

Your Journey Towards Open Science

Open Science Pathways in the Earth, Space,
and Life Sciences

9 May 2022

Shelley Stall
Vice President, Data Leadership
0000-0003-2926-8353
@ShelleyStall | sstall@agu.org

Chris Erdmann
Director, Data Leadership
0000-0003-2554-180X
@LibCCE | cerdmann@agu.org

AGU's position statement on data affirms that

“Earth and space science data are a world heritage, and an essential part of the science ecosystem. All players in the science ecosystem—researchers, repositories, publishers, funders, institutions, etc.—should work to ensure that relevant scientific evidence is processed, shared, and used ethically, and is available, preserved, documented, and fairly credited.”

AGU's position statement on data affirms that

“Earth and space science data are a world heritage, and an essential part of the science ecosystem. All players in the science ecosystem—researchers, repositories, publishers, funders, institutions, etc.—should work to ensure that relevant scientific evidence is processed, shared, and used ethically, and is available, preserved, documented, and fairly credited.”

150 years of *Nature*

A century and a half of research and discovery.

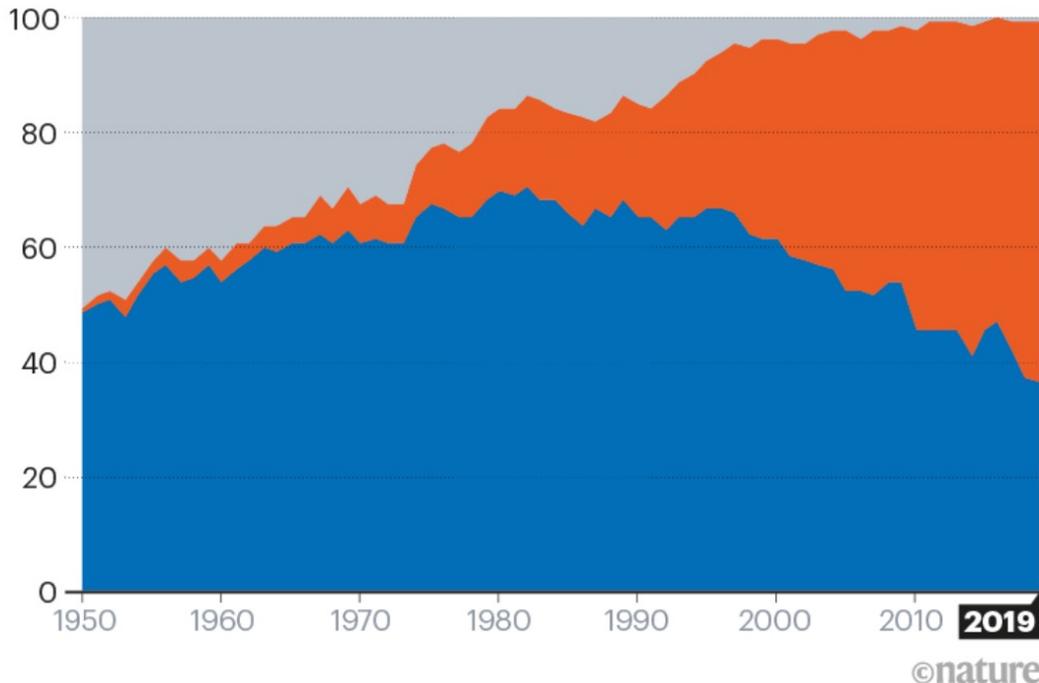


INTERNATIONAL COLLABORATIONS

Author lists on research publications show a shift towards multinational teams; fewer teams are composed entirely of researchers from one country.

Proportion of papers

■ Multinational ■ Domestic ■ Single author



Monastersky, R., & Van Noorden, R. (2019). 150 years of Nature: a data graphic charts our evolution. *Nature*, 575(7781), 22–23.
<https://doi.org/10.1038/d41586-019-03305-w>

The Future of your Research

- Research Teams (not individuals)
- International Collaborations
- Robust **tools to discover** relevant research worldwide
- Good **documentation** to understand that research, data, and/or software
- Data that is **interoperable**, no matter which research team created it
- Software that is **accessible** and developed in current tools (e.g., Jupyter Notebooks)
- Licenses that support **reuse**.

FAIR Guiding Principles

FAIR is...

Findable

Accessible

Interoperable

Reusable

Article in Nature journal *Scientific Data*:
Wilkinson, M. D. *et al.* The FAIR Guiding Principles
for scientific data management and stewardship.
Sci. Data 3:160018 doi: 10.1038/sdata.2016.18
(2016).





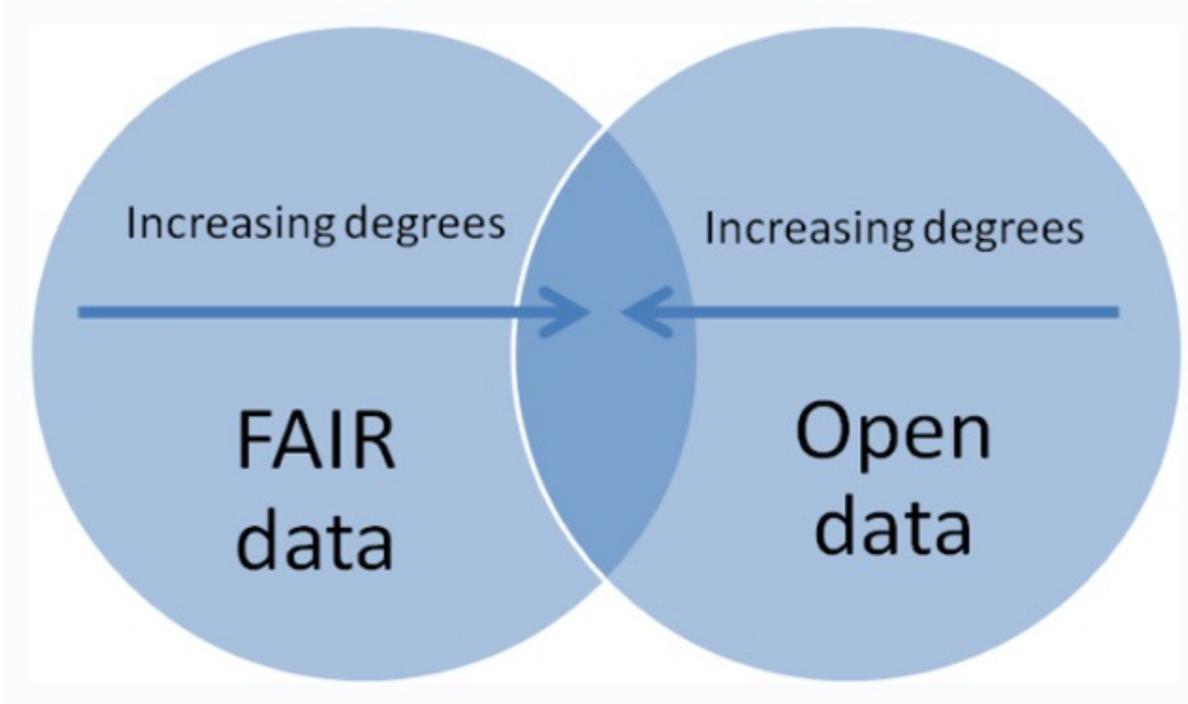
FAIR Guiding Principles

- Findable
- Accessible
- Interoperable
- Reusable

Photo Credit: The Magnifying glass, Tap, Gears set, Recycle sig, Storage, Infinity, Discussion, Shield, and Man User icons made by [Freepik](https://www.flaticon.com) from www.flaticon.com are licensed by [CC 3.0 BY](https://creativecommons.org/licenses/by/3.0/). All other icons made by ARDC. Entire FAIR resources graphic is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

Source: <https://www.andis.org.au/working-with-data/fairdata/training>

Is FAIR Open? In short, “It depends.”



Data can be FAIR or Open, both or neither.

The greatest potential for reuse comes when data are both FAIR and Open.

Higman, Rosie, Daniel Bangert, and Sarah Jones. 2019. “Three Camps, One Destination: The Intersections of Research Data Management, FAIR and Open”. *Insights* 32 (1): 18. DOI: <http://doi.org/10.1629/uksg.468>

Data should be as open as possible, as closed as necessary.

Why do we Care
about FAIR?



RESEARCHERS ♥ DATA MANAGEMENT!

HMM... WHAT DATA IS OUT THERE ...

F



... TO SUPPORT MY RESEARCH?

F-A

ULTRA SPEC METADATA SEARCH CRITERIA

Wow! SUPER RELEVANT USEFUL DATA!



F-A-R

F-A-I-R

AMAZING RESEARCH HAPPENS.

Now... TO SUBMIT MY unique & valuable RESEARCH DATA.



F-A-I-R

Why DOES THE DATA MANAGER INSIST ON SO MUCH METADATA?



F-A-I-R

DATA MANAGER

Sigh... I THOUGHT FOR SURE THIS ONE HAD MADE THE CONNECTION.

DATA MANAGERS ... MAKING DATA MORE USEFUL ... AND PAVING THE WAY FOR FUTURE RESEARCHERS TO FIND VALUE IN YOUR DATA!

Shelley

The Future of your Research

- Research Teams (not individuals)
- International Collaborations
- Robust **tools to discover** relevant research worldwide
- Good **documentation** to understand that research, data, and/or software
- Data that is **interoperable**, no matter which research team created it
- Software that is **accessible** and developed in current tools (e.g., Jupyter Notebooks)
- Licenses that support **reuse**.



F-A



F-A-R



F-A-I-R



R

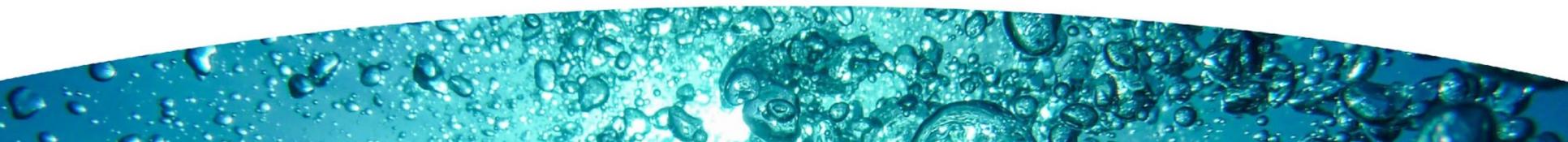
Open Science – Transparency and Trust

“Increased openness leads to increased transparency and trust in scientific information...”

Source: UNESCO
Recommendation on Open
Science; adopted
November 2021



<https://en.unesco.org/science-sustainable-future/open-science>



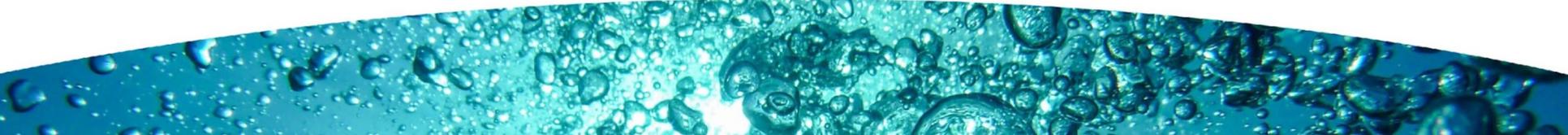
Open Science – The Key Pillars

- Open scientific knowledge
 - **Scientific publications**
 - **Open research data**
 - **Open source software and source code**
 - **Open hardware**
- Open science infrastructures
- Science communication
- Open engagement of societal actors
- Open dialogue with other knowledge systems.

Source: UNESCO
Recommendation on Open
Science; adopted
November 2021



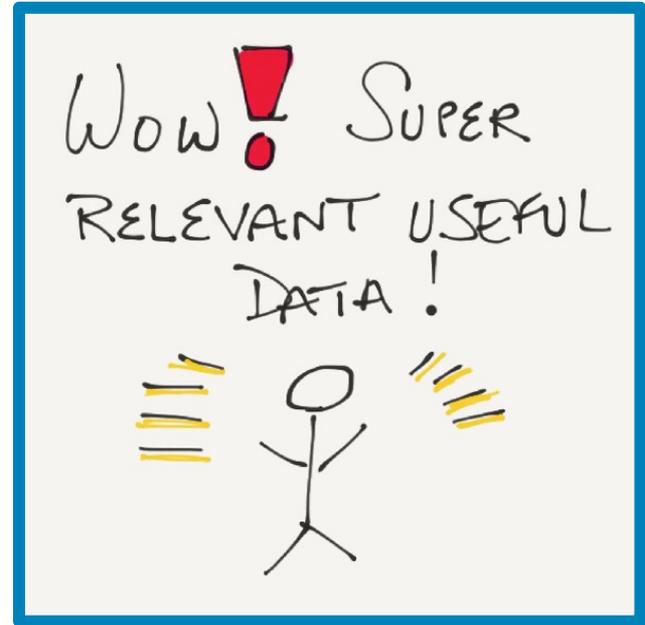
<https://en.unesco.org/science-sustainable-future/open-science>



From the very beginning of the research process,

the researcher **both contributes** to open science and

takes advantage of the open science practices of other members of the research community.





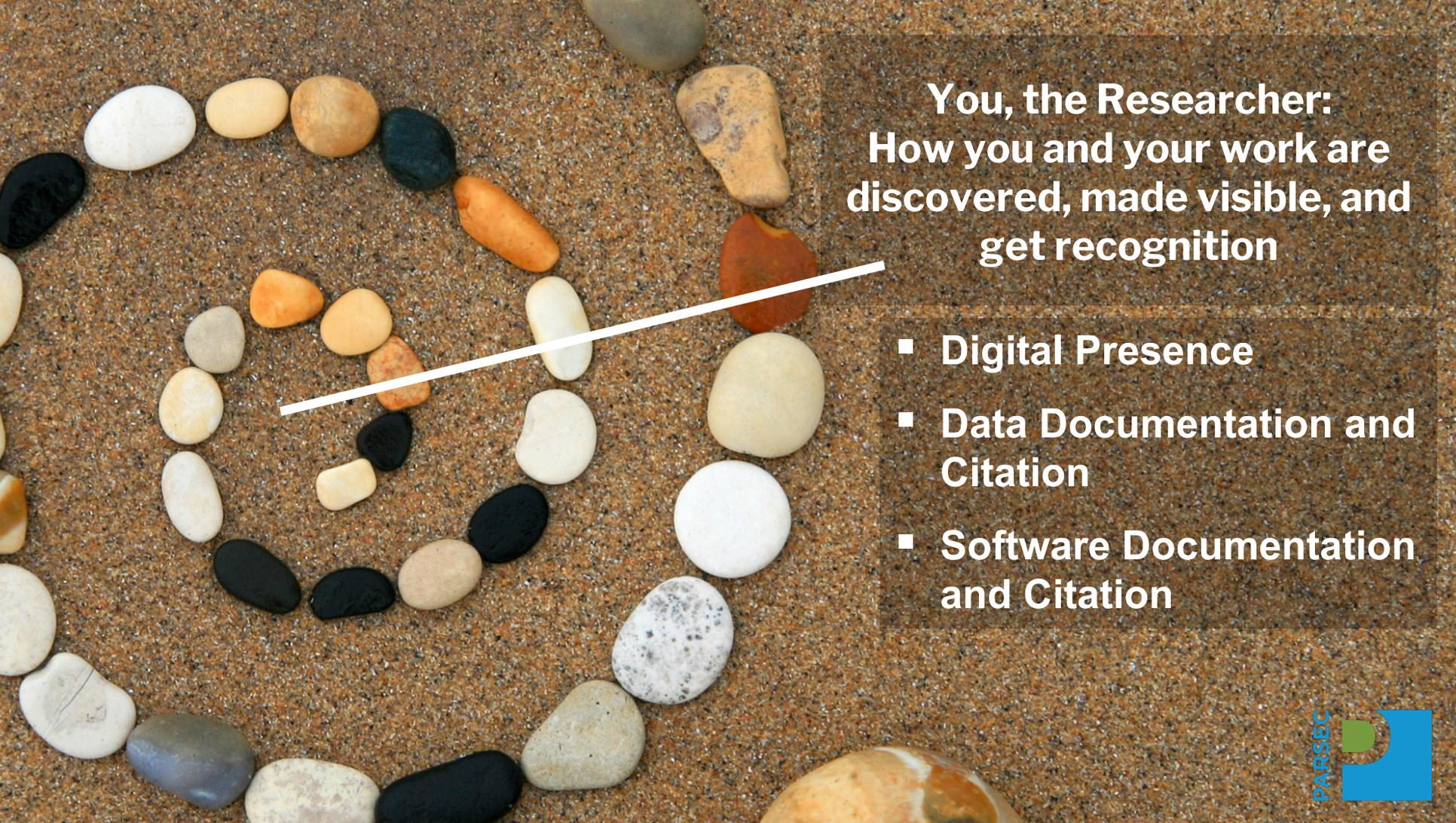
Wai...wai...wait.

This is all new to me!

What do you mean
this matters to my
research?!?

Building Your Open Science Skills

- **You, the Researcher:** How you and your work are discovered, made visible, and get recognition
- **Your Research Team/Lab:** Working Openly
- **Your Community:** Improving interoperability, sharing, and reuse beyond your team
- **Beyond Your Community:** Preparing for Cross-domain challenges



**You, the Researcher:
How you and your work are
discovered, made visible, and
get recognition**

- **Digital Presence**
- **Data Documentation and Citation**
- **Software Documentation and Citation**



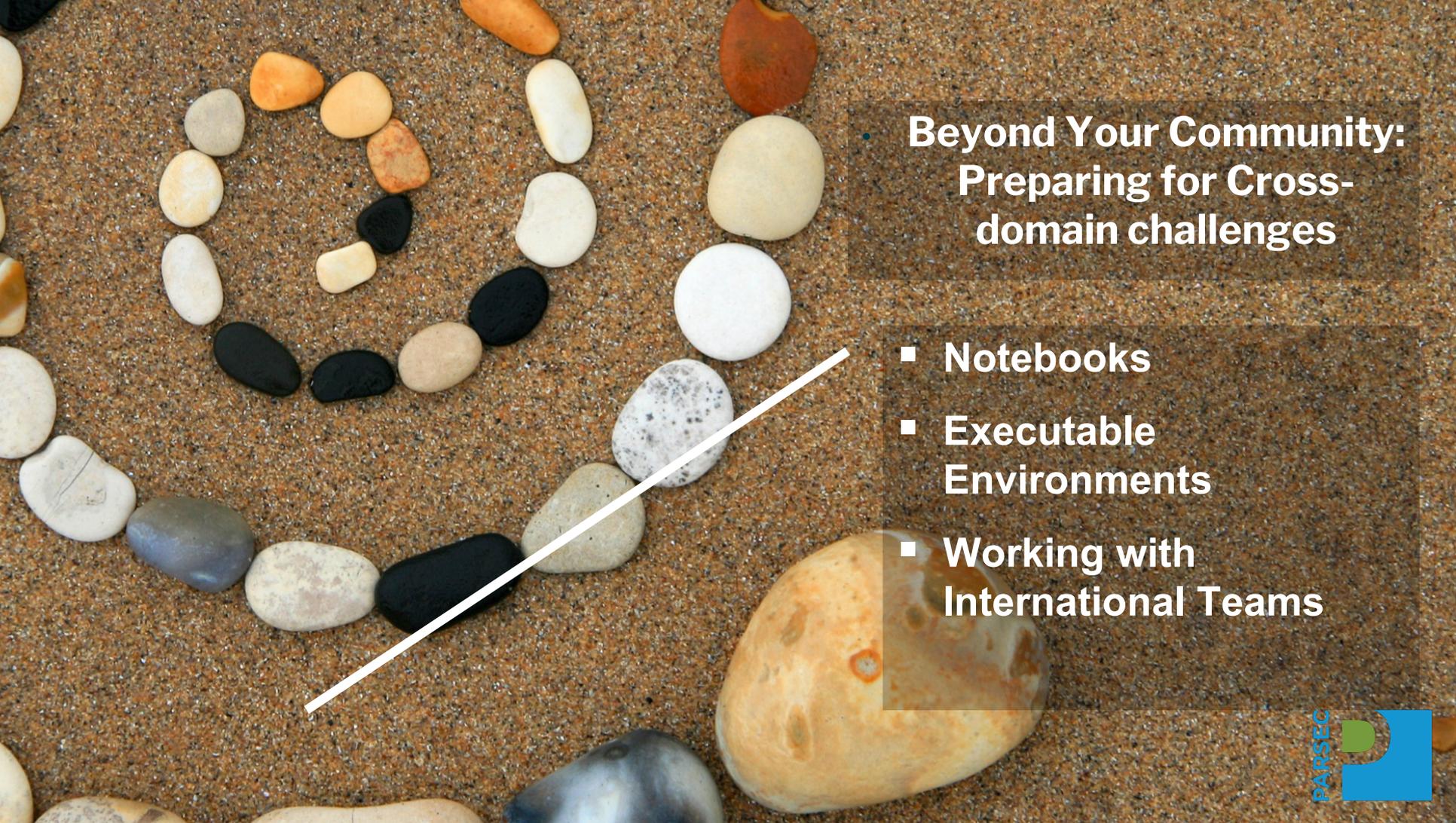
Your Research Team/Lab: Working Openly

- Research Team Open Science Practices
- Resources and Lab Guidelines
- Project Archiving



Your Community:
Improving interoperability,
sharing, and reuse beyond
your team

- Transparency and Understanding
- Usability
- Trust and Persistence



Beyond Your Community: Preparing for Cross- domain challenges

- Notebooks
- Executable Environments
- Working with International Teams

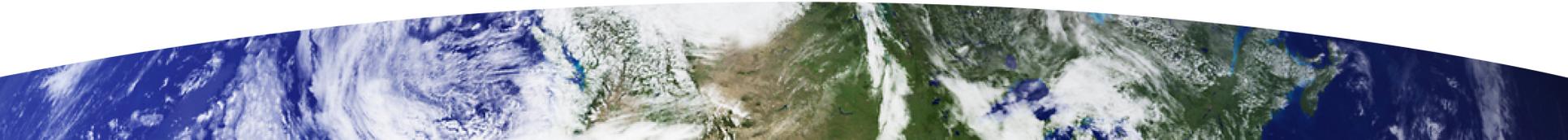
Let's Start the Journey to Open Science...

- Digital Presence
- Data Documentation and Citation
- Software Documentation and Citation



What is your Digital Presence?

1. How **you and your research** appear in online content.
2. How well **your work is integrated and connected** in the scientific record through your publications, datasets, software, and other digital objects and content.





[https://orcid.org/
0000-0003-2926-8353](https://orcid.org/0000-0003-2926-8353)

Emails >

sstall@agu.org

Website and social links >

[AGU Data Leadership](#)

 Is this you? [Sign in to start editing](#)

 [Printable version](#)

Name

Shelley Stall

Biography

Shelley Stall is the Senior Director for the American Geophysical Union's Data Leadership Program. She works with AGU's members, their organizations, and the broader research community to improve data and digital object practices with the ultimate goal of elevating how research data is managed and valued. Better data management results in better science. Shelley's diverse experience working as a program and project manager, software architect, database architect, performance and optimization analyst, data product provider, and data integration architect for international communities, both nonprofit and commercial, provides her with a core capability to guide development of practical and sustainable data policies and practices ready for adoption and adapting by the broad research community.

Shelley's recent work includes the Enabling FAIR Data project (<https://copdess.org/enabling-fair-data-project/>) engaging over 300 stakeholders in the Earth, space, and environmental sciences to make data open and FAIR targeting the publishing and repository communities to change practices by no longer archiving data in the supplemental information of a paper but instead depositing the data supporting the research into a trusted repository where it can be discovered, managed, and preserved.

Activities

[Collapse all](#)

 **Employment (1)**

 Sort

American Geophysical Union: Washington, DC, US

2015-06-22 to present | Senior Director (Data Leadership)
Employment

[Show more detail](#)



Quarterly

Founded: 2018

E-ISSN: 2641-435X

More About *Data Intelligence* 

Journal Resources

Editorial Info

Abstracting and Indexing

Growing the FAIR Community at the Intersection of the Geosciences and Pure and Applied Chemistry

[Shelley Stall](#) , [Lesley McEwen](#) , [Lesley Wyborn](#), [Nancy Hoebelheinrich](#) and [Ian Bruno](#)

Posted Online November 01, 2019

https://doi.org/10.1162/dint_a_00036

© 2019 Chinese Academy of Sciences Published under a Creative Commons Attribution 4.0 International (CC BY 4.0) license

ORCID

Crossref DOI

DIGITAL PRESENCE CHECKLIST



Connect your research to your data, software, institution, and more. Use this checklist to optimize your digital presence, increase discovery of your work to potential collaborators and partners, and receive credit when others use your work.

You. Your ORCID.

- Have your own ORCID.** It provides a persistent digital identifier that distinguishes you from other researchers and supports automated linkages between you and your research activities. Go here to register: <https://orcid.org>, and select “For Researchers”.
- Include your ORCID on all scholarly work.** This includes your publications, datasets, software, presentations, posters, signature block of your emails. Everything. This helps with linking to your ORCID profile.
- Keep your ORCID profile current.**
 - Enable automatic updates from Crossref and DataCite. AGU [Digital Presence blog post](#) has the detailed steps.
 - Set a reminder every three months to ensure all your work is connected and current in your ORCID profile. Make sure your current affiliation and email are included and public for viewing. Add a second email (which can be private) to ensure account access should one become locked.

This work is part of the **Building New Tools for Data Sharing and Re-use through a Transnational Investigation of the Socioeconomic Impacts of Protected Areas (PARSEC)** project with funding provided by the Belmont Forum through the National Science Foundation, Grant 1929464 as well as the **Accelerating Open and FAIR Data Practices Across the Earth, Space, and Environmental Sciences: A Pilot with the NSF to Support Public Access to Research Data** project funded by the National Science Foundation, Grant 2025364.



Checklist

- <http://doi.org/10.5281/zenodo.4706118> (English)

Tutorial – 15 min

- <https://doi.org/10.5281/zenodo.4706146> (slides, link to recording)

AGU Data & Software Sharing Guidance

What is covered:

- What data needs to be available?
- Repository Selection
- Availability Statement
- Data & Software Citation
- Citation Formatter
- Models & Simulations
- Journal Specific Guidance
- International Geo Sample Numbers
- Data Help Desk

AGU ADVANCING EARTH AND SPACE SCIENCE

JOIN RENEW GIVE LOGIN Q

Data & Software for Authors

WHAT IS NEEDED?

AGU requires that the underlying data needed to understand, evaluate, and build upon the reported research be available at the time of peer review and publication. Additionally, authors should make available software that has a significant impact on the research. This entails:

1. Depositing the data and software in a trusted repository, as appropriate, and preferably with a DOI
2. Including an [Availability Statement](#) as a separate paragraph in the Open Research section explaining to the reader where and how to access the data and software
3. And including [citation\(s\)](#) to the deposited data and software, in the Reference Section.

Click on the headings below for detailed information on:

- [Models & Simulations](#)
- [Journal-Specific Data Guidance](#)
- [International Geo Sample Numbers](#)

Most of your questions regarding data and software should be answered by the resources below. Just in case, if you still have questions, you can contact DataHelp@agu.org.

WHAT DATA NEEDS TO BE AVAILABLE?

Primary and processed data used for your research should be preserved and made available. Generally, the underlying data are considered to be the types of data usually preserved in domain repositories for each discipline. These may include raw data, but are usually the processed or refined data that support and lead to the described results and allow other readers to assess your conclusions and build off your work.

In your paper, cite these data, as well as any data you used from other sources, and include information about access to the data in the availability statement. For [model or simulation data](#), follow [journal specific guidance](#) on prioritizing preserved output; in general, availability of software is most important.

Very large data (greater than 1 terabyte or TB) can be a challenge to preserve as there often fees and additional resources required. One option to consider, institutions often offer solutions for data preservation and compliance. Again, refer to the [journal specific guidance](#) for more information or email DataHelp@agu.org.

<https://www.agu.org/Publish-with-AGU/Publish/Author-Resources/Data-and-Software-for-Authors>

Resources



Manage your Digital Objects – Research
Team Member Checklist



Software Citation - 5 Tips



Digital Presence - Connect your research
for better discovery



Guidance for AGU Authors - Jupyter
Notebooks



Data and Software Sharing Guidance for
Authors Submitting to AGU journals

Thank you

Shelley Stall

AGU Sr. Director, Data Leadership

sstall@agu.org

@ShelleyStall

<https://orcid.org/0000-0003-2926-8353>



This work is generously funded through grants from:

