Practicing Open Science

Principles, Ecosystems and Tools

Shelley Stall
Vice President, Open Science Leadership
ORCID: 0000-0003-2926-8353
@ShelleyStall | sstall@agu.org

Kristina Vrouwenvelder
Program Manager, Open Science Leadership
ORCID: 0000-0002-5862-2502
@kvrouwen | kvrouwenvelder@agu.org

Open Science and Data Help Desk during EGU 2023

24-28 April 2023

- Engage with Informatics Experts
- Learn Skills and Techniques
- Make your Data Open and FAIR

Visit us in person at **Booth X224**

Ask questions through twitter via **#DataHelpDesk**

View our **online gallery of resources**

Details and Full Schedule:

A Program of EGU, AGU, ESIP, and their partners

EGU European Geosciences Union

AGU ADVANCING EARTH AND SPACE SCIENCE

ESIP MAKING DATA MATTER

Open Science and Data Help Desk

EGU Booth Hours

- **Monday-Thursday:** 10:00-19:00
- **Friday:** 10:00-13:00

Bring your questions for our experts!

- How and where can I publish my data or software?
- Where can I find a certain type of data?
- What do FAIR and open really mean?
- How can I write a good data management plan?
- How can I comply with my funder and publisher requirements for my data and software?
- And more...

We look forward to seeing you!

EGU European Geosciences Union

ESIP MAKING DATA MATTER

AGU ADVANCING EARTH AND SPACE SCIENCE

What will you learn in this course?

- *What Open Science is & why it's important*
- *How to manage your digital presence*
- *How to manage your research data & software, from project beginning to publication*

AGU
ADVANCING EARTH AND SPACE SCIENCE

What IS Open Science?

- Opening access for **discovery**
- Opening opportunities for **equitable participation**
- Opening the scientific process to invite **other ways of knowing**
- Opening new channels for **collaboration**

Open Science is – in a word – *transformative*!

UNESCO Recommendation on Open Science; adopted November 2021

Image by WikiImages from Pixabay

About Open Science

Open science has the potential of making the scientific process more transparent, inclusive and democratic.

Open science:

- increases scientific collaborations and sharing of information for the benefits of science and society;
- makes multilingual scientific knowledge openly available, accessible and reusable for everyone; and
- opens the processes of scientific knowledge creation, evaluation and communication to societal actors beyond the traditional scientific community.

https://www.unesco.org/en/open-science/about?hub=686

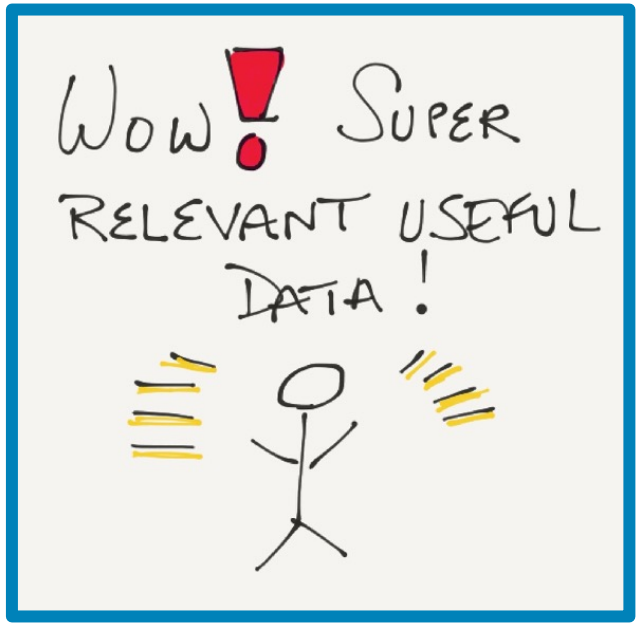
UNESCO

AGU
ADVANCING EARTH AND SPACE SCIENCE

The Key Pillars of Open Science

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

UNESCO

AGU
ADVANCING EARTH AND SPACE SCIENCE

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.**"

https://www.agu.org/Share-and-Advocate/Share/Policymakers/Position-Statements/Position_Data

Why should we care about open science (and data, and software)?

There is a practical benefit for researchers:

- Your research is **easier to evaluate** by others (including peer reviewers).
- Your work can be **discovered in different ways** than just through your paper.
- Your data will be **preserved** as part of the scientific record and **linked** to both you and your publication. (not true for supplemental information)

Citations: Colavizza et al, (2020) PLoS ONE 15(4): e0230416. https://doi.org/10.1371/journal.pone.0230416

Why should we care about open science (and data, and software)?

There is a practical benefit for researchers:

Papers that cite data are up to 25% more likely to be cited by others.

Open access articles are cited more often, whether published open access or archived openly.

Open access articles get more media coverage – in some cases, 2.5 – 5x more page views.

Citations: 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.

AGU
ADVANCING EARTH AND SPACE SCIENCE

Why should we care about open science (and data, and software)?

Open science matters to policymakers and the public.

“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

Why should we care about open science (and data, and software)?

Open science helps us address big issues.

The data behind the new IPCC report

1 year, 8 months ago

Climate Change 2021: The Physical Science Basis - the Working Group I Contribution to the Intergovernmental Panel on Climate Change's Sixth Assessment Report - has been released today. This regular assessment provides policymakers with information on the physical science of climate change.

https://www.ceda.ac.uk/blog/the-data-behind-the-new-ipcc-report/

Image by Dominic Wunderlich from Pixabay

Take the first steps on your journey to Open Science

Useful checklists for you, the researcher

Your Digital Presence

Data Documentation and Citation Checklist

Software Documentation and Citation Checklist

Guidelines for your research team

Open Science Practices for Teams

Resources and Guides for Teams

Digital Objects Preservation Checklist for Teams

PARSEC

AGU
ADVANCING EARTH AND SPACE SCIENCE

Take the first steps on your journey to Open Science

Useful checklists for you, the researcher

Your Digital Presence

Data Documentation and Citation Checklist

Software Documentation and Citation Checklist

Guidelines for your research team

Open Science Practices for Teams

Resources and Guides for Teams

Digital Objects Preservation Checklist for Teams

PARSEC

AGU
ADVANCING EARTH AND SPACE SCIENCE

Activity

Discuss one advantage open science practices might have…

- *for your personal career*
- *for the outcomes of your work.*

Managing Your Digital Presence

Learn:

- ✓ Why should I manage my digital presence?

- ✓ What tools can I use to manage my digital presence?

AGU
ADVANCING EARTH AND SPACE SCIENCE

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.

AGU
ADVANCING EARTH AND SPACE SCIENCE

Who searches your Digital Presence?

- ✓ Researchers doing background or literature review
- ✓ Possible collaborators
- ✓ Your future advisor or employer
- ✓ Your institution
- ✓ Your funder
- ✓ Your society and associations
- ✓ Community Groups, Local Government Decision Makers, and on…

AGU
ADVANCING EARTH AND SPACE SCIENCE

Why Do I Care?

- Help others **discover you and your research!**
  - Increase **Citations** of your work
  - Connect with possible **Collaborators**
  - Connect to researchers in **Other Disciplines**
- Improve your **Open Science** practices by linking to your research

AGU
ADVANCING EARTH AND SPACE SCIENCE

First step. You.

The most important element of your work is YOU.

How do others find you?

Google

Mark Parsons

All | News | Images | Videos | Shopping | More | Settings | Tools

About 43,400,000 results (0.53 seconds)

https://en.wikipedia.org › wiki › Mark_Parsons

Mark Parsons - Wikipedia

Mark Parsons (born 8 August 1986) is an English-American soccer manager and currently the head coach of the Portland Thorns in the National Women's ...

Managerial career · Washington Spirit · Portland Thorns · Managerial statistics

https://twitter.com › mparsons_1

Mark Parsons (@mparsons_1) | Twitter

The latest Tweets from **Mark Parsons** (@mparsons_1). Head Coach of @ Thornsfc. Portland, OR.

https://www.youtube.com › channel

Mark Parsons YouTube Channel

Mark Parsons. Friends are those rare people who ask us to punch them and then wait to hear the answer. Chutney Glaze created by @Chutney Glaze Used by ...

Videos

WATCH | Mark Parsons takes questions from the media ahead ...

Portland Timbers
4 days ago

TRAINING | Mark Parsons talks preseason progress ...

Portland Timbers
1 week ago

Mark Parsons from the Thorns' draft room on the selection of ...

Portland Timbers

Mark Parsons

American soccer manager

Mark Parsons is an English-American soccer manager and currently the head coach of the Portland Thorns in the National Women's Soccer League. Wikipedia

Born: August 8, 1986 (age 34 years), England, United Kingdom

Teams coached: Portland Thorns FC (Association football manager, since 2015), MORE

Manage: Portland Thorns FC

Profiles

Twitter | LinkedIn

People also search for

View 1+ more


Paul Riley


Jim Gabarra


Alex Morgan


Verónica Boquete

AGU
ADVANCING EARTH AND SPACE SCIENCE

ORCID
Connecting research and researchers

SIGN IN/REGISTER English

mark parsons

ABOUT FOR RESEARCHERS MEMBERSHIP DOCUMENTATION RESOURCES NEWS & EVENTS

ADVANCED SEARCH

Showing **50** of **30485** results.

Items per page: 50 1 – 50 of 30485

ORCID ID	First Name	Last Name	Other Names	Affiliations
0000-0001-5613-0507	Mark	Parsons		Reed Eye Associates, Saint John Fisher College, Salus University Pennsylvania College of Optometry, Wilmington VA Medical Center
0000-0001-5490-226X	Mark	Parsons	Mark T Parsons, MT	Brigham Young University, The University of Arizona College of Medicine Phoenix
0000-0002-7723-0950	Mark	Parsons	Mark Parsons, MA Parsons, M. A. Parsons, Mark Alan	AScI Corporation, Coe-Truman Technologies, Cornell University, Data Science Journal, National Snow and Ice Data Center, Rensselaer Polytechnic Institute, The University of Alabama in Huntsville, University of Colorado Boulder, University of Edinburgh
0000-0003-4097-7468	Mark	Parsons		University of Edinburgh
0000-0001-8874-2487	Mark	Parsons		University of New South Wales
0000-0002-3474-1830	Mark	Parsons		

AGU ADVANCING EARTH AND SPACE SCIENCE

0000-0002-7723-0950

All Maps Shopping News Images More Settings Tools

About 27,800 results (0.43 seconds)

https://orcid.org › 0000-0002-7723-0950

Mark A. Parsons (0000-0002-7723-0950) - ORCID ...

Mark A. Parsons. ORCID iD. https://orcid.org/**0000-0002-7723-0950**.

https://rd-alliance.org › group › data-citation-wg › post

Fwd: Invitation - April 8 - Data Citation Community of Practice ...

Mar 25, 2021 — University of Alabama in Huntsville +1 303 941 9986. Mail: 1550 Linden Ave, Boulder, CO 80304, USA https://orcid.org/**0000-0002-7723-0950**.

https://www.force11.org › users › mark-parsons

Mark Parsons | FORCE11

Organization/Institution: Data Science Journal; ORCID ID: **0000-0002-7723-0950** . Edit. Badges. Help support the FORCE11 community with a donation ...

http://tw.rpi.edu › media › 2016/12/15 PPT

PowerPoint Presentation - Tetherless World Constellation

orcid.org/**0000-0002-7723-0950** & orcid.org/0000-0002-1009-7163, xsd:2012, http://dx.rpi.edu/10833/4199-5811-3221-0002-CC/?dc:title, jns:10.2481, ...

http://www.copdess.org › enabling-fair-data-project › ena...

Enabling FAIR Leadership – COPDESS

Mark A. Parsons, Rensselaer Polytechnic Institute, https://orcid.org/**0000-0002- 7723-0950**. Lesley Wyborn, National Computational Infrastructure ...

AGU
ADVANCING EARTH AND SPACE SCIENCE

https://orcid.org/
0000-0002-7723-0950

Websites & social links

LinkedIn

Other IDs

Scopus Author ID: 7202453217

ResearcherID: C-3037-2013

Keywords

Arctic, data management , data sharing, informatics, international collaboration, research data, geography, data curation, persistent identifiers

Countries

Is this you? Sign in to start editing

Printable version

Published name

Mark A. Parsons

Name
Mark Parsons

Also known as
Mark Parsons, MA Parsons, M. A. Parsons, Mark Alan Parsons

Biography

Mark A. Parsons is a Research Scientist and geographer at the University of Alabama in Huntsville working to help align data, software, and information standards and processes across NASA's science divisions. Mark has more than 25 years of experience in researching and developing data stewardship policies, practices, and systems. He has repeatedly and effectively built dynamic, functional collaborations across all sorts of differences in language and professional cultures. Mark was the first Secretary General of the Research Data Alliance. He has helped coordinated stewardship of a broad range of data from satellite remote sensing to Indigenous knowledge of Arctic change. He led the data management effort for the International Polar Year and helped establish the Exchange for Local Observations and Knowledge of the Arctic (ELOKA). His published work has guided national data policies and practice and has contributed to educational programs. Mark lives in Colorado and likes to ride bicycles, bake bread, and play outside.

Activities

Collapse all

Employment (8)

Sort

The University of Alabama in Huntsville: Boulder, CO, US

2020-12 to present | Research Scientist
Employment

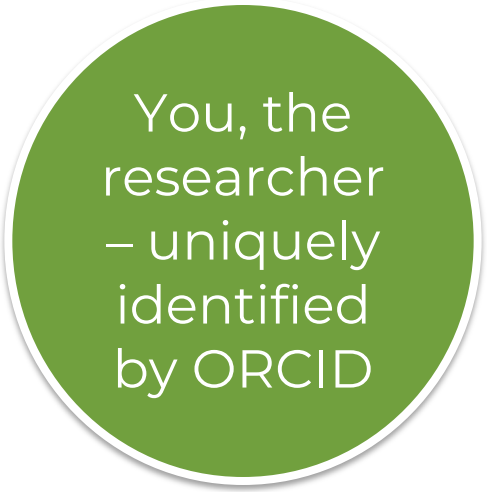
Show more detail

Source: Mark A. Parsons

First step. You.

Your ORCID

ORCID provides a persistent digital identifier that distinguishes you from other researchers and supports automated linkages between you and your research activities.

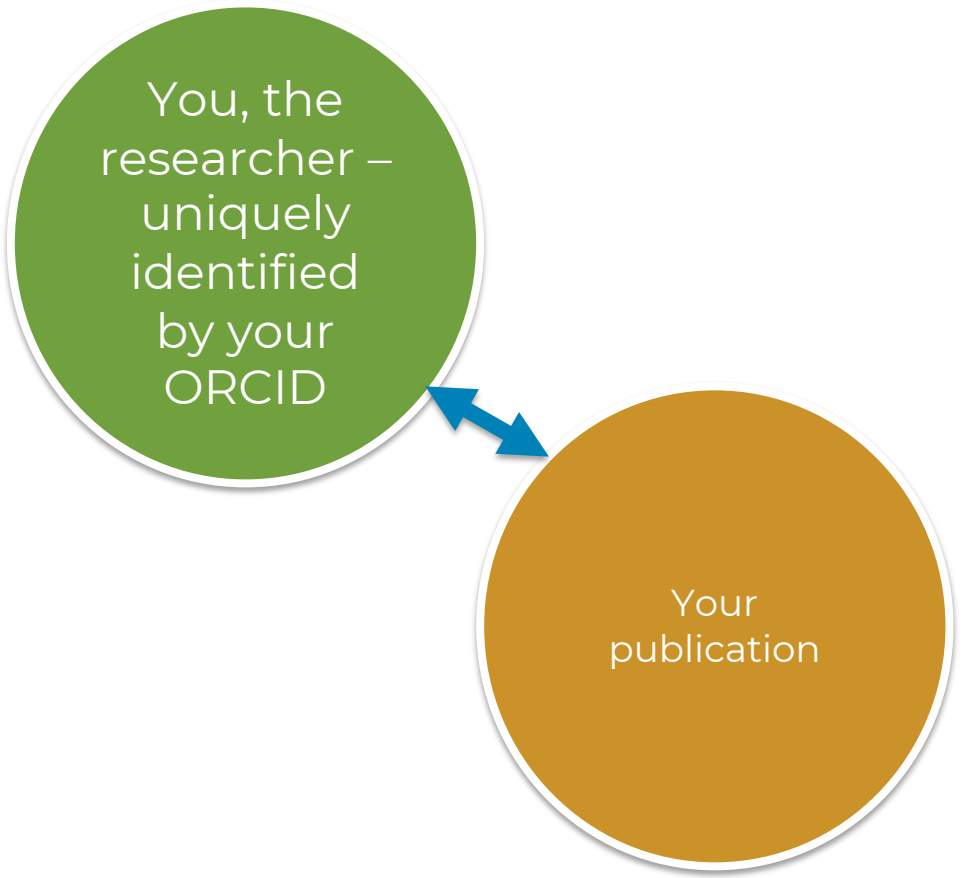


AGU
ADVANCING EARTH AND SPACE SCIENCE

Second step.
Your publications.

Your publications are paramount for your career.

They MUST be associated with you and your co-authors to receive the necessary credit.



AGU
ADVANCING EARTH AND SPACE SCIENCE

Quarterly
Founded: 2018
E-ISSN: 2641-435X

More About *Data Intelligence*

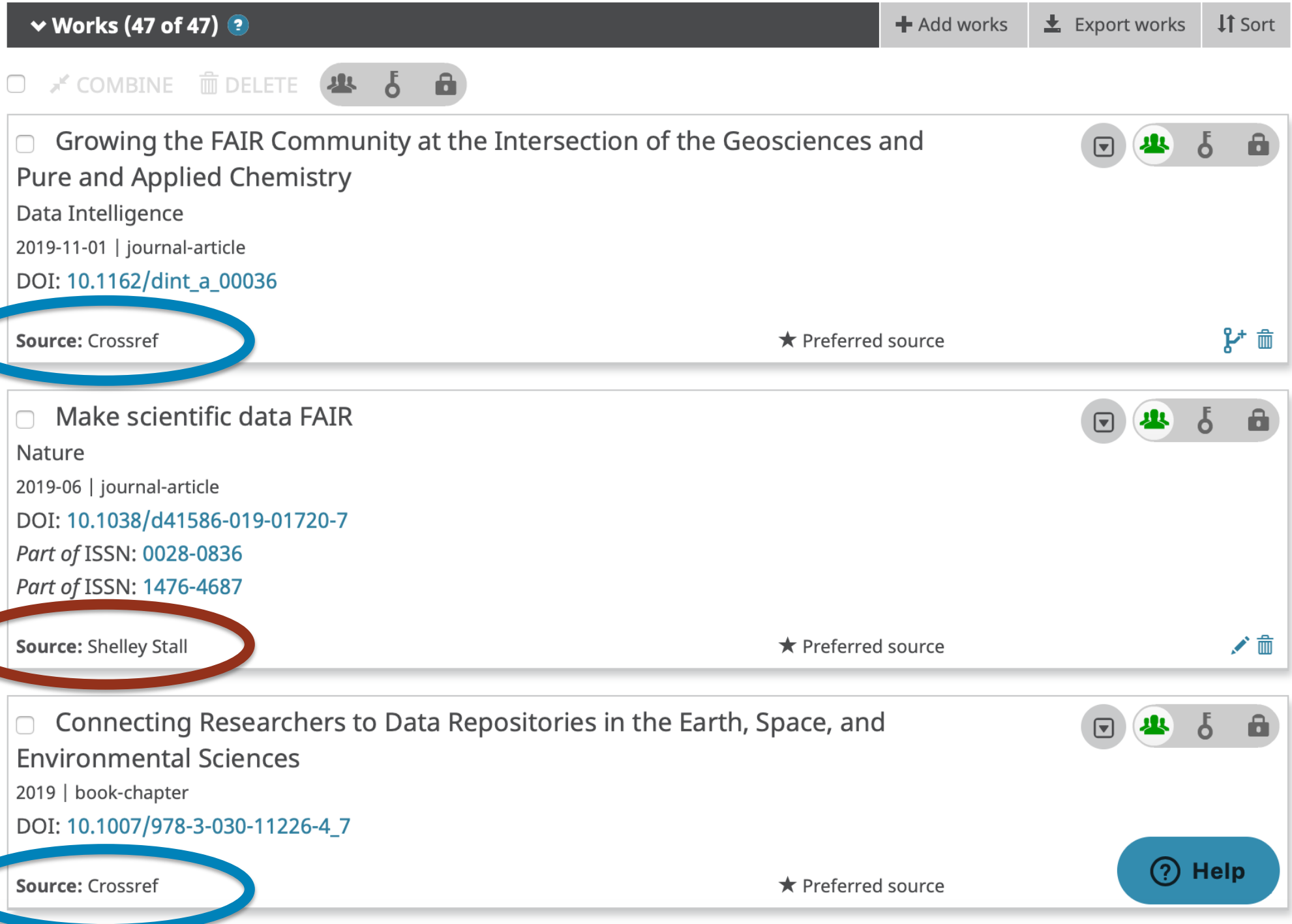
Journal Resources

Editorial Info
Abstracting and Indexing

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

Shelley Stall, Leah McEwen, Lesley Wyborn, Nancy Hoebelheinrich and Ian Bruno

Posted Online November 01, 2019
https://doi.org/10.1162/dint_a_00036


Works (47 of 47)
Add works
Export works
Sort
COMBINE
DELETE
Growing the FAIR Community at the Intersection of the Geosciences and Pure and Applied Chemistry
Data Intelligence
2019-11-01 | journal-article
DOI: 10.1162/dint_a_00036
Source: Crossref
★ Preferred source
Make scientific data FAIR
Nature
2019-06 | journal-article
DOI: 10.1038/d41586-019-01720-7
Part of ISSN: 0028-0836
Part of ISSN: 1476-4687
Source: Shelley Stall
★ Preferred source
Connecting Researchers to Data Repositories in the Earth, Space, and Environmental Sciences
2019 | book-chapter
DOI: 10.1007/978-3-030-11226-4_7
Source: Crossref
★ Preferred source
Help

Just published a paper, data, or software?

Create a trust relationship and allow Crossref, DataCite to update your ORCID automatically!

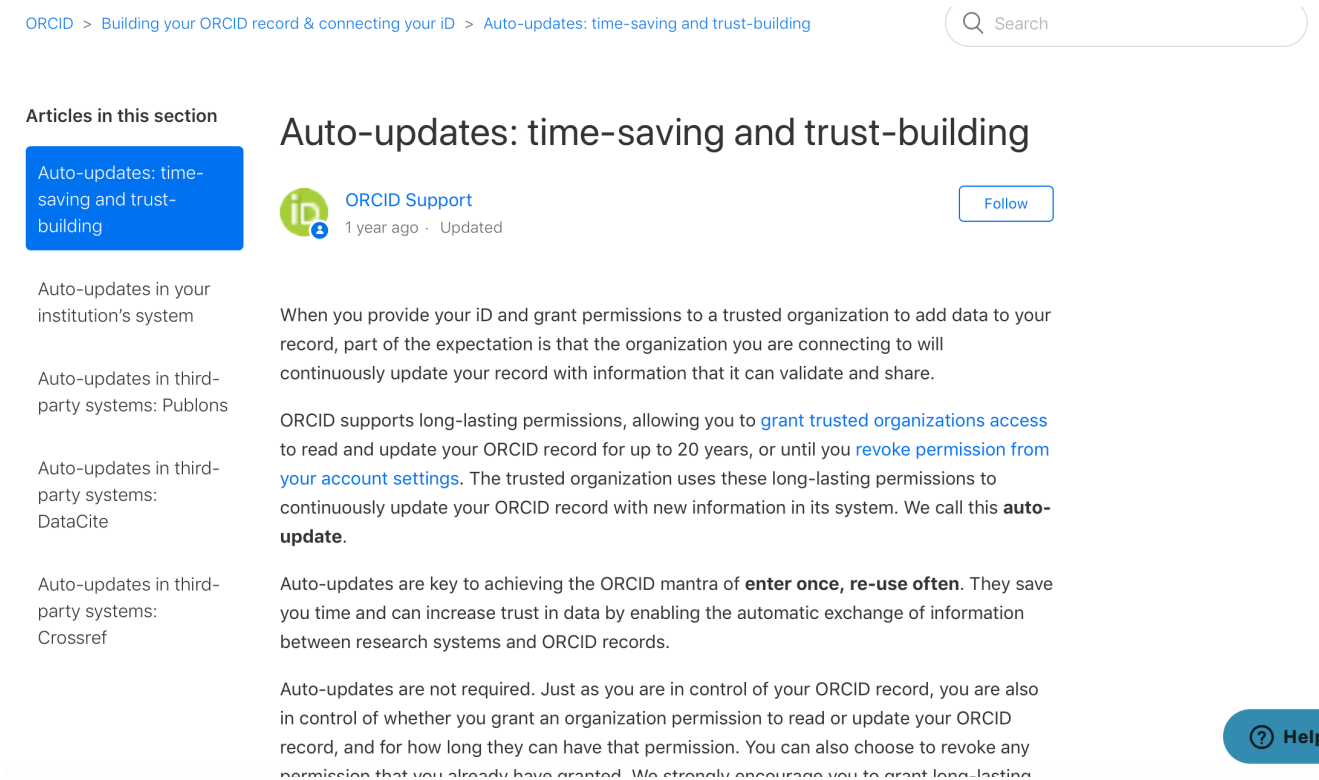
Establish a trust relationship between your ORCID profile and two of the primary DOI registry agencies.

Use this link for detailed steps:

https://data.agu.org/2021/09/21/digital-presence-orcid.html

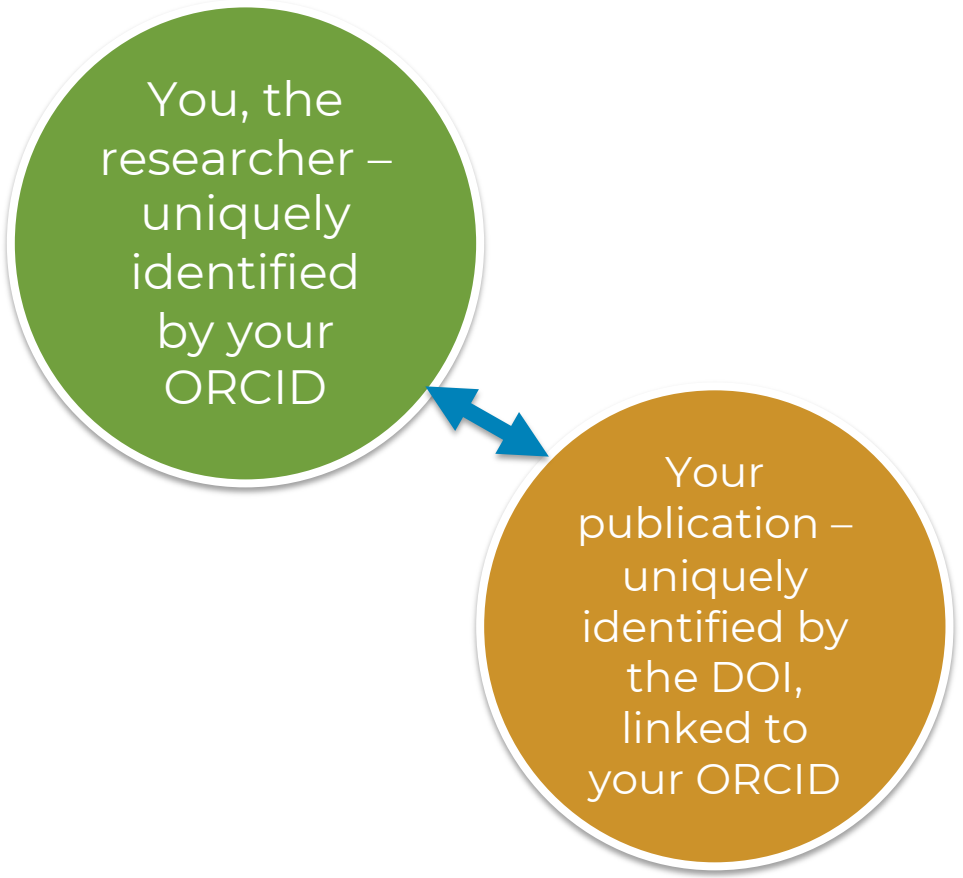
Do this

Do this



Second step.
Your publications.

Your publication's Digital Object Identifier (DOI) is most likely registered through **Crossref** by your journal when you publish.

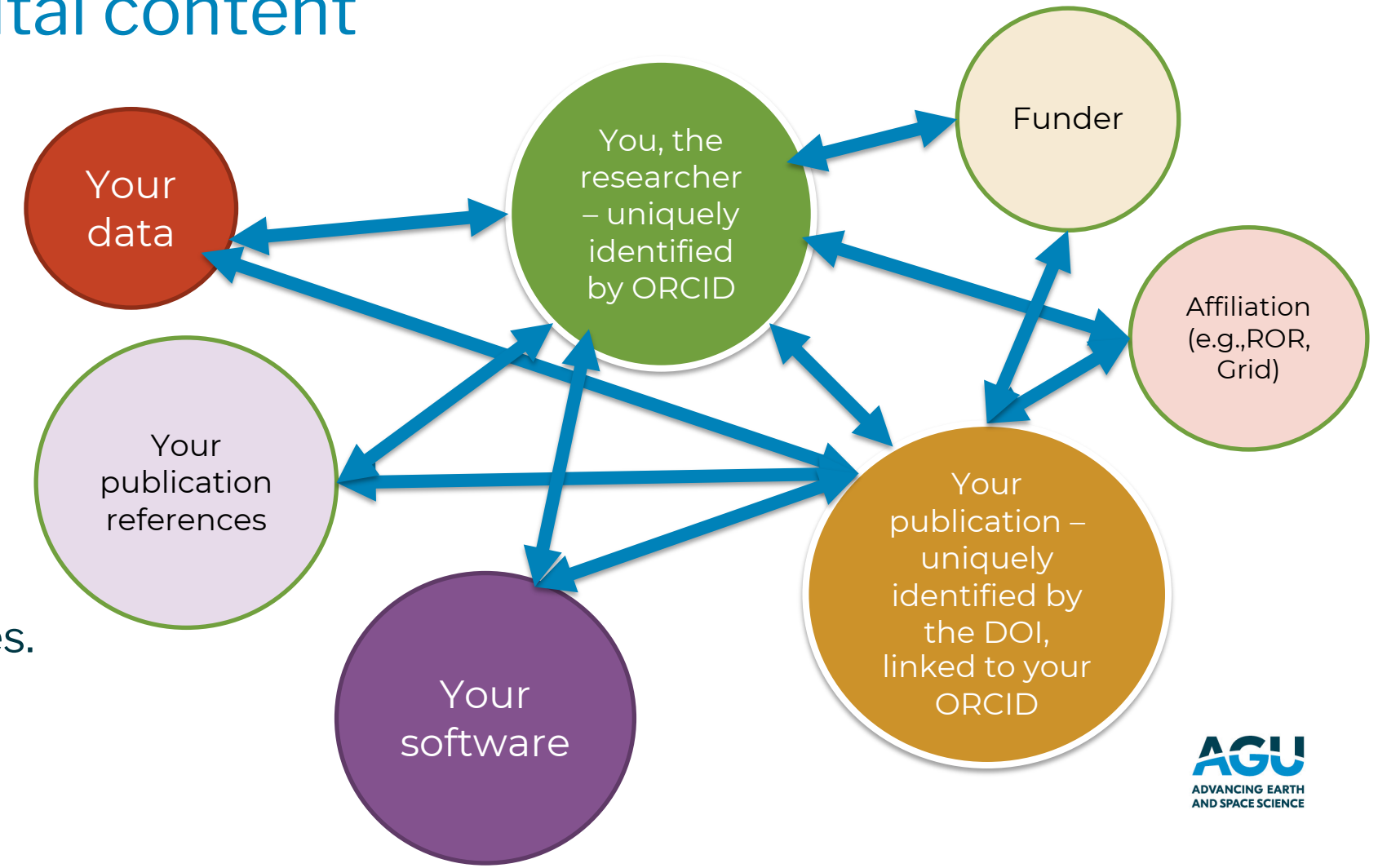


AGU
ADVANCING EARTH AND SPACE SCIENCE

Third step.
Building your digital content and connections.

Your publication needs to include **citations with the persistent identifier** to all related research objects.

Shown here are a few examples: data, software, other publication references.



AGU
ADVANCING EARTH AND SPACE SCIENCE

How can I work towards the best digital presence for my papers and research objects?

1. Optimize linking and discovery
2. Optimize understanding.
3. Optimize reuse.

AGU
ADVANCING EARTH AND SPACE SCIENCE

Optimize linking and discovery

Every research object you link…

1. …has your **ORCID**
2. …is **preserved** in an open, trusted repository that supports discovery
   - ✓ persistent identifier
   - ✓ machine readable
3. …includes **links to other related research objects**

AGU
ADVANCING EARTH AND SPACE SCIENCE

Optimize understanding

Every research object you link…

1. …is **well-documented** using metadata appropriate for the research object and community.
   - ✓ Include a README file (when using general repositories)
   - ✓ Include information that supports:
     - discovery,
     - understanding,
     - and reuse

*Note: **Metadata** is the "data about the data" – and might include information about the data format, method of collection, location, and more.*

AGU
ADVANCING EARTH AND SPACE SCIENCE

Optimize reuse and credit

Every research object you link…

1. …should have a **clear usage license**.
2. …should have a **recommended citation**.

Make sure your ORCID has trusted relationships with…

- CrossRef
- DataCite

…using the guidance found here:

https://data.agu.org/2021/09/21/digital-presence-orcid.html

AGU
ADVANCING EARTH AND SPACE SCIENCE

Other Scholarly Work to be Recognized, Linked, and Promoted in your ORCID profile

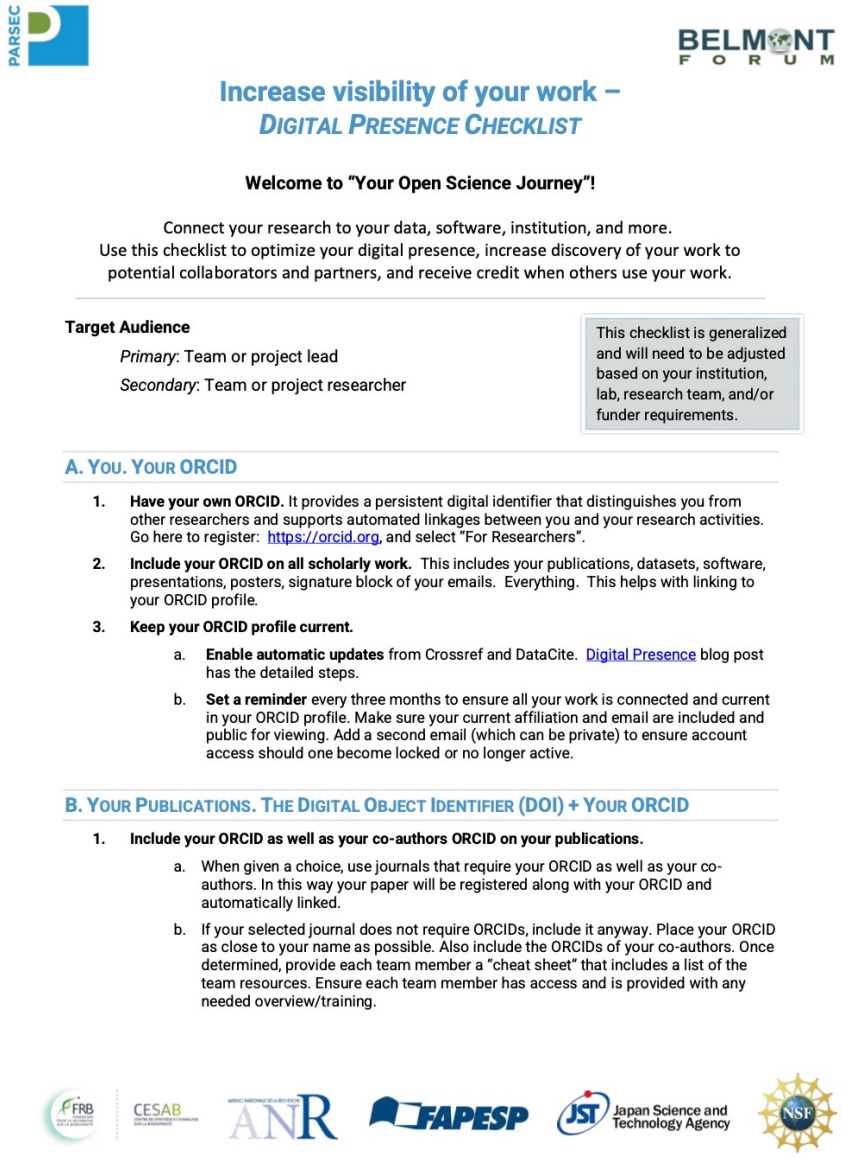
- Oral Presentation
- Poster
- Pre-Print
- Webinar/Workshop
- Society Volunteer Work
- Local Community Support
- International Standards Development

AGU
ADVANCING EARTH AND SPACE SCIENCE

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.



Increase visibility of your work – *Digital Presence Checklist*

Welcome to "Your Open Science Journey"!

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.

Target Audience

Primary: Team or project lead

Secondary: Team or project researcher

This checklist is generalized and will need to be adjusted based on your institution, lab, research team, and/or funder requirements.

A. You. Your ORCID

1. **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".
2. **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.
3. **Keep your ORCID profile current.**
   a. **Enable automatic updates** from Crossref and DataCite. Digital Presence blog post has the detailed steps.
   b. **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 or no longer active.

B. Your Publications. The Digital Object Identifier (DOI) + Your ORCID

1. **Include your ORCID as well as your co-authors ORCID on your publications.**
   a. When given a choice, use journals that require your ORCID as well as your co-authors. In this way your paper will be registered along with your ORCID and automatically linked.
   b. If your selected journal does not require ORCIDs, include it anyway. Place your ORCID as close to your name as possible. Also include the ORCIDs of your co-authors. Once determined, provide each team member a "cheat sheet" that includes a list of the team resources. Ensure each team member has access and is provided with any needed overview/training.

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.

Activity

Pair up – each of you will Google the other's name. What comes up? What *doesn't* come up?

Brainstorm a few of your own outputs or activities that you'd like to see highlighted in your digital presence.

How to Manage Your Data

AGU
ADVANCING EARTH AND SPACE SCIENCE

How to Manage Your Data: **During Your Research**

Planning and Documentation:

Data Management & Software Management Plans
E.g. Data Stewardship Wizard, DMP Online
Software Sustainability Institute

Organize Your Data
E.g. Dryad Good Data Practices,
The Carpentries (Tidy Data),
Earth Lab data science

DSW
DMP ONLINE
Software Sustainability Institute
Earth Lab
THE CARPENTRIES
CORNELL UNIVERSITY FOUNDED A.D. 1865
DRYAD
AGU ADVANCING EARTH AND SPACE SCIENCE

Roberts-Pierel, Ben; Davis, Eleanor; Rao, Yuhan (2021): A Graduate Student's Road Map to Data Management Training. ESIP. Educational resource. https://doi.org/10.6084/m9.figshare.14384456.v1

A Graduate Student's Road Map for Data Management Training

Benjamin Roberts-Pierel[1], Eleanor Davis Pierel[2], Yuhan Rao[3]

1. Oregon State University; 2. University of South Carolina; 3. North Carolina State University

** All authors were Earth Science Information Partners (ESIP) Community Fellows for 2018-2020.*

1. Mapping Data Lifecycle into Graduate School Road Map

- Plan - Coursework/Proposal writing
- Collect - Data collection
- Assure - Data collection
- Describe - Data collection
- Preserve - Defense & Publication of manuscripts
- Discover - Coursework/Proposal writing
- Integrate - Data analysis & Dissertation/Thesis Writing
- Analyze - Data analysis & Dissertation/Thesis Writing

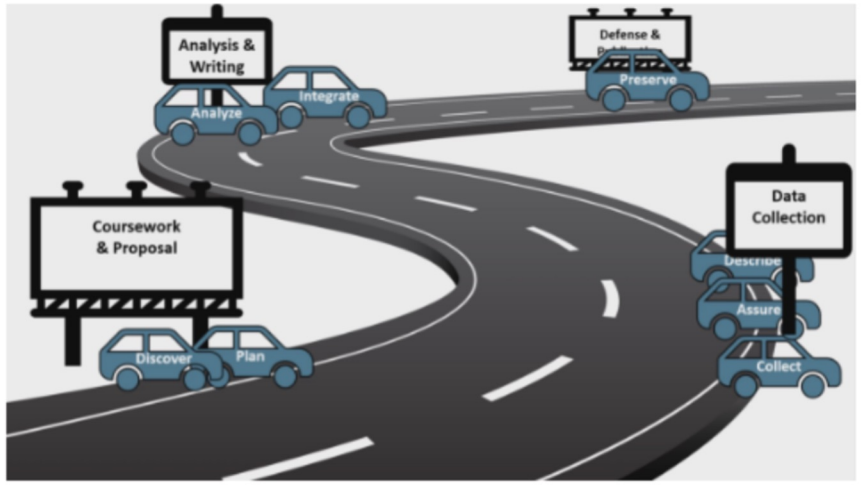


Figure 1. Mapping data lifecycle (DataONE) to different graduate school stages.

2. Your Graduate School Data Road Map

Congratulations on starting graduate school! It can often feel like a marathon and data

nature ecology & evolution

openscapes

Published: 23 May 2017

Our path to better science in less time using open data science tools

Julia S. Stewart Lowndes, Benjamin D. Best, Courtney Scarborough, Jamie C. Afflerbach, Melanie R. Frazier, Casey C. O'Hara, Ning Jiang & Benjamin S. Halpern

Nature Ecology & Evolution **1**, Article number: 0160 (2017) | Cite this article

https://doi.org/10.1038/s41559-017-0160 and https://www.openscapes.org/

Main Goal: Avoid parachuting researchers into your data/software and instead guide them as best as you can

Learning to get from A to B, one windy afternoon in the New Forest by Annie Spratt
Paragliding above the Chartreuse massif by Nicolas Tissot

How to Manage Your Data: **When Project is Complete**

Plan to deposit and preserve your created data/software in a **trusted, community-accepted repository** that supports citation.

AGU
ADVANCING EARTH AND SPACE SCIENCE

BELMONT FORUM

Repository Guidelines

PARSEC

Benefits of Storing Research Data in a Repository

There are many advantages to you as both a data producer and data user if you and your peers choose to preserve data in a repository. Of course, not all repositories are created equal, and these potential benefits are only realized by selecting a repository that does its job correctly, as described in the next section.

If you are a…

Data Producer/Depositor

- ✓ Your Data Management Plan is fulfilled (i.e., satisfies funders/Open Data requirements).
- ✓ The initial investment of collecting your data is preserved.
- ✓ You have the satisfaction that your data are being stewarded correctly and remain useful and meaningful.
- ✓ Your data are looked after long term, even if the data service discontinues.
- ✓ The ease of discovery of your data is increased.
- ✓ Publication, reuse or repurposing, and citation[1] is facilitated for your data.
- ✓ Recognized expertise is available to assist you with technicalities.
- ✓ It can be ensured that any necessary/wanted conditions on access and use, as well as licensing, are adhered to. (N.B. This is especially important for sensitive data.)

Data User

- ✓ You can easily discover data.
- ✓ You can easily understand your access and usage rights
- ✓ You can reuse/repurpose data without the costs of collection/production.
- ✓ You can verify (and thus build on) others results, accelerating scientific knowledge.
- ✓ You can cite peers, knowing that the data will still exist into the future.
- ✓ You have the satisfaction that the data are original/uncorrupted, and that any changes are recorded (provenance).
- ✓ (Re)Use of the data is made easier through full/appropriate metadata in an international or community standard.
- ✓ Ability to give feedback to the data producer/holder.

Edmunds, Rorie, Specht, Alison, Stall, Shelley, David, Romain, Mabile, Laurence, O'Brien, Margaret, Murayama, Yasuhiro, Correa, Pedro, Machicao, Jeaneth, & Miyairi, Nobuko. (2022). Repository Guidelines. Zenodo.

Guide including steps to consider when choosing a trusted, community-accepted repository:

https://doi.org/10.5281/zenodo.6542494

AGU
ADVANCING EARTH AND SPACE SCIENCE

Domain-Discipline Repositories Useful to AGU Journals

The data that supports the research reported in your paper must be deposited in a community-accepted, trusted preservation repository. Additionally, authors should make available software that has a significant impact on the research. A repository that specializes in dommain-discipline specific data and software will maximize the probability that the deposited data and software will be findable, accessible, interoperable and reusable (FAIR). Repsoitories that use persistent identifier links (e.g. DOI or digital object identifier over URLs (and not to the home page) are recommended. Note, an English language translation is necessary in order for the data/software to be accessible to the wider community. Domain-discipline repositories useful to AGU journals below may also be at different stages in supporting the FAIR principles. For any additional domain-discipline repositories recommendations, contact datahelp@agu.org or submit a GitHub issue/pull request. Otherwise, look to your institutional repository, your computing center, a general repository (e.g., Zenodo, Dryad, figshare), or search for a repository using re3data, OpenAire, or DataOne. Consult Data and Software for Authors and Data and Software Sharing Guidance for Authors Submitting to AGU Journals for more in-depth guidance.

https://data.agu.org/resources/useful-domain-repositories

Generalist Repository Comparison Chart

doi: 10.5281/zenodo.3946719

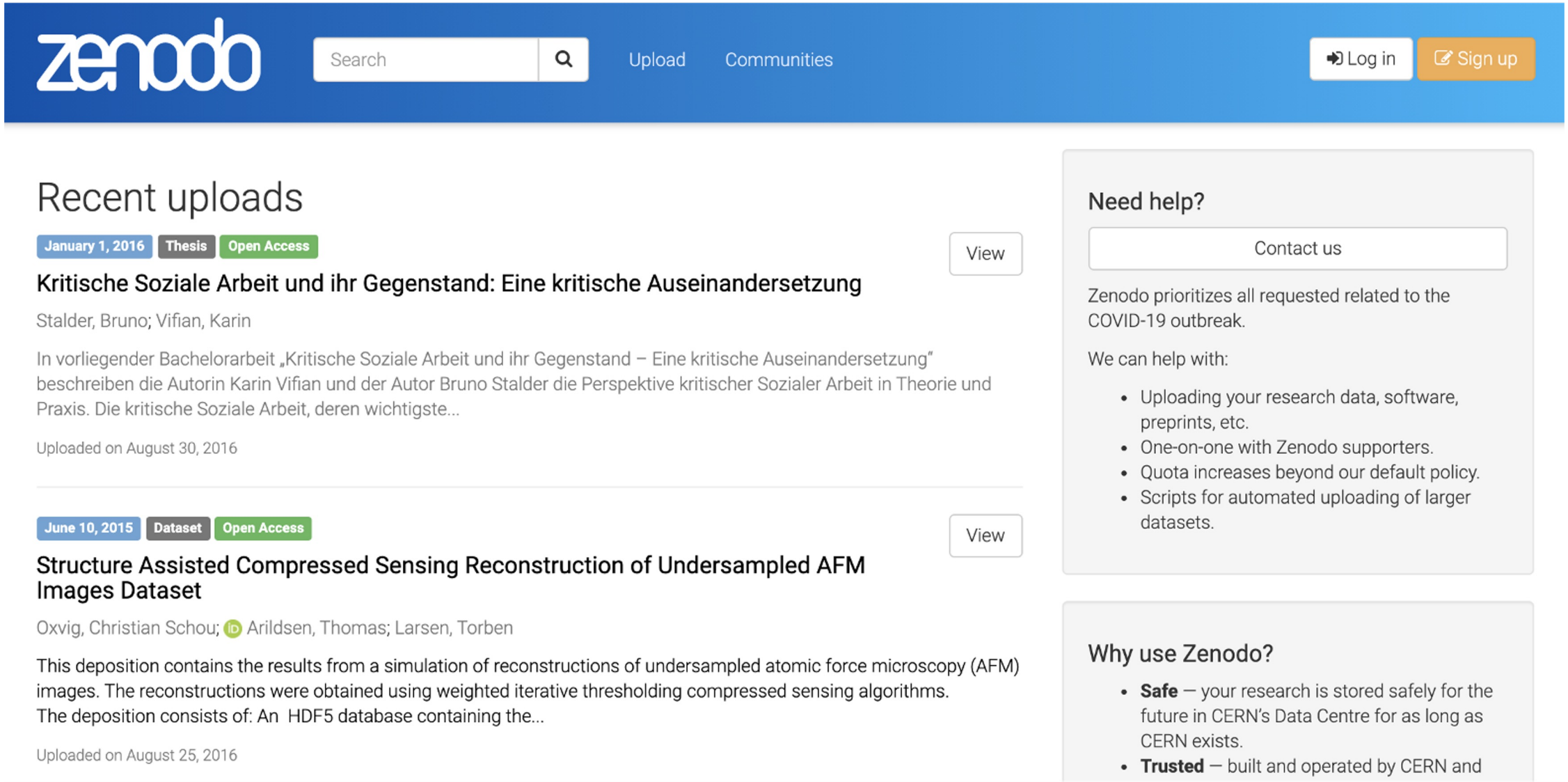
This chart is designed to assist researchers in finding a generalist repository should no domain repository be available to preserve their research data. Generalist repositories accept data regardless of data type, format, content, or disciplinary focus. For this chart, we included a repository available to all researchers specific to clinical trials (Vivli) to bring awareness to those in this field.

https://fairsharing.org/collection/GeneralRepositoryComparison

AGU
ADVANCING EARTH AND SPACE SCIENCE

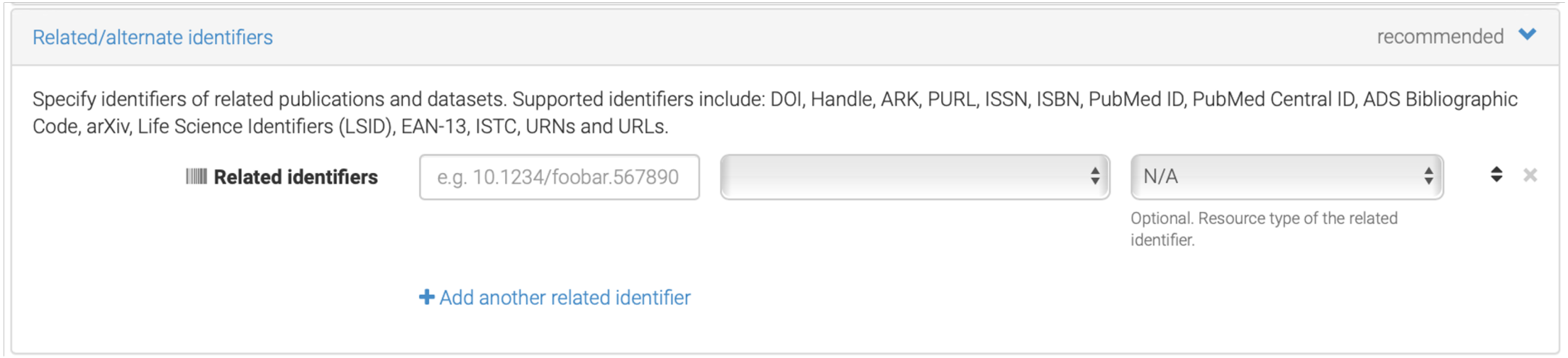
TOPIC	HARVARD DATAVERSE REPOSITORY	DRYAD	FIGSHARE	MENDELEY DATA	OSF	VIVLI	ZENODO
Brief Description	Harvard Dataverse Repository is a free data repository open to all researchers from any discipline, both inside and outside of the Harvard community, where you can share, archive, cite, access, and explore research data.	Open-source, community-led data curation, publishing, and preservation platform for CC0 publicly available research data Dryad is an independent non-profit that works directly with: • researchers to publish datasets utlizing best practices for discovery and reuse • publishers to support the integration of data availability statements and data citations into their workflows • institutions to enable scalable campus support for research data managment best practices at low cost	A Data publishing platform for all researchers. Some of figshare's Core Beliefs: Academic research outputs should be as open as possible, as closed as necessary Academic research outputs should never be behind a paywall Academic research outputs should be human and machine readable/ query-able Academic infrastructure should be interchangeable Academic researchers should never have to put the same information into multiple systems at the same institution Identifiers for everything The impact of research is independent of where it is published and what type of output it is Figshare+ (https://plus.figshare.com/) supports sharing larger datasets.	Mendeley Data is a free repository specialized for research data. Search more than 20+ million datasets indexed from 1000s of data repositories and collect and share datasets with the research community following the FAIR data principles.	OSF is a free and open source project management tool that supports researchers throughout their entire project lifecycle in open science best practices.	Vivli is an independent, non-profit organization that has developed a global data-sharing and analytics platform. Our focus is on sharing individual participant-level data from completed clinical trials to serve the international research community.	Powering Open Science, built on Open Source. Built by reserachers for researchers. Run from the CERN data centre, whose purpose is long term preservation for the High Energy Physics discipline, one of the largest scientific datasets in the world
Size limits	No byte size limit per dataset. Harvard Dataverse Repository currently sets a file size limit of 2.5GB.	300GB/dataset	20GB for free figshare.com accounts. Figshare+ offers storage in tiers beginning at 100GB up to 10TB+ per dataset. System limit of 5TB/file.	10GB per dataset	Projects currently have not storage limit. There is a 5GB/file upload limit for native OSF Storage. There is no limit imposed by OSF for the amount of storage used across add-ons connected to a given project.	If more than 1TB of study data, reach out to us at support@vivli.org so we can help transfer your data.	50GB per dataset, contact us via https://zenodo.org/support for higher limits

Stall, Shelley, Martone, Maryann E., Chandramouliswaran, Ishwar, Crosas, Mercè, Federer, Lisa, Gautier, Julian, … Zigoni, Alberto. (2022). Generalist Repository Comparison Chart. Zenodo. http://doi.org/10.5281/zenodo.3946719



Zenodo, a commonly used general repository, features a sandbox where you can test how your data/software will be deposited and described.
https://sandbox.zenodo.org/

For instance, you can test in the Sandbox how to link your data/software to your paper



See Author Carpentry > Related Indentifiers
https://authorcarpentry.github.io/dois-citation-data/01-register-doi.html

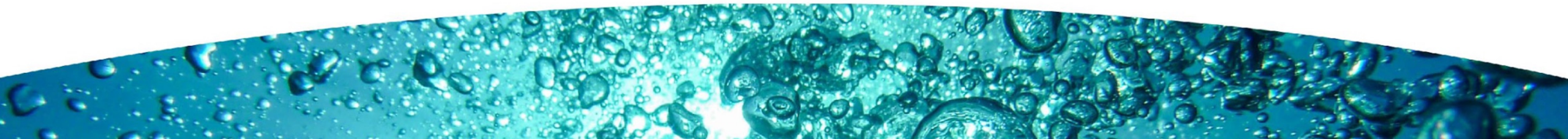
How to Manage Your Data: **At Publication**

Plan to properly **cite your data (and software)** in your paper and write an **Availability Statement** that includes needed information to find and use your data (and software).

AGU
ADVANCING EARTH AND SPACE SCIENCE

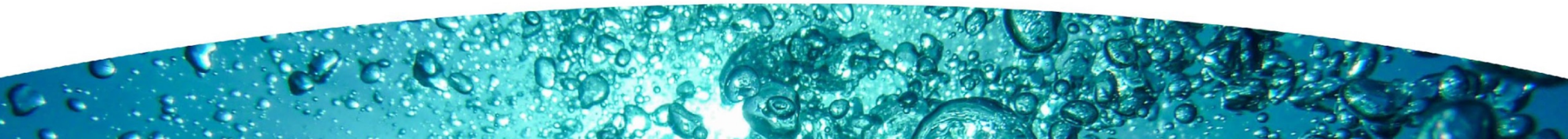
What is included in an Availability Statement?

1. A brief description of the type(s) of data or software
2. Repository Name(s) where they are deposited
3. Version (of software)
4. DOI, Persistent Identifier Link to Data or Software (and Identifier)
5. Link to publicly accessible development platform (in the case of Software, e.g. GitHub)
6. Access Conditions (e.g. if Registration is Required)
7. Licensing/Permissions (e.g. Creative Commons Attribution)
8. In-text citation in References (optional)



Sample Availability Templates

- The [type of data] data used for [brief context, description] in the study are available at [repository, source name] via [DOI, persistent identifier link] with [license, access conditions] [optional in-text citation in References]
- [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], available via [license type, access conditions] and developed openly at [software development platform link].* [optional in-text citation in References]

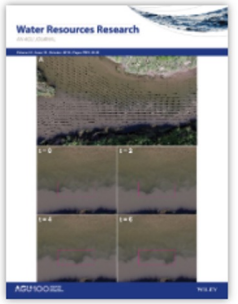


Data Availability Statement Example

Water Resources Research

Research Article | Free Access

Volume 54, Issue 10
October 2018
Pages 7619-7630



Flow, Flux, and Feeding in Freshwater Mussels

Rakesh Mistry, Josef D. Ackerman

First published: 19 September 2018 | https://doi.org/10.1029/2018WR023112 | Citations: 9

SECTIONS PDF TOOLS SHARE

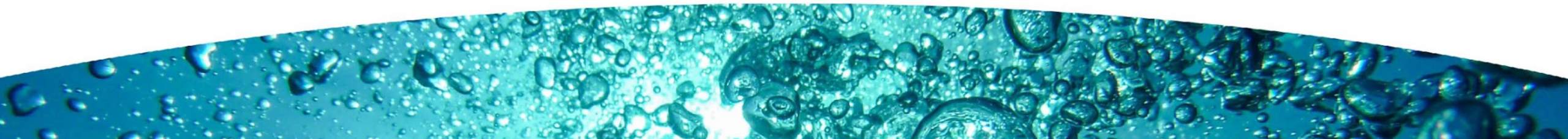
Abstract

Unionid mussels are important constituents of aquatic systems that are affected by anthropogenic changes in hydrology and concomitant increases in suspended solids, yet little is known about the effects of flow on their suspension feeding. We examined the clearance rates (*CRs*) of four species of freshwater mussels (*Lampsilis siliquoidea*, *Lampsilis fasciola*, *Ligumia nasuta*, and *Villosa iris*) to determine whether they feed selectively on river seston and how this may vary with algal flux (concentration × velocity). The *CR* for

https://doi.org/10.1029/2018WR023112

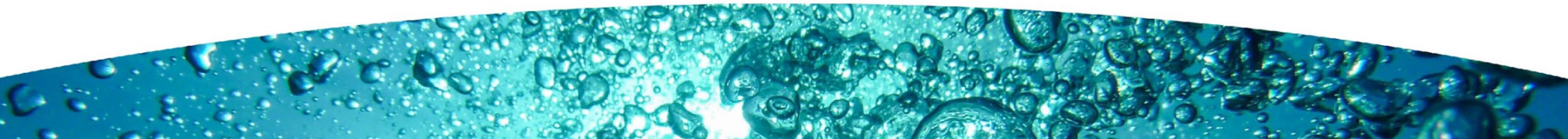
Open Research Section:
Data Availability Statement Example

The authors are grateful for the support provided by Colette Mesher, Mike Nishizaki, Shaylah Tuttle-Raycraft, Aidin Jabbari, and the research assistants of the Ackerman lab, Matt Cornish and Mike Davies in the Hagen Aqualab, Steve Wilson and Ian Moore in the Physics Workshop. Funding for this research was provided by grants from the Ontario Ministry of Natural Resources and Forestry (Species at Risk Research Fund Ontario and Species at Risk Stewardship Fund), Fisheries and Oceans Canada (SARCEP), and the Natural Sciences and Engineering Research Council of Canada (Discovery) to J.D.A. Data reported in this paper can be accessed in the Dryad Digital Repository (https://doi.org/10.5061/dryad.v18jj97; Mistry & Ackerman, 2018).



Data Citation Example in the References Section (for Credit)

Mistry, R., & Ackerman, J. D. (2018). Data from: Flow, flux and feeding in freshwater mussels. *Water Resources Research*, Dryad Digital Repository, https://doi.org/10.5061/dryad.v18jj97 , https://doi.org/10.1029/2018WR023112
Wiley Online Library | Web of Science® | Google Scholar



Data from: Flow, flux and feeding in freshwater mussels

Mistry, Rakesh, University of Guelph
Ackerman, Josef D., University of Guelph

Publication date: September 19, 2019
Publisher: Dryad
https://doi.org/10.5061/dryad.v18jj97

Data Creators

Persistent ID needed for Citation

Citation

Mistry, Rakesh; Ackerman, Josef D. (2019), Data from: Flow, flu
https://doi.org/10.5061/dryad.v18jj97

Recommended Citation to use in you paper

Abstract

Unionid mussels are important constituents of aquatic systems that are affected by anthropogenic changes in hydrology and concomitant increases in suspended solids, yet little is known about the effects of flow on their suspension feeding. We examined the clearance rates (CR) of four species of freshwater mussels (Lampsilis siliquoidea, Lampsilis fasciola, Ligumia nasuta and Villosa iris) to determin
this may vary with algal flux (concentration × velocity). The CR
seston particle size, particle fluorescence, and algal taxon. The CR of all species increased linearly with flow chamber velocity, but exhibited saturation-like kinetics with increasing algal flux. The CR of Lampsilis species were higher for larger (>10 um) vs. smaller (<10 um) particles, the latter of which were numerically dominant in river seston. The CR of Lampsilis mussels on most of the algal taxa declined (linearly or non-linearly) with algal flux indicating that mussels

Abstract Describing Data

Data Files

Download dataset

September 19, 2019

Related Works

Article

https://doi.org/10.1029/2018wr023112

Metrics

30 views

2 downloads

1 citations

Activity

Write a “Data Availability Statement” for data that you’ve generated or used in the past. Your Data Availability Statement should include…

- Who was involved in data production (individuals, teams, organizations)
- Where the data can be found
- How the data was collected, or how to find out more information about the data collection
- How the data can be reused (license? Permission needed?)
- What else might someone need to know before reusing your dataset?

How to Manage Your Software

When should you preserve and share your software?

AGU
ADVANCING EARTH
AND SPACE SCIENCE

How to Manage Your Software: **During Your Research**

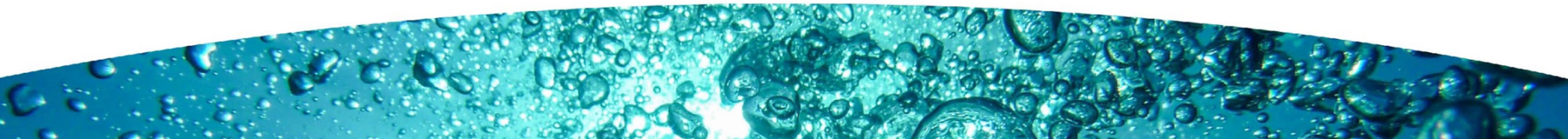
Planning and Documentation:

Data Management & Software Management Plans
E.g. Data Stewardship Wizard, DMP Online
Software Sustainability Institute

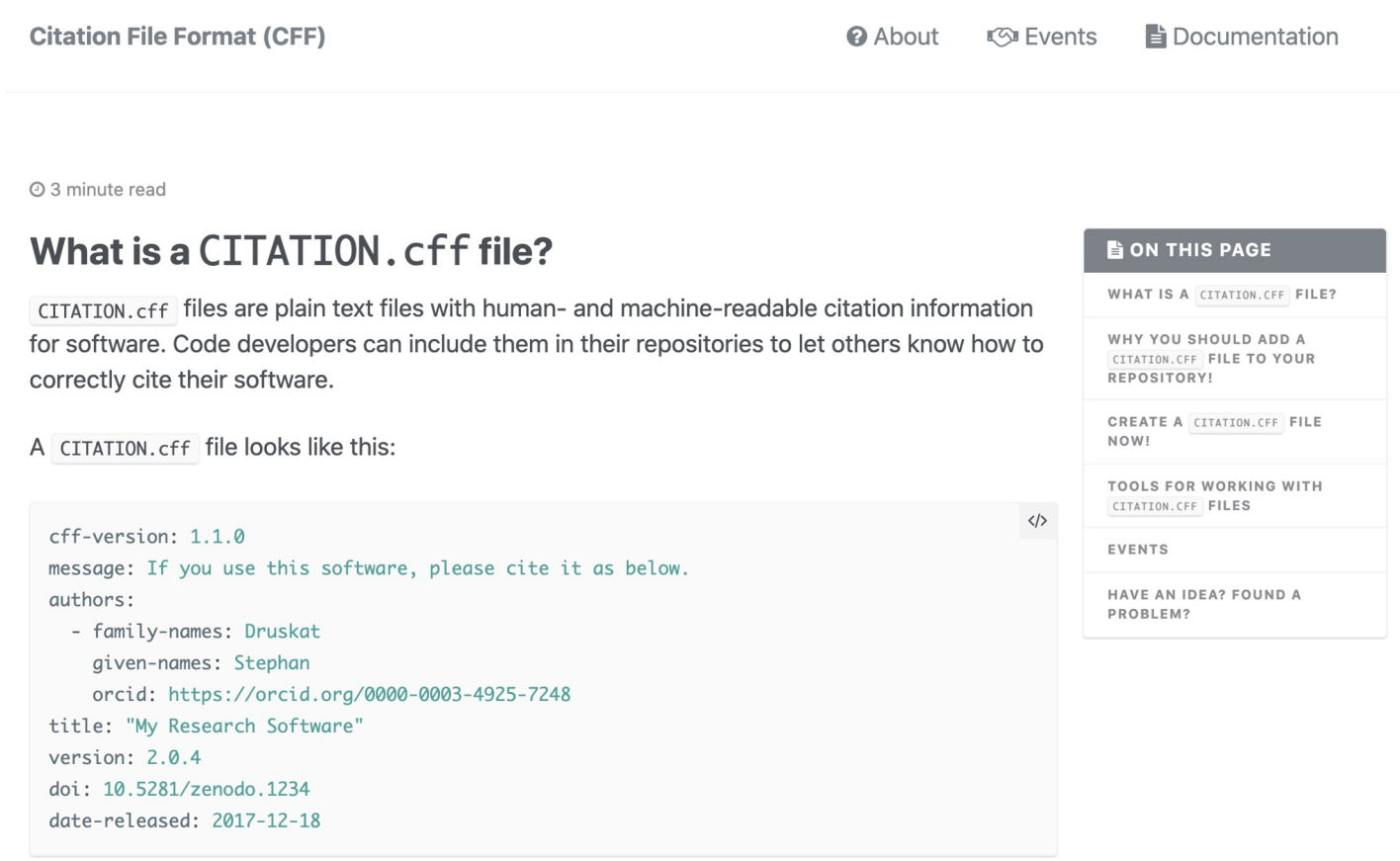
Keep Track of Versions and Work Collaboratively
Use platforms like GitHub to manage your software and code development efficiently, then link with Zenodo to publish with a DOI when you're ready.

Software Citation Examples

- Lab for Exosphere and Near Space Environment Studies. (2019, March 20). lenses-lab/LYAO_RT-2018JA026426: Original Release (Version 1.0.0) [Software]. Zenodo. http://doi.org/10.5281/zenodo.2598836
- Bell, S. W. (2020). samwbell/saturn_counts: April 26, 2020 Release (Version 1.1.0) [Software]. Zenodo. https://doi.org/10.5281/ZENODO.3766959
- Shaoqian Hu. (2019, December 25). Direct surface wave radial anisotropy tomography package (Version 1.0) [Software]. Zenodo. http://doi.org/10.5281/zenodo.3592528



The importance of citing your software on GitHub and other places…



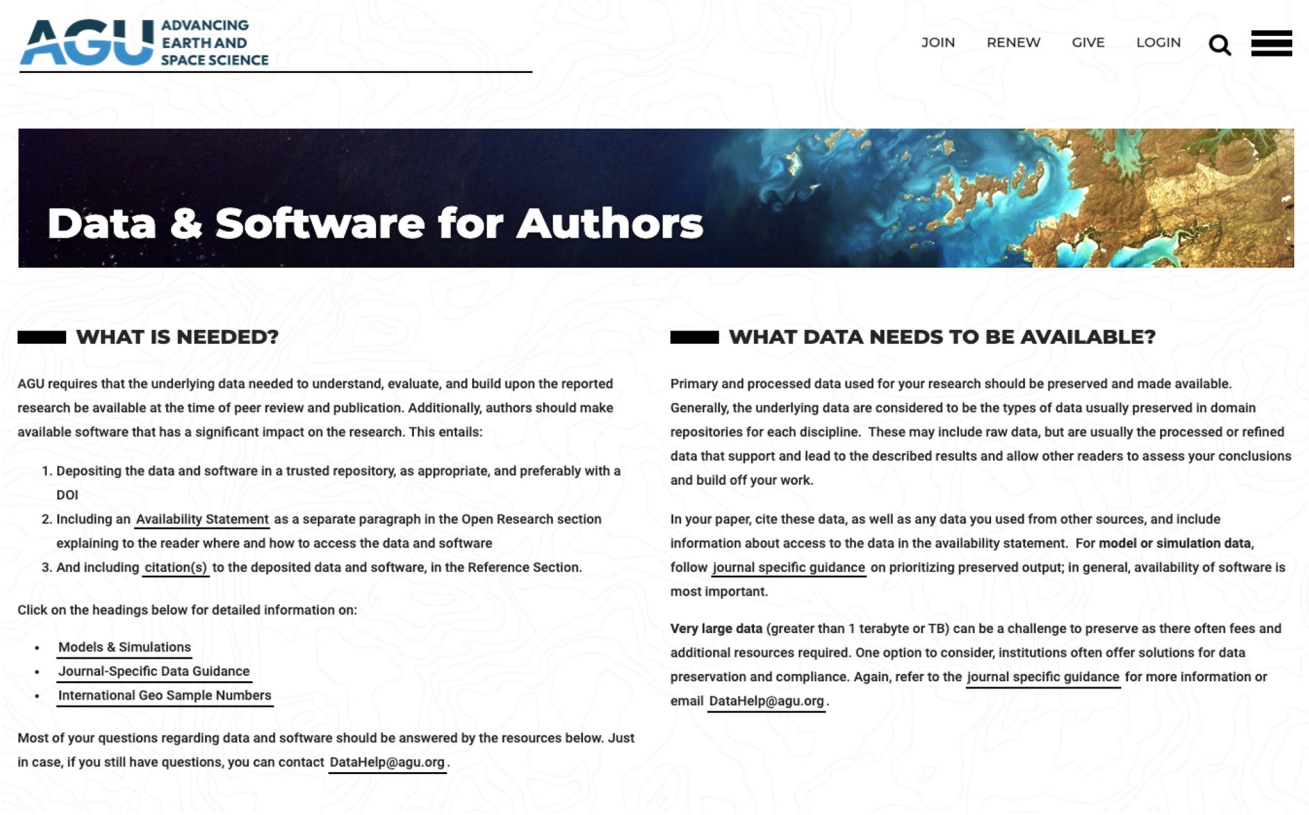
https://citation-file-format.github.io/#/what-is-a-citation-cff-file

AGU
ADVANCING EARTH AND SPACE SCIENCE

AGU Data & Software Sharing Guidance

What's 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 (datahelp@agu.org)



AGU ADVANCING EARTH AND SPACE SCIENCE

JOIN RENEW GIVE LOGIN

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

AGU ADVANCING EARTH AND SPACE SCIENCE

Activity

Write a “Software Availability Statement” for data that you’ve generated or used in the past. Your Software Availability Statement should include…

- Who helped with your code or where your software is from (individuals, teams, organizations)
- Where the software version you used can be found
- How the software was used and under what conditions (e.g. software version, PC vs Mac vs other, etc)
- How the software can be reused (license? Permission needed?)
- What else might someone need to know before reusing your software?

THANK YOU

For Resources and Information About Open Science at AGU:
https://data.agu.org/

Contact:
Shelley Stall, sstall@agu.org
Kristina Vrouwenvelder, kvrouwenvelder@agu.org