children's Tumor Foundation (CTF)
- Gathering for Open Science Hardware (GOSH)

- Health Research Alliance (HRA)
- Higher Education Leadership Initiative for Open Scholarship (HELIOS)
- Life Sciences Editors Foundation (LSEF)
- PREreview
- Open Research Funders Group (ORFG)
- Creative Commons
- Scholarly Publishing and Academic Resources Coalition (SPARC)
- The Michael J. Fox Foundation for Parkinson's Research
- GO FAIR US
- Springer Nature
- STM
- Federation of American Scientists (FAS)
- The Wilson Center
- Frontiers

For You

Are you just beginning your open science journey? Perhaps you have only just begun to post your code or data online, share your pre-prints, or share your null hypothesis as part of your grant application? Or perhaps you are exploring science communication on a personal blog. All of these behaviors exemplify open science and we would like for you to join us in the Year of Open Science!

If you are excited to learn and encourage others to learn about open science as part of the Year of Open Science, then we invite you to explore how to take Individual Actions for the Year of Open Science!

For Your Team

Is your team interested in adopting open science principles and practices, but unsure of how to begin? The Year of Open Science team activities could help you reflect together on how to equip the people at your organization with practical knowledge in open science, as well as introduce them to a wider community of open science practitioners.

Learn how to get your team involved in Team Actions for the Year of Open Science!

For Your Entire Organization

Is your organization is ready to devote people, funds, and other resources to adopting open science throughout the organization. If your leaders are committed to becoming known as an open science organization, and excited to facilitate open science research through its policy and actions, then we invite you to explore Organization Actions for the Year of Open Science!

Open Science 101

TOPS Open Science Curriculum: Open Science 101

TOPS aims to meet everyone where they are at on their open science journey. While some members of our growing community may have practiced open science for many years, others may be trying to transition to a new way of conducting research or are students who are looking to begin their scientific careers. As such, TOPS' first priority is to develop the infrastructure to train 20,000 scientists and researchers as part of our five-year program. This community-developed open science curriculum will introduce those beginning their open science journey to important definitions, tools, and resources; and provide participants at all levels recommendations on best practices from subject matter experts. TOPS will initially focus on developing the Open Science 101, designed as five modules that mimic a scientific workflow.

For the 2023 Year of Open Science, TOPS is developing strategic partnerships with scientific associations to teach open science during large annual meetings, special science team summer schools, and other events. The vision for the Open Science 101 and other learning resources is for a CC-BY licensed online, open course, hosted on the OpenEDx platform, that can be used to train scientists and award NASA open science badges.

The five Open Science 101 modules are:

1. Ethos of Open Science
2. Open Tools and Resources
3. Open Data
4. Open Software
5. Open Results

Content for the Open Science 101 is being derived from several amazing resources, including: - OpenSciency - OpenbyDesign - Utrecht - Mozilla - and more....

Learning Objectives

The Open Science 101 aims to introduce learners to a nuanced understanding of open science, enabling participants to better understand an open science workflow from end to end. The focus of the curriculum will be on providing learners with a basic understanding of open science, its ethos and benefits, and how to actively participate in open science communities. The TOPS curriculum will also be used to support researchers looking to engage with NASA as NASA moves to adopt more open science requirements (e.g., Scientific Information Policy and ROSES opportunities). Scientists will need to acquire the new skills highlighted in the curriculum to participate in open science effectively and to demonstrate those skills when applying for NASA funding opportunities.

Researcher core open science skills

- Have an ORCID
- Familiar with data management and software management plan best practices and resources
- How to find and identify community accepted data and software repositories
- How to openly license and share FAIR data and assign a DOI
- Apply a permissive license and share open-source software and assign a DOI

- Organizing open meetings

Interested in learning these skills now? Head over to our Open Science Guide and begin in Section 1 of "Guide for Your Open Science Journey".

Curriculum

The design is for a 12.5-hour course consisting of five, 2.5-hour modules that could be taught in-person during large society meetings or during science team meetings, completed individually online, or completed via group-organized online events.

Additional information about the module structure and suggestions for design are provided.

Course Creation

For the workshop and online course materials, Open edX has many partners that specialize in working with subject matter experts to create online courses. To train 20,000 scientists, the MOOC will need to be engaging and interactive, similar to Elements of AI, with animations rather than videos of people. Using animations rather than videos will make the MOOC easier to translate into additional languages.

Instructor-led Workshop Materials

For instructor-led workshops (in-person and virtual), the materials will be organized as follows:

- Each module is to be organized into five, 30-minute lessons for a total of 2.5 hours of content. We recommend scheduling the module for three hours and taking two 15-minute breaks.
- Each lesson has content suitable for in-person or virtual workshops (slides with talking points).
- Each lesson ends with an activity or knowledge check that can be completed by both in-person or virtual participants.

- Participants visit the MOOC to take assessment and earn module badge.
- All content will be assigned a CC-BY license.

Online Course Materials

For online coursework, including the Open edX course, the materials will be organized as follows:

- Each module is to be organized into five lessons, each approximately 30 minutes in length, for a total of 2.5 hours of content. All five modules, if taken together, will be 12.5 hours of content.
- Each lesson has content suitable for online learning, graphics, and/or animations.
- Each lesson ends with an activity or knowledge check that can be completed online.
- Every module except the "Ethos of Open Science" module will offer a "test-out" mechanism, such as an assessment, to allow for advanced learners to earn the credit for that module.
- All content will be assigned a CC-BY license.

Open edX platform

For the MOOC, Open edX has many partners that specialize in customizing and deploying an Open edX platform.

- Single sign-on using ORCID
- Track user engagement
- Auto-grading of assessment questions
- Badging
- Data analytics
- Mobile-ready
- 10,000+ users, autoscaling

Badging

There will be a badging process for in-person, virtual, and Open edX module completion. Initial details are here

In order for the TOPS Open Science badge to be offered as part of in-person or virtual workshops, the following minimum requirements must be met.

- The event must be registered with the TOPS team.
- Teach at least one module in full (Note: Although modules can stand alone, we recommend teaching Ethos of Open Science alongside your chosen module, as it provides the foundational understanding for Open Science practices and benefits to users).
- Have at least one certified instructor or apply for a waiver (application link TBD).
- Must use TOPS Open Science 101 curriculum materials.
- Survey course participants before and after completion, using a survey that TOPS will provide for your use.
- Abide by the TOPS Code of Conduct.

What is open science?

Open Science is 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.

NASA and Open Science

AT NASA, we believe that open science is also a commitment to a collaborative culture, enabled by technology, that empowers the open sharing of data, information, and knowledge within the scientific community and the wider public to accelerate scientific research and understanding. Open science is accessible, reproducible, and inclusive and has positive effects such as more

citations and usage of publications, increased scholarly collaborations, greater transparency of research, widening access to "hidden knowledge," and a broadening of participation in science which encourages more equitable systems.

Through programs such as the Open-Source Science Initiative (OSSI) and the Transform to Open Science (TOPS) mission, NASA looks to pursue an ethos of *open-source science*; open-source science accelerates discovery by conducting science openly from project initiation through implementation. The result is the inclusion of a wider, more diverse community in the scientific process as close to the start of research activities as possible. This increased level of commitment to conducting the full research process openly and without restriction further enhances transparency and reproducibility, which promotes trust in the scientific process. It also represents a cultural shift that encourages collaboration and participation among practitioners of diverse backgrounds, including scientific discipline, gender, ethnicity, and expertise.

Open science and open-source science definitions are always influenced by the particular lived experience, academic training, and research background of those putting forth that definition. A good overview of five, distinct yet interrelated definitions is provided in Fecher, B., Friesike, S. (2014). Open Science: One Term, Five Schools of Thought. In: Bartling, S., Friesike, S. (eds) Opening Science. Springer, Cham. https://doi.org/10.1007/978-3-319-00026-8_2.

Here are some of the definitions of open science, as put forth by different teams at NASA:

- “When properly implemented, open science policies promote ‘efficiency, greater collaboration, reduced duplication, . . . a broadening of the user community, improved code testing, more reproducible research, and enhance[d] research transparency’ (NASEM, 2018, p. 3).”
    - From NASA Science article “Why Do Open Science”
- “NASA’s Earth Science Data Systems (ESDS) Program defines open science as a collaborative culture enabled by technology that empowers the open sharing of data, information, and knowledge within the scientific community and the wider public to accelerate scientific research and understanding. A system based on open science aims to make the scientific process as transparent (or open) as possible by making all elements of a claimed discovery readily accessible, which enables results to be repeated and validated.”
    - From NASA’s Earth Data Systems Open Science Resources
- “Open science is an evolving paradigm that seeks to foster greater inclusivity, diversity, and participation in the scientific process while increasing transparency and reproducibility.”
    - From NASA’s Open-Source Science for Data Processing and Archives Workshop
- “Open-source science is a commitment to the open sharing of software, data, and knowledge (algorithms, papers, documents, ancillary information) as early as possible in the scientific process. The principles of open-source science are to make publicly funded scientific research transparent, inclusive, accessible, and reproducible.”
    - From NASA’s Open-source Science Initiative Overview

What is open-source science (OSS)?

Open-source science is a commitment to the open sharing of software, data, and knowledge (algorithms, papers, documents,

ancillary information) from the start of research activities. The principles of OSS are to make publicly funded scientific research transparent, inclusive, accessible, and reproducible. OSS is enabled by advances in technology, including collaboration tools and cloud computing. More information is available from NASA's Science Mission Directorate (SMD) Policy Document (SPD-41) on science information policy.

What is the difference between open-source science and open science?

Ramachandran et al. define open science as "a collaborative culture enabled by technology that empowers the open sharing of data, information, and knowledge within the scientific community and the wider public to accelerate scientific research and understanding." The primary difference is that open-source science commits to making the scientific process open from the start of research activities rather than making research results open once the research is complete and papers are published. The commitment to conduct research in the open supports greater participation in answering fundamental scientific questions and the use of publicly funded research, data, and analysis for societal benefit.

What is the difference between open-source science and open data?

Open data are a critical component of open-source science. Other components of OSS include open documentation, publications, citizen science, challenges/prizes, open-source software, open peer review, open notebooks, and open educational resources among others.

Is the lack of open science a cultural or technical issue?

Both. Open science is more than just the open sharing of data and code. It also is a cultural shift in the scientific process that encourages collaboration among people of diverse backgrounds, including scientific field, gender, location, ethnicity, and expertise. By removing

barriers to participation in the scientific process, open-source science is inherently inclusive and collaborative. NASA's vision is to use open science principles to expand participation in the scientific process, improve reproducibility, and accelerate scientific discovery for societal benefit. Technological considerations include use of existing investments in infrastructure and mechanisms for community contributions, while limiting the proliferation of unvalidated data.

Does open science mean "free" science?

Open science is the commitment to the full, free, and open sharing of data, code and knowledge as early in the research process as possible. In terms of activities related to NASA's Open-Source Science Initiative, Science Mission Directorate (SMD) Policy Document (SPD-41) consolidates existing guidance for the openness and accessibility of data, software, papers, and ancillary information resulting from SMD-funded efforts

What is the first step to getting involved with open science?

The first step is to support open-source science efforts within your communities. The Turing Way is a community-driven guide that provides more details on how to design open projects. TOPS is developing an open science curriculum that will become available in late 2022. The UNESCO recommendations have detailed definitions and suggestions on areas of action to support open science.

What incentives and disincentives are being used to encourage open science?

NASA awards funding based on the strength of the scientific ideas and the ability to advance those for societal benefit. Traditionally, the enterprise has awarded individuals, but without a focus on openness; incentives have been based on publishing papers in big-release journals. However, NASA will take the next year to shift incentive structures from what they have been in the past, to include incentivizing open science activities (eg. collaborations, team-

building, open data, open software, and open-access publications). This shift requires NASA to engage with professional organizations, academia, etc. to make it happen and be appropriately recognized. It is important to note that this shift is not automatic but rather will require some experimentation to see what works and what does not. It is also imperative that NASA get feedback from the community to ensure the effectiveness and efficiency of its OSS efforts.

How do I know that my code will not be taken without proper attribution?

One of the principles of open science is proper attribution of previous work, collaborations, and knowledge used from various sources. Making your code open, appropriately licenced, and assigning it a DOI will help researchers track code being developed. The open development of code should make it easier to identify when people aren't properly attributing their work.

Can you recommend some resources to help people develop their research using open science principles?

One of the goals of TOPS is to develop resources to help researchers, organizations, and citizen scientists do their work using open science practices and principles. TOPS is developing an open science curriculum that will become available in late 2022. Please sign up for our newsletter to receive the latest updates!.

How can I advocate for open science?

A few ways to get started are: Make data non-proprietary and available in an open repository; Provide datasets in standardized formats and assign them DOIs; Develop open-source software and code, using best practices and rigorous version control, so that people can reuse it; Support community development and encourage reuse; Publish in open-access journals; Actively engage the public through storytelling (blogging, social media), hack-a-thons, and citizen science; and cite your data, software, and documentation.

Can you expand a bit more on how NASA is thinking about open science in terms of the research outputs domain?

TOPS is advocating a vision of open science in which the entire research workflow--from inception to the creation of data and software artifacts and publishing results--is as open as possible.

What kind of training and educational initiatives could we implement to make open science more accessible?

ScienceCore is just the beginning; TOPS will announce in late 2022, the upcoming Year of Open Science which will include a comprehensive plan on engaging with the scientific community through hackathons and summer/winter schools, and at all the large science society conferences. Check out our calendar and we hope to see you there!

How can research and data-driven artists further science?

One way to broaden participation is through making science more accessible and the arts are one way to do that. We hope to have a Space Apps challenge focused on STEAM initiatives to get people from all of the world thinking about this.

How does citizen science connect to open science?

We realize that the road to making open science a reality doesn't begin and end with academics and NASA scientists. We want to reach science-interested populations too! Citizen science provides an opportunity for the general science-interested public to get involved with scientific research to address societal needs, particularly those at a regional or local level, and to advance innovation.

How can data scientists and machine learning experts help further open science?

There are open science principles that those working with code and data can incorporate into their work, even if it is not "traditional" scientific research. They can make the underlying data findable, accessible, interoperable and reusable (which is known as the FAIR

principles). Any code which is developed should be as open as possible (e.g., open-source or white-listing); including the creation of clear documentation so that others can build on your work.

Start doing open science!

Are you new to open science *today*?

Dive right in by doing these activities!

–> Get an ORCID ("Open Researcher and Contributor ID")

–> Sign up for a GitHub account!

–> Sign up for a Zenodo account! *Tip: You can use your newly made GitHub or ORCID account to register for Zenodo!*

Are you new to open science *this month*?

–> Make your first pull request (PR) on GitHub! We will shortly have an interactive page on GitHub where you can make a PR and see your addition once your PR is approved! You can also follow these guidelines to contribute directly to the TOPS GitHub repository or our this website!

–> Start learning an open-source programming language!
  –> **Python**: Project Pythia's "Zero to Python" where you can run code interactively in your browser!
  –> **R**: "R Programming: Zero to Pro"

–> Find an open science community to join! Some communities are centered around scientific discipline, while others are focused around certain aspects of the open science process (e.g. open-source software). While each community functions differently, they often have guidelines about which software packages are used within a field, and forums for questions and discussions. See this list of open science communities - we

have only just started this list, so please add your community if you don't see it listed!

Are you new to open science *this year*?

–> Use what you've learned so far to help give back to the open science community! Answer questions in the TOPS GitHub Discussions space (and in other open science community forums) or contribute to open-source software packages by opening issues and making pull requests.

–> With a year under your belt, consider helping TOPS promote open science at upcoming conferences and events! Making sure everyone is aware of other benefits of open science will help the scientific community transform to a truly open and inclusive effort. TOPS has created some templates that you can follow to organize a townhall, host a workshop, and much more!

Core Data and Computing Services Program

CONTENTS

The Core Data and Computing Services Program (CDCSP) will provide a layered architecture on which SMD science Divisions can seamlessly and efficiently integrate their discipline-specific services such as data archives. Central to the development of Core Data and Computing Services will be the integration of new and existing data and computing capabilities, including existing Division capabilities, into a modular and secure architecture such that they are reusable by all SMD Divisions.

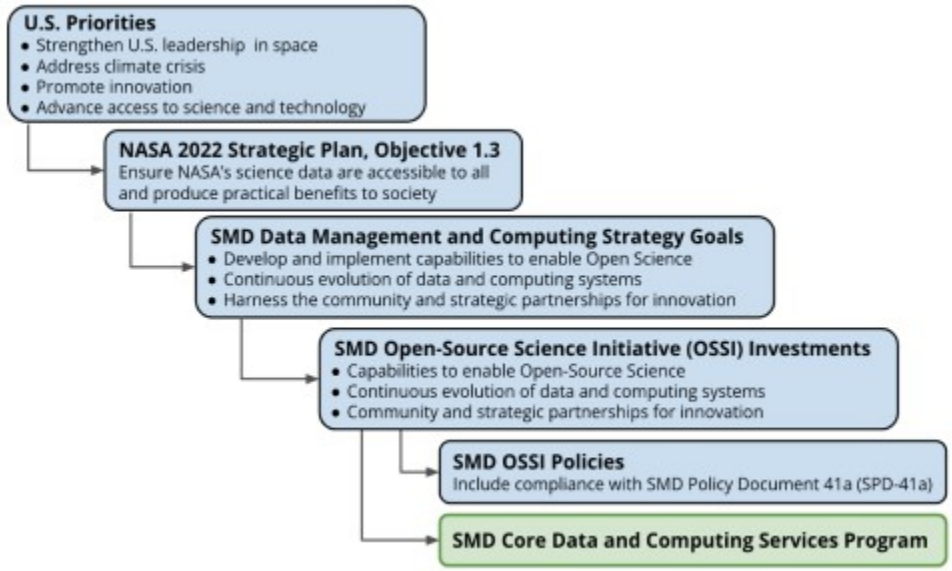
SMD Core Data and Computing Services will:

- Develop SMD-wide data and computing infrastructure to support Open Science
- Develop services to support the adoption of SPD-41a by SMD Divisions

Core Data and Computing Services Background

Open Science is 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. Broad adoption of Open Science best practices will be critical to increasing the transparency, pace, and inclusivity of the scientific process.

SMD's Chief Science Data Office (CSDO) is leading efforts to develop an SMD-wide data and computing infrastructure to support Open Science by initiating SMD Core Data and Computing Services. Core Data and Computing Services will provide a foundation for a layered architecture (Figure 2) over which SMD science Divisions can seamlessly and efficiently integrate their discipline-specific services. Central to the development of Core Data and Computing Services will be the integration of new and existing data and computing capabilities into a modular and secure architecture such that they are reusable by all SMD Divisions. The infrastructure and services provided by this program will make it easier for SMD Divisions to support Open Science in alignment with SMD Policy Document 41a (SPD-41a).



The Core Data and Computing Services Program (CDCSP) flows from U.S., NASA, and SMD priorities/strategies. The program will define a Core Architecture of general Core Services across SMD Divisions and unique Core Services for each Division.

What will the SMD Core Data and Computing Services Program do?

- Develop and operate an extensible SMD-wide cloud computing framework that is harmonized with existing high-end computing capabilities
  - Technical support including onboarding, cost modeling, transition planning and best practices
  - Account and configuration management (e.g. tools supporting configuration management, egress, billing and metrics)
  - Security and financial controls
  - Backup solution for irreplaceable data
- Provide funding and technical support for Division prototypes and testing of cloud storage and compute while the SMD-wide cloud computing framework is being developed
- Provide access to and management of public data programs for hosting and distribution of Division-defined high-value NASA science datasets within commercial public data programs
- Provide resources to help grantees gain no-overhead access to cloud computing, including training, design, cost modeling and design of cloud-native analysis software
- Provide resources and platforms to support discovery and access of scientific data, scientific information, and software

SMD Core Data and Computing Services Division and Community Engagement

Community Engagement

- RFI: “Scientific Data and Computing Architecture to Support Open Science” released January 5 2023, closed Feb 21 2023; 75 responses received (internal and external), final report in progress
- Public Virtual Townhall for RFI: Scientific Data and Computing Architecture (January 16, 2023)
- 17 Open Workshops (Fall 2022) that included talks from relevant internal and external organizations
- 2-day SMD User Needs Open Workshop (Oct 12-13 2022, hybrid at GSFC)

SMD Division Engagement

- 4 Technical Interchange Meetings (TIMs) held with Division archives (BPS, HPD, APD, and PSD) to gather specialized Division needs
- Bi-weekly Study Steering Committee meetings held with Division representatives (Fall 2022)
- Regular briefings to Division on Study and Core Data and Computing Services development activities