Skills for Open Science: Sharing Your Software, Data, & More

Kristina Vrouwenvelder, Ph.D
Program Manager, Open Science
ORCID: 0000-0002-5862-2502

kvrouwenvelder@agu.org

AGU OPEN SCIENCE

24 journals
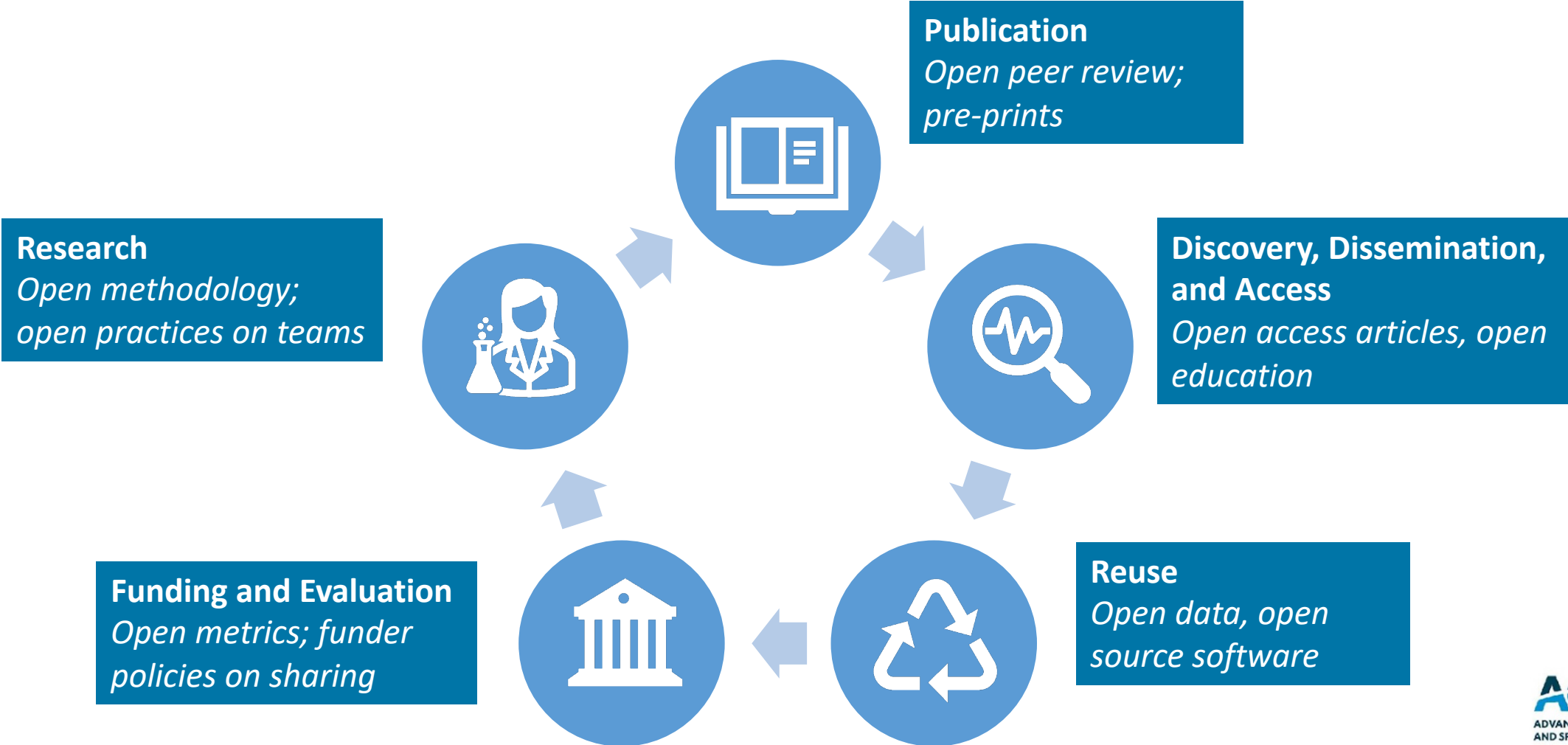
60,000 members across 137 countries

Annual meeting of ~25k researchers

———

Nonprofit scientific society governed by Board of Directors and Council.

AGU
ADVANCING EARTH AND SPACE SCIENCES

Open Science in the Research Ecosystem

Publication
Open peer review; pre-prints

Discovery, Dissemination, and Access
Open access articles, open education

Reuse
Open data, open source software

Funding and Evaluation
Open metrics; funder policies on sharing

Research
Open methodology; open practices on teams

AGU
ADVANCING EARTH AND SPACE SCIENCE

Open Science's Benefits for Research

Accelerates Discovery

Connects researchers to datasets to better understand and address major geoscience questions[1]

Broadens Research Impact

More citations for research and more connections for researchers (and the public)[2]
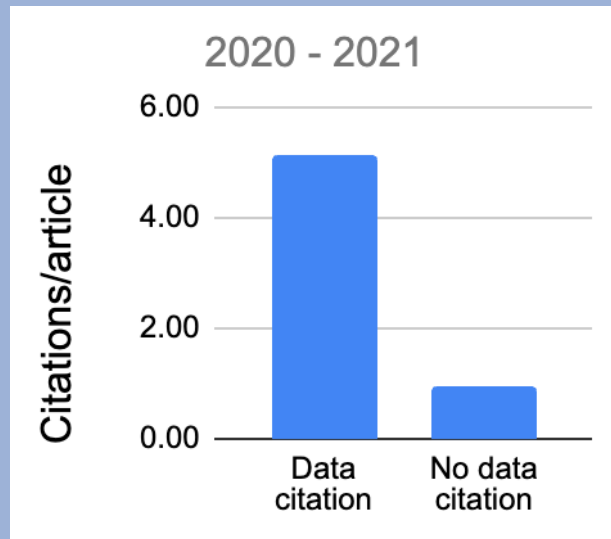
Builds Trust in Science

Enables greater research reproducibility, transparency, and access[2]

AGU
ADVANCING EARTH
AND SPACE SCIENCE

1: Maeda, 2012 (10.1007/s13280-012-0283-4); Huston, 2019 (10.14745/ccdr.v45i10a01); Sielemann, 2020 (10.7717/peerj.9954).
2: Sears, 2011 (https://ui.adsabs.harvard.edu/abs/2011AGUFMIN53B1628S); McKiernan, 2016 (10.7554/eLife.16800)

Open Science's Benefits for Research

2020 - 2021

Citations/article — Data citation vs No data citation

It's not just AGU: Papers citing data are **25%** more likely to be cited.[1]

Open access articles get up to **5x** more page views, more media coverage, and more citations.[1]

1: Colavizza et al, (2020) PLoS ONE 15(4): e0230416. https://doi.org/10.1371/journal.pone.0230416; McKiernan et al. eLife 2016;5:e16800. DOI: 10.7554/eLife.16800.

Open Science Policy in the U.S.

2013 - White House Office of Science & Technology Policy (OSTP) Holdren Memo: *Expanding Public Access to the Results of Federally Funded Research*

- 1 year embargo for research publications
- Importance of public access to research data sets

2022 - OSTP Nelson Memo: *Ensuring Free, Immediate, and Equitable Access to Federally Funded Research*

- Dec. 31st, 2025 deadline
- Publications and their supporting data publicly accessible, no embargo
- Data access for all federally-funded research data

Open Science at AGU
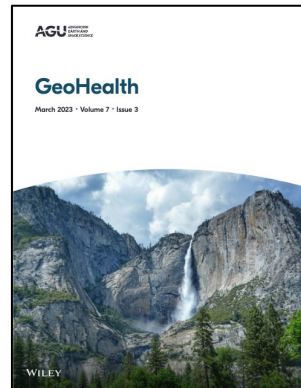
2017: AGU partners with community, COPDESS on enabling FAIR Data project (copdess.org).

**2021**: Data and Software Sharing Guidance for Authors Submitting to AGU Journals published.

2022: First AGU journal *(GeoHealth)* enters **Data Citation Pilot.**
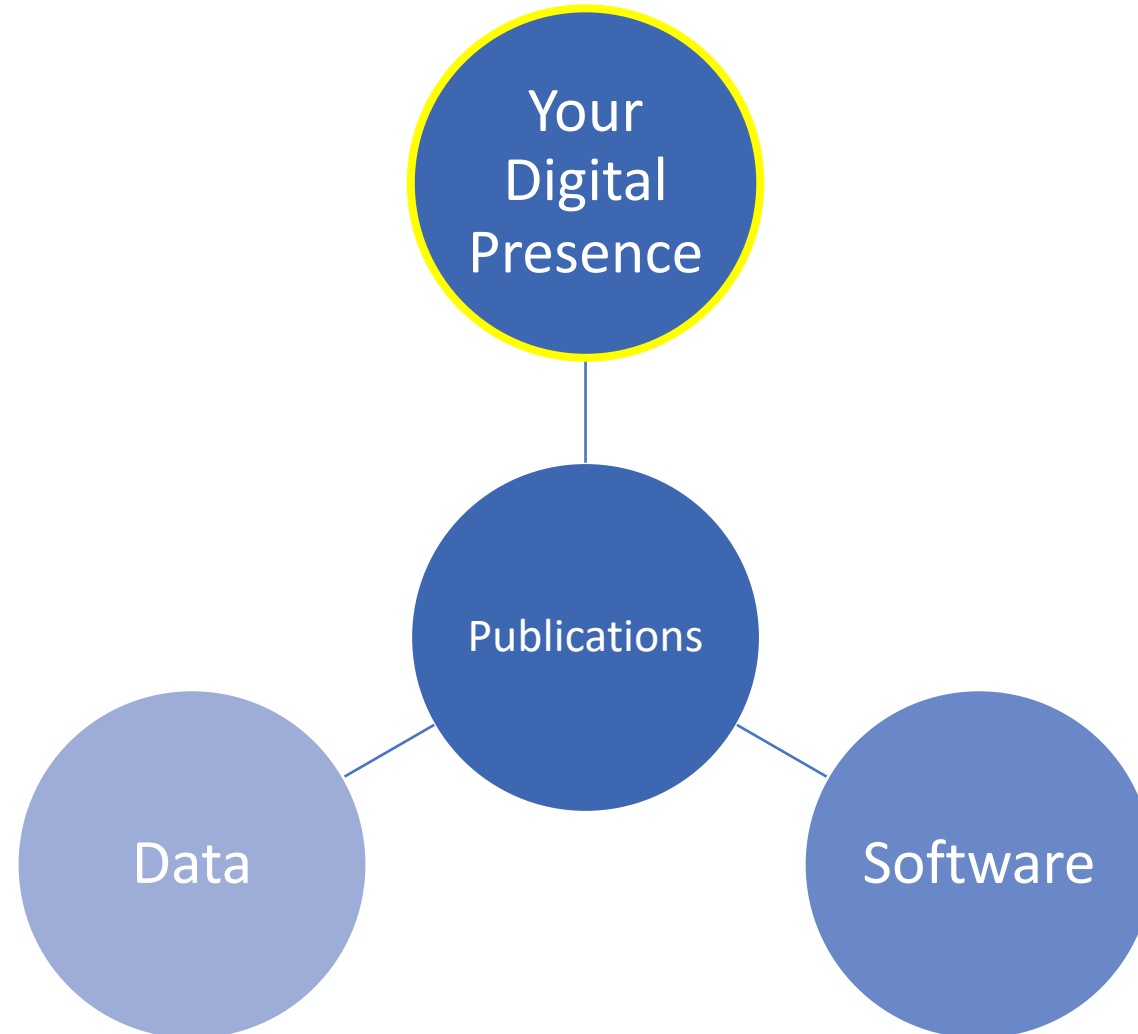
Aug 2023: All AGU journals entered in Data Citation Pilot.

Dec 2023: First **Open Science Recognition Prizes** are awarded at AGU23.

All of AGU's journals offer open access options; 12 are fully open.

Open Science at AGU: the Data Citation Pilot
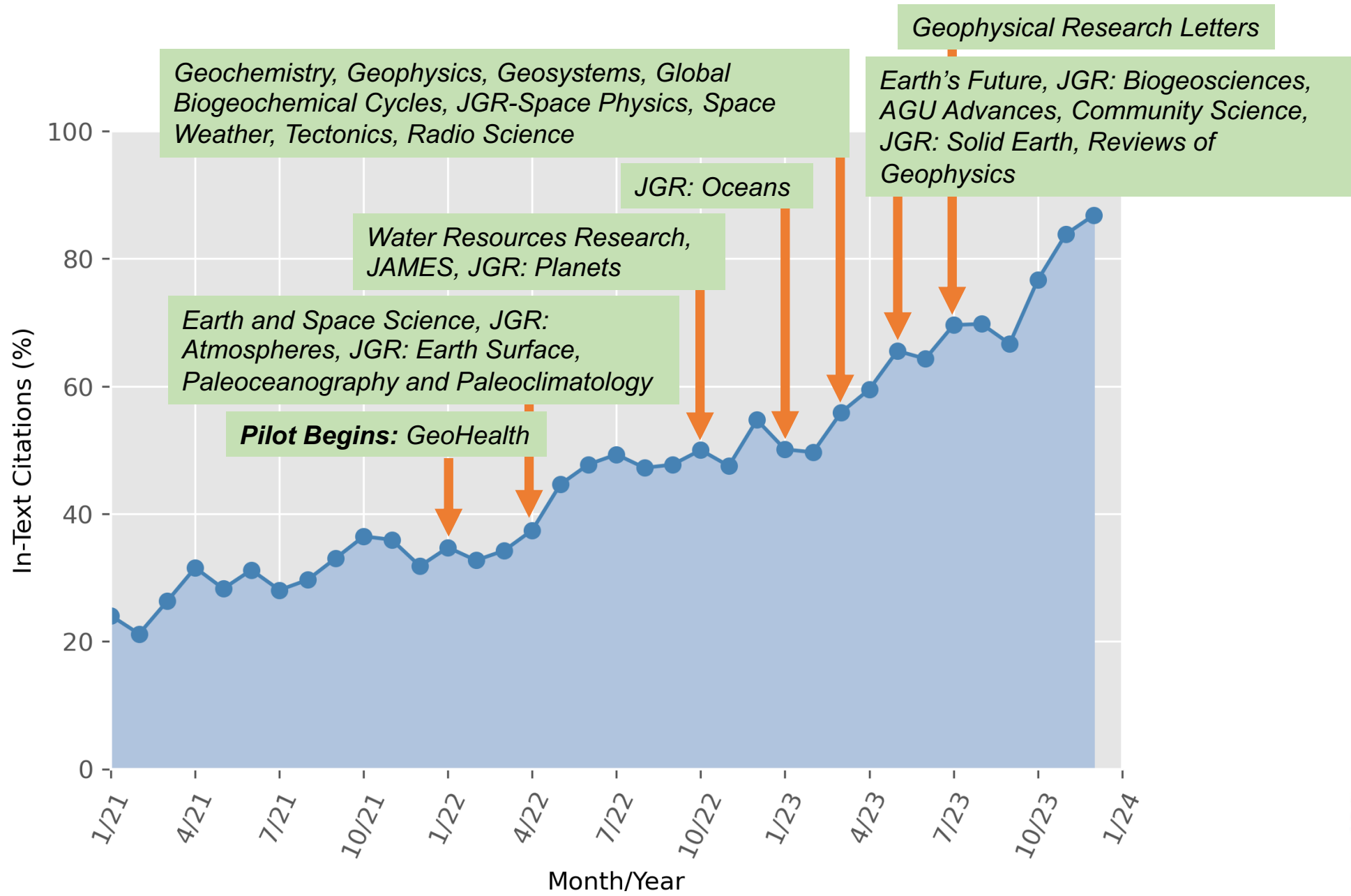
Skills for Opening Your Science

# Who Searches Your Digital Presence?

✓ Researchers reviewing the literature

✓ Possible collaborators

✓ Advisor, institution, or employer

✓ Journal editors looking for reviewers

✓ Funders

✓ Societies and associations

✓ Community groups, local government, ...
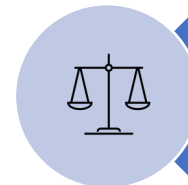
Who Searches Your Digital Presence?

✓ Researchers reviewing the literature

✓ Possible collaborators

✓ Advisor, institution, or employer

✓ Journal editors looking for reviewers

✓ Funders

✓ Societies and associations

✓ Community groups, local government, ...
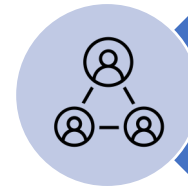
Connect with new collaborators
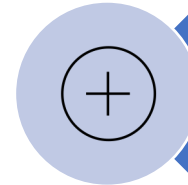
Increase citations to your work and enhance your research impact

Highlight your contributions to science *beyond* the publication

AGU
ADVANCING EARTH
AND SPACE SCIENCE

Why Do I Care?

LinkedIn
https://www.linkedin.com › charlie-shobe-46b12a202

Charlie Shobe - Marine Safety - City of Laguna Beach

Irvine, California, United States · Marine Safety · City of Laguna Beach
Charlie Shobe · Student at Chapman university · View mutual connections with Charlie · Welcome back · Activity · Experience · Education · People also ...

Mutual Art
https://www.mutualart.com › Artist › Charlie-Shobe

Charlie Shobe | 4 Artworks at Auction

Charlie Shobe was an **artist** who was born in 1940. The **artist's** work has been offered at auction multiple times, with realized prices ranging from 27 USD to ...

Charlie shobe
https://www.charlieshobe.com

Charlie shobe – Research Geomorphologist at the USFS ...

On 7/14/2023 I left West Virginia University to join the U.S. Forest Service Rocky Mountain Research Station as a Research Geomorphologist. Given my new job, I ...

charlie shobe

charlie shobe
Charles M. Shobe

charlie shobe **artist**

charlie **schlatter characters**

charlie **in good luck** charlie **now age**

Google Search   I'm Feeling Lucky

Report inappropriate predictions

ORCID
Connecting research and researchers

charles shobe

ABOUT   FOR RESEARCHERS   MEMBERSHIP   DOCUMENTATION   RESOURCES   NEWS & EVENTS   SIGN IN

Search

ADVANCED SEARCH ˅

Showing **50** of **37814** results.

Items per page: 50 ˅   Page 1 of 757   ‹ ›

ORCID ID	First Name	Last Name	Other Names	Affiliations
0000-0003-3015-1283	Charles	Shobe		College of William and Mary, Helmholtz Centre Potsdam - GFZ German Research Centre for Geosciences, University of Colorado, West Virginia University
0000-0001-8818-0541	William	Shobe	Bill Shobe	Lewis & Clark Law School, Regents of the University of Minnesota, University of Virginia
0000-0002-7563-7482	Elizabeth	Shobe		Stockton University, University of Toledo
0000-0001-6137-5058	Gladriel	Shobe		

Using your ORCID

- Education
- Work History
- Funding
- Publications
- Peer review & service
- Presentations, posters, pre-prints
- Datasets
- Software

AGU Advances

Research Article | 🔓 Open Access

Strong Physical Contrasts Across Two Mid-Lithosphere Discontinuities Beneath the Northwestern United States: Evidence for Cratonic Mantle Metasomatism

Tianze Liu ✉, Emily J. Chin, Peter Shearer

CORRESPONDING AUTHOR
Tianze Liu

✉ tianzeliu@gmail.com

ⓘ orcid.org/0000-0001-7953-2939

.org/10.1029/2023AV001014

cle is available as a PDF in the Supporting Information.

📄 PDF   🔧 TOOLS   ⤴ SHARE

Barriers to Data and Software Sharing: The Reference Manager Gap

🅰 THE UNIVERSITY OF ARIZONA

Natalie Raia (0000-0003-4939-3282), Postdoctoral Research Associate, School of Information, University of Arizona
Kristina Vrouwenvelder (0000-0002-5862-2502), Program Manager, Open Science Leadership, American Geophysical Union

AGU ADVANCING EARTH AND SPACE SCIENCES

Motivation: Are reference managers adequately developed to enable data and software citation?

Method: Compare citation import, export, and

Dataset Citation

Q: Are repositories providing citation guidance on their

Q: What are common citation errors among reference managers?

Zotero (browser plug-in)

Using your ORCID – everything else!

Your ORCID Record

Activities   Expand all

> **Employment** (1)   ☰ Sort   ← Add your affiliations and employers!

> **Professional activities** (2)   ☰ Sort   ← Your service

> **Funding** (4)   ☰ Sort   ← Grant information
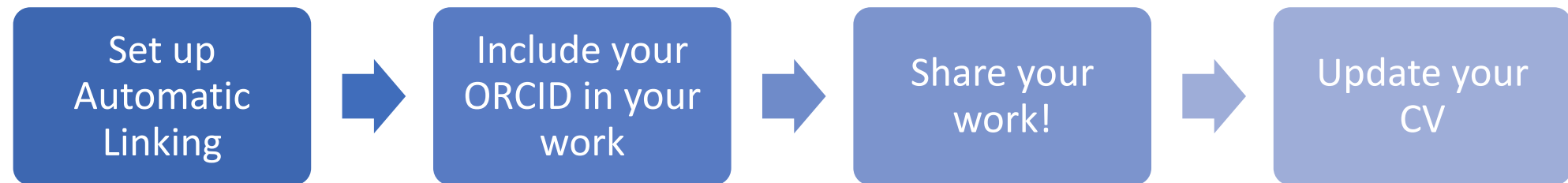
> **Works** (50 of 209)   ☰ Sort   ← Journal articles, presentations, datasets, software...

> **Peer review** (1 review for 1 publication/grant)   ☰ Sort   ← Even your peer reviewing!

AGU
ADVANCING EARTH
AND SPACE SCIENCE

Keeping your Digital Presence Updated

Set up Automatic Linking → Include your ORCID in your work → Share your work! → Update your CV

Set Up Automatic Linking:

For CrossRef, to link your publications:

For DataCite, to link other works (presentations, posters, datasets, software):

If Something's Missing:

- **Auto-Linking**: Add works by direct import from other systems to import links to your publications and other works to your ORCID record.
- Add works using an identifier
- Import works from a BibTeX file
- Add works manually

Include your ORCID in...

- **manuscripts** & submission systems

- grant applications (as prompted)

- your institution's research information management system

- Your **other shared works**, including datasets, software/code, presentations and posters

- Your business cards, website, and **CV** or resume

Article title

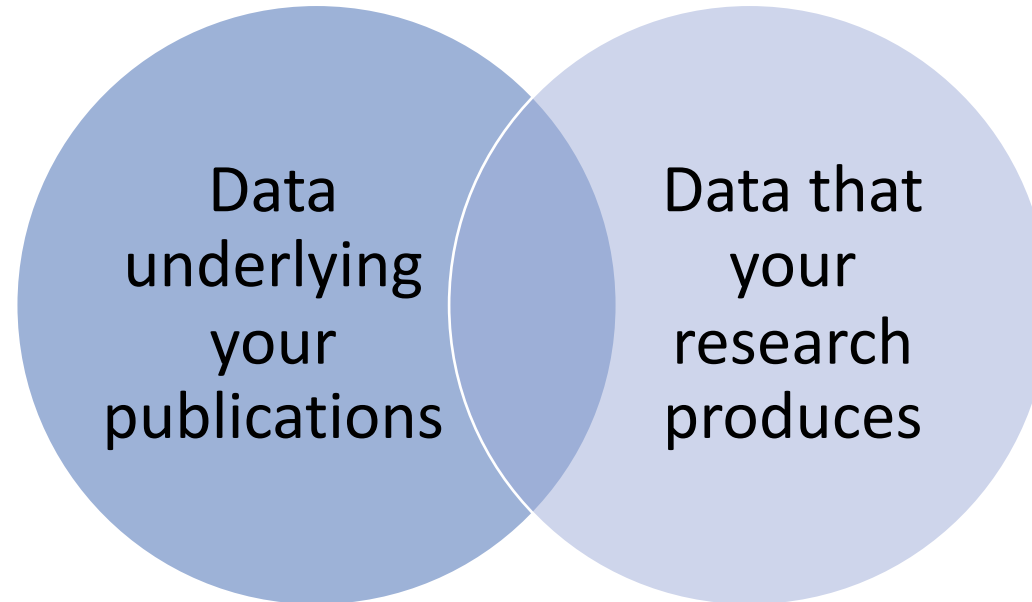Author A 🆔 , Author B 🆔 , Author C, Author D, Author E 🆔

Connect your ORCID to ScienCV to import your BioSketch and Current and Pending Support!
Also compatible with Canadian Common CV, or export your works from ORCID to a BibTex file to update your CV!

ORCID records will export directly to Digital Measures-- used by many universities.

AGU
ADVANCING EARTH
AND SPACE SCIENCE

Skills for Opening Your Science

What data should I share?

Data underlying your publications

Data that your research produces

This might include:
- Data from **observations** collected in the field;
- Data from **satellites** (primarily level 2 or 3);
- Data from **laboratory experiments**;
- All data displayed in the figures of the paper

Get your data ready

Sharing your data enables reuse and collaboration (more impact!)

Documentation:

- Include good metadata – the data about the data

- Create standardized, logical file structures and file names

- Include README files with your datasets
    - Define abbreviations, leave notes about methodologies, …

Storage

- Use non-proprietary file formats designed for long-term use

- Set up automatic backups!

- Preserve your data in a repository with a persistent identifier (DOI) and metadata appropriate for your discipline/domain

Discovery

- Cite your data in your papers!

- Add your data to data portals, as appropriate

See also: https://www.usgs.gov/data-management/why-share-your-data ;

Choosing a repository

Your data's forever home

- Are there repositories specific **to your discipline**?

- Other musts:
  - Datasets should receive a citable & unique persistent ID, like a **DOI**

- Google Drive, Box, FTP servers, etc.
  - Often incompatible with peer review
  - Not citable
  - Not permanent storage

Skills for Opening Your Science

Get your software ready

Sharing software builds collaboration and increases transparency.

- Use **open source software tools** when possible

- Create **standardized, logical variable names** and file structures

- Choose a **license** for your code
  - Visit https://choosealicense.com/

- Use a system for **version control!**
  - Platforms such as **GitHub**, **GitLab**, or **Open Science Framework** (OSF), provide additional collaborative and online backup capabilities

- Add a citation.cff file to your repository so others can **cite your work.**

If you need to get started with version control, check out the Software Carpentries lesson on Git!

AGU
ADVANCING EARTH
AND SPACE SCIENCE

Other resources: https://datascience.nih.gov/tools-and-analytics/best-practices-for-sharing-research-software-faq

Supported by:

Zotero

Documenting your software

- Read the **style guide** for your programming language for best practices

- **Comment your code** and use a standardized, meaningful format for variable and function names.
  - Imagine you'll need to explain your code to a colleague.

- To **share your code** with other scientists, you may need to include other information alongside your code.
  - E.g., scenarios when your model can be applied or conditions for initial parameters

- Include README files with your code
  - Information about versions, code contributors, coding standards, dependencies…

Check out this guide to documenting code for a list of related resources!

You might find a tool like https://readme.so/ helpful for creating README files.

For an example of extensive documentation, see: DOI: 10.1029/2020EA001584 and associated notebook.

Sharing Your Software

Issuing a persistent identifier for your repository with Zenodo 🔗

To make your repositories easier to reference in academic literature, you can create persistent identifiers, also known as Digital Object Identifiers (DOIs). You can use the data archiving tool Zenodo to archive a repository on GitHub.com and issue a DOI for the archive.

> **Tips:**
>
> - Zenodo can only access public repositories, so make sure the repository you want to archive is public.
> - If you want to archive a repository that belongs to an organization, the organization owner may need to approve access for the Zenodo application.
> - Make sure to include a license in your repository so readers know how they can reuse your work.

https://docs.github.com/en/repositories/archiving-a-github-repository/referencing-and-citing-content

Notebooks Now!: Computational notebooks as scholarly communication
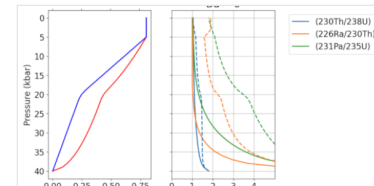
Skills for Opening Your Science
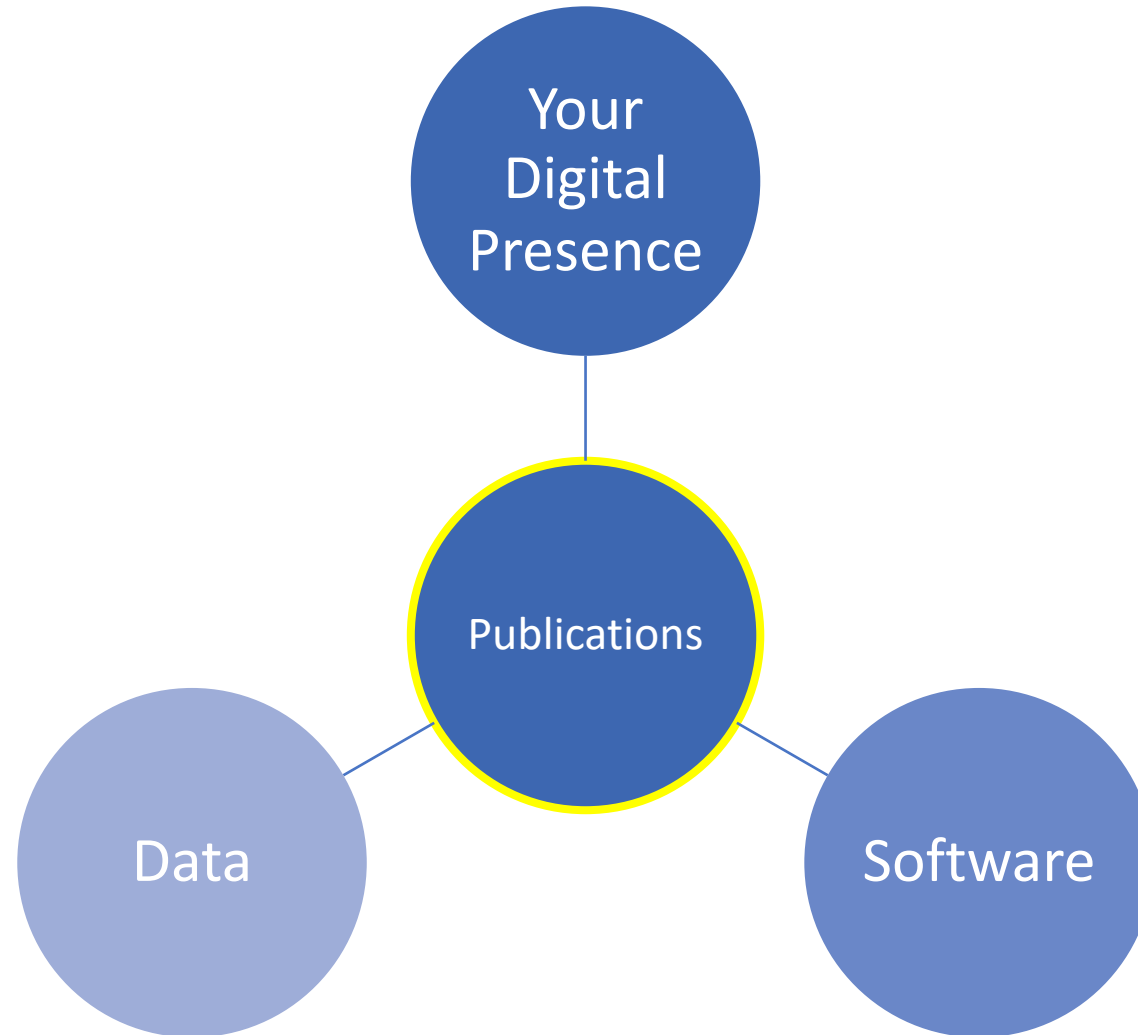
Data and software in your scientific papers

Plan where to share your data and software early!
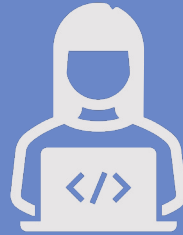
Make sure **reviewers and editors** have access to your data and software.

Include a citation to your data and software in your paper!

A quick note about software & scientific papers

What do we mean by software? This could include…

- software or code that you used for analysis and visualization of the data
- software or code used to produce a model output

- **Proprietary programs** that can't be shared:

  - Mention in your Methods section and Availability Statement in the Open Research Section

- If your work depends on scripting in **Python, R** or another scientific programming language, and/or builds off an existing model/analysis package, you should…

  - **If it's your code:** publish it in a repository like Zenodo so it can be used by others
  - **If it's someone else's code:** cite their code, including the version!

How to write a Data(Software) Availability Statement

Appears at the end
of your paper!

Open Research

Data Availability Statement

All software and processed data required to enable the reader to reproduce our results are published via Zenodo at https://doi.org/10.5281/zenodo.8398701 (Baker et al., 2023). The MITgcm (Marshall et al., 1997) is documented at https://mitgcm.readthedocs.io and available at https://doi.org/10.5281/zenodo.1409237 (Campin et al., 2019). Observational CTD data along the SR1b section, collected by the ACCLAIM project and used in Figures 1e and 2 are available from https://www.bodc.ac.uk/data/bodc_database/ctd/search/ by selecting cruise identifier JR20091118. CTD and LADCP data along the SR1b section collected by the DIMES project used in Figure 4 are available from https://www.bodc.ac.uk/data/bodc_database/nodb/search/ by selecting cruise identifier JR20130317.

Supporting Information

References

How to write a Data(Software) Availability Statement

Include the following elements:

- **A brief description** of the type(s) of data or software

- **Repository Name(s)** where they are deposited

- **DOI (Persistent Identifier)**; or, if no DOI is available, Link to Data or Software

- **Citation in References section** [for all data and software with DOIs]

- **Version** of data or software

- **For Software Only**: link to publicly accessible development platform, if you'd like to encourage reuse

- **Access Conditions** (e.g. if Registration is Required)

- **Licensing/Permissions** (e.g. Creative Commons Attribution)

Check out AGU's full author guidance!

Example: Availability Statement

Data:

The [type of data] data used for [brief context, description] in the study are available at [repository, source name] via [DOI, persistent identifier link, OR URL if no persistent identifier is available] with [license, access conditions] [in-text citation in References, required for each DOI]

Software:

[Version number] of the [software name] used for [brief context, description of what the software was used for] is preserved at [DOI, persistent identifier link, OR URL if no persistent identifier is available], available via [license type, access conditions] and developed openly at [software development platform link]. [in-text citation in References, required for each DOI]

Check out AGU's full author guidance!

Example: Data or Software Citation

Data:

Edmunds, P. J., Didden, C., & Frank, K. (2021). Mean percentage cover of corals and Porites astreoides at each site by year at St. John, VI from 1992 to 2019 (Version 1) [Dataset]. Biological and Chemical Oceanography Data Management Office (BCO-DMO). https://doi.org/10.26008/1912/BCO-DMO.843284.1

Software:
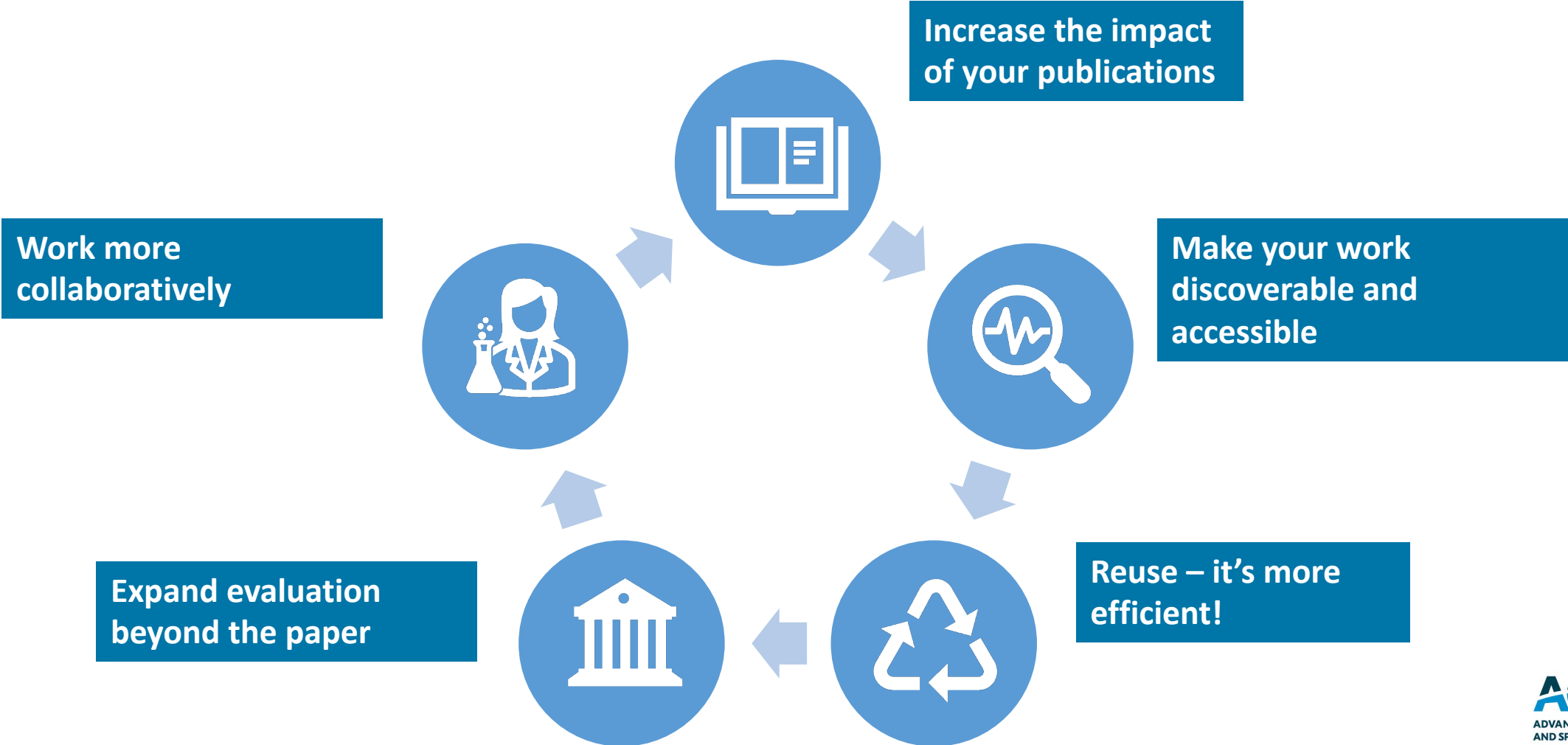
Include bracketed description!

Shobe, C. (2023). Code and data for "The uncertain future of mountaintop-removal-mined landscapes 1: How mining changes erosion processes and variables" (v1.0) [Software]. Zenodo. https://doi.org/10.5281/zenodo.10059514

Check out AGU's full author guidance!

AGU
ADVANCING EARTH AND SPACE SCIENCE

Open Science skills can help you...

Increase the impact of your publications

Make your work discoverable and accessible

Reuse – it's more efficient!

Expand evaluation beyond the paper

Work more collaboratively

AGU
ADVANCING EARTH
AND SPACE SCIENCE

Thank you!

Questions or comments?
Email kvrouwenvelder@agu.org

For more Open Science content at AGU:
data.agu.org

AGU
OPEN SCIENCE

Begin Your Open Science Journey:

BUILDING YOUR DIGITAL PRESENCE

BUILD YOUR DIGITAL CV:

☐ **Register for an ORCID** at https://orcid.org. It provides a persistent digital identifier that distinguishes you from other researchers and supports automated linkages between you and your research activities.

> 💡 *TIP: Include your ORCID on all scholarly work: publications, datasets, software, presentations, posters, and the signature block of your emails.*

☐ **Enable automatic updates** from Crossref and DataCite in your ORCID profile. Follow the QR code on this postcard for a tutorial!

☐ **Set a calendar reminder** every three months to check that all your work is connected and current in your ORCID profile.

Flip this card for more tips!

Get this digital postcard:

Begin Your Open Science Journey:

CITING AND MANAGING YOUR DATA

MANAGE YOUR DATA:

☐ **Store your data in non-proprietary formats** where possible. Use compression file formats that don't lose information (e.g., TIFF) and use open source software where possible.

> *TIP: New habits take time! Commit to incorporating a few changes into your workflow and make a calendar event to check on progress.*

☐ **Adopt standardized file and folder naming conventions** for your files, projects, and groups.

☐ **Create README files and keep them up-to-date** with a running log of actions and changes. Include basic elements that will help you create metadata for a dataset.

☐ **Set up automatic backups** for your data and other files.

Flip this card for more tips!

Get this digital postcard:

Find more resources:

PARSEC

BELMONT FORUM

anr©

FAPESP

JST

NSF

AGU ADVANCING EARTH AND SPACE SCIENCES

AGU ADVANCING EARTH AND SPACE SCIENCE