

PIDs: the view from AGU

*Kristina Vrouwenfelder, Ph. D.
Program Manager, Open Science*

23 journals

60,000 members across 137
countries

Annual meeting of ~25k
researchers

Nonprofit scientific society governed by Board of
Directors and Council.



First:

A short story about data and software citations.

The Data Citation Journey at AGU:

AGU and COPDESS identify FAIR data issues in ESS...

2021: Issue with machine-actionable data and software citations in AGU papers discovered.

2023: All AGU journals entered in Data Citation Pilot, enabling full author guidance and compliance checks for data and software citations.

2020: AGU and partners awarded NSF Public Access to Research 2.0 grant to enable Open and FAIR data practices across ESS.

Jan 2022: First AGU journal (*GeoHealth*) enters Data Citation Pilot.

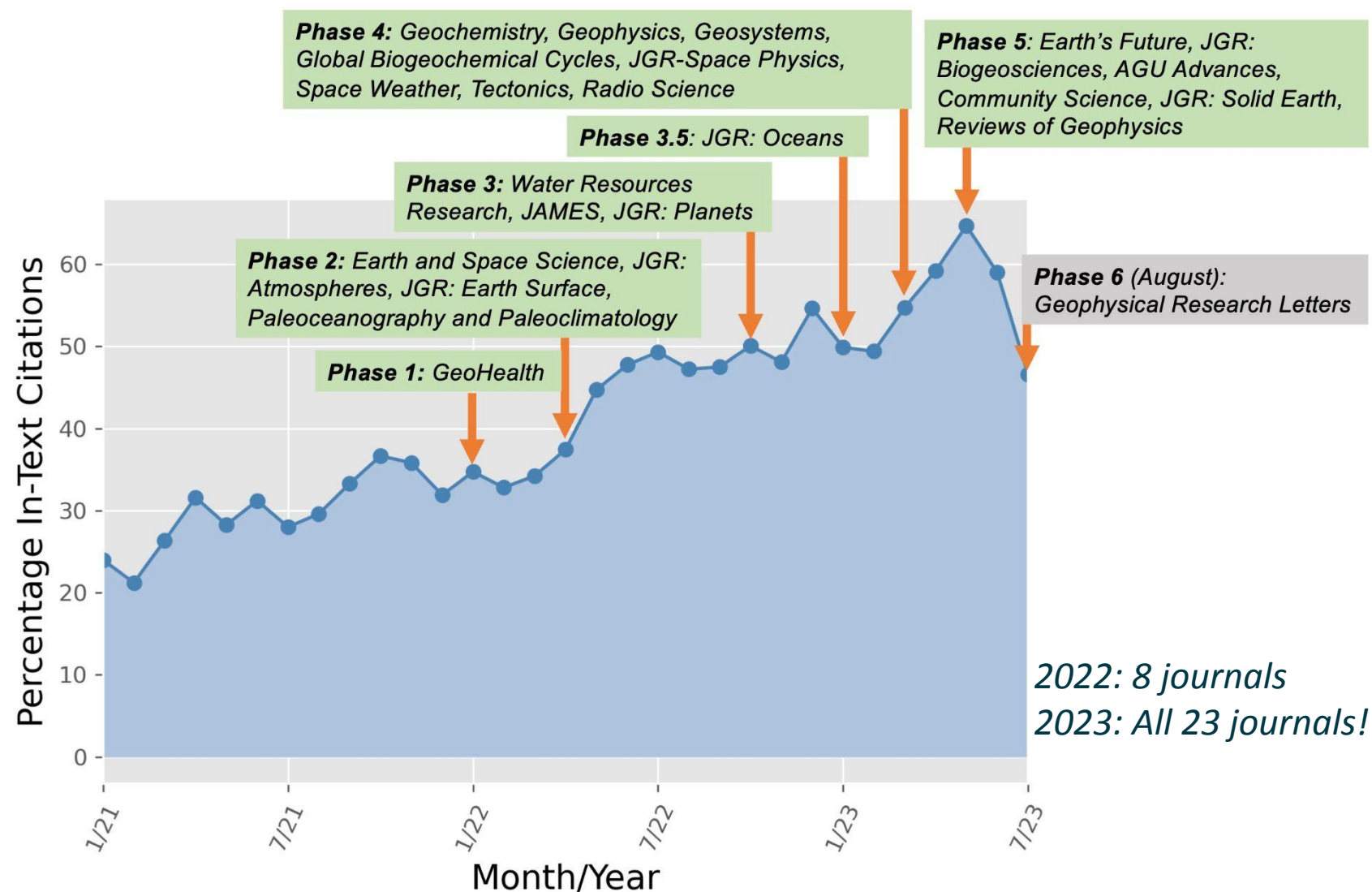
May 2022: Data and software citation issue resolved.

DOI: 10.1038/s41597-023-02491-7



Data Citation Pilot and Author Guidance at AGU:

- 1) Data and/or software are deposited in a community-accepted, trusted repository, as appropriate, and preferably with a DOI,
- 2) Availability Statement is included as a separate paragraph in the Open Research section explaining to the reader where and how to access the data and software,
- 3) Citation(s) to the deposited data and software are included in the Reference Section.



Pieces of the (Data Citation) Puzzle:

Measurable Compliance

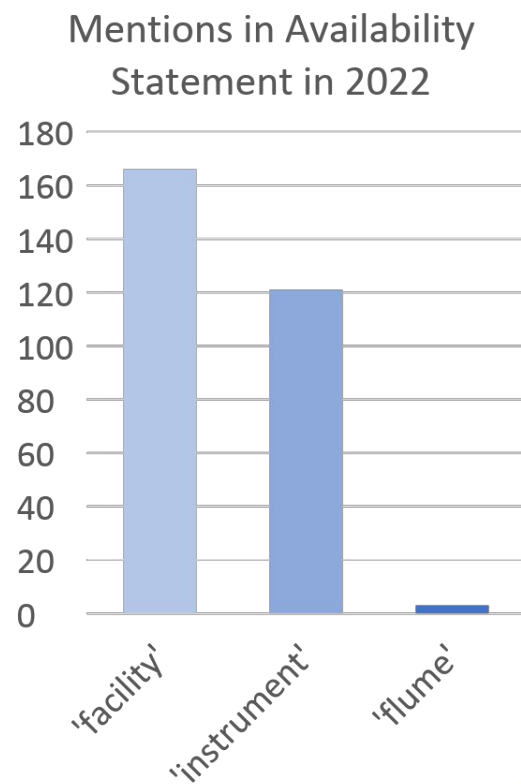
Author Guidance and Publication
Policy

Technology

Education and Outreach

**Data and Software Citation at
AGU!**

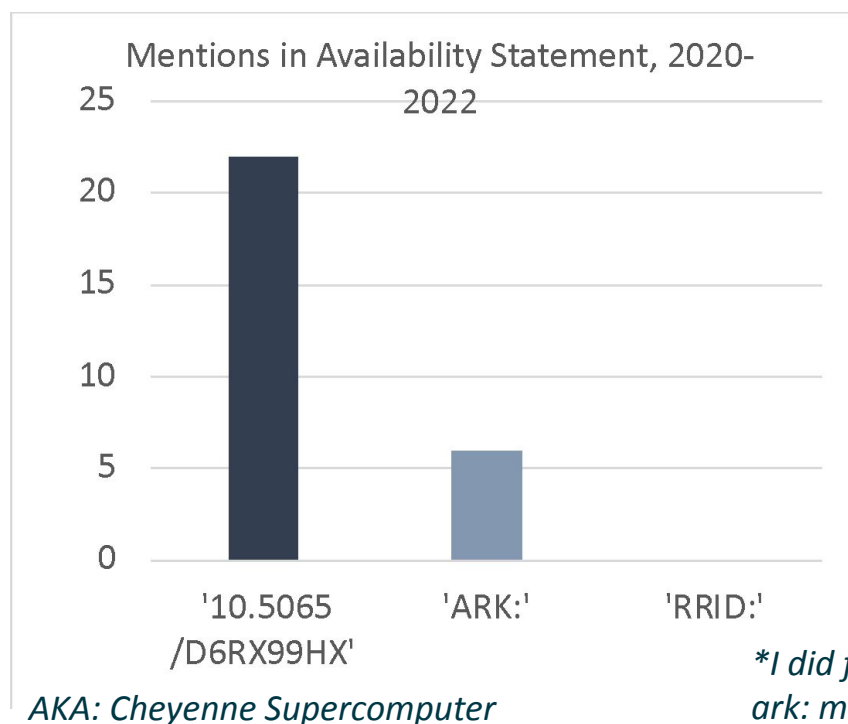
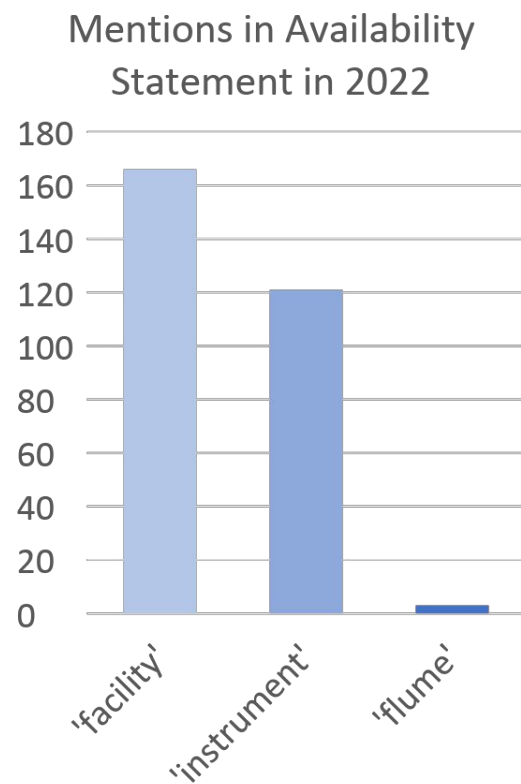
What about PIDs for Instruments and Facilities?



For context: ~8000 papers in AGU journals in 2022; > 70% include a DOI in the Availability Statement.

Church, M., Dudill, A., Venditti, J. G., & Frey, P. (2020). Are results in geomorphology reproducible?. *Journal of Geophysical Research: Earth Surface*, 125, e2020JF005553. [https://doi.org/ 10.1029/2020JF005553](https://doi.org/10.1029/2020JF005553)

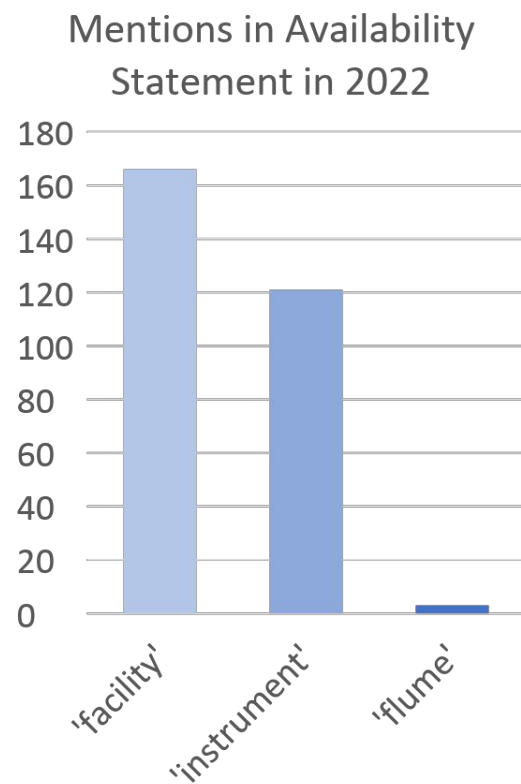
What about PIDs for Instruments and Facilities?



**I did find 2 RRID: mentions and ~30ish ark: mentions in the acknowledgements section, same time period*



What about PIDs for Instruments and Facilities?



Open Research

Data Availability Statement

This work made use of the MESSENGER E/V/H XRS CALIBRATED (CDR) SPECTRA v1.0 data set (Starr, [2010](#)) archived by the NASA Planetary Data System at the PDS Geosciences Node, Washington University, St Louis, MO, USA. This work used the ALICE High Performance Computing Facility at the University of Leicester.

DOI: [10.1029/2021JA029675](https://doi.org/10.1029/2021JA029675)



What about PIDs for Instruments and Facilities?

Open Research

Data Availability Statement

The data analyzed for this paper are available as Marshall et al. (2022). Data loggers used in flume experiments are Onset Computing HOBO U-24 fluid electrical conductivity loggers and HOBOWare Pro version 3.7.8 (Onset Computing Onset Computer Corporation, 2022) was used with a software license to configure loggers and download logger data, available at <https://www.onsetcomp.com/products/software/hoboware>. Numerical modeling simulations were run in HydroGeoSphere (Aquanty Inc., 2015), available under a license at <https://www.aquanty.com/hydrogeosphere>. Temporal moment calculations were done in Matlab (MATLAB, 2020) under an academic license of version 9.9.0 or later, available at <https://matlab.mathworks.com/>. Statistical analyses and figure creation were done in open-source R Studio software (R Core Team, 2022) using version 4.2.1, available at <https://www.R-project.org/>.

DOI: 10.1029/2022WR033139

ATED (CDR) SPECTRA v1.0 data set at the PDS Geosciences Node, the ALICE High Performance Computing



What would AGU need?

PIDs in Availability Statement: Not (currently) machine-actionable – but we *can* include links to a landing page with the PID. DOIs can be cited...

Make format discoverable, and, please, consider the number of standards you're creating!

Different domains will need different solutions – we need something universal enough to implement at the publisher level



Time:

- Compliance checks for data and software citations add ~10 min per paper of staff time!

Consider implementation strategy!

Guidance for all levels:

- Publishers & staff
- Editors and reviewers
- Researchers
- Facilities / Institutions



300k AGU researchers



Many institutions



Few publishers, but must address legacy tech for best results

Thank you!



Kristina Vrouwenvelder (kvrouwenvelder@agu.org)

Shelley Stall (sstall@agu.org)

