

What is the existing landscape and what gaps exist in that landscape for data producers and users?

Title:

FAIR-aligned Scientific Repositories: Essential Infrastructure for Open and FAIR Data

Abstract:

In the Earth, space, and environmental sciences, a coalition of publishers, repositories, and research data infrastructure organizations are working together in an effort led by the American Geophysical Union to ensure data are open and FAIR, cited in publications, and discoverable for analysis and reuse. Common policies and practices are codified in a Commitment Statement currently accepting signatories at COPDESS.org.

The stakeholders of this 'Enabling FAIR Data' project include repositories with varying capability and researcher services. Some are well-prepared for these new policies and others will be challenged with resource limitations in both staff and technology to meet the new demands.

Today, many researchers across a wide range of domains still tend not to share their data (37% across domains, and 54% for specific domains) when they submit a manuscript for publication (Stuart et al. 2018). As we transition to this new policy, researchers will have to navigate the complex process of locating an appropriate repository that is FAIR-aligned, community-accepted, and able to receive and curate datasets.

Using information in re3data.org, of the repositories identified within the geosciences (over 600), nearly half are closed for researcher submitted datasets. The remaining repositories have some level of restriction (nearly 400). Of these, only 50% use persistent identifiers and less than 7% have a relationship with ORCID that supports automated attribution and credit.

In short, researchers do not have a clear path to a repository that can meet their needs for open and FAIR data. Many will be left with a 'best choice' option, or even, a 'last resort' option with general repositories being the last and possibly only choice a researcher has for making their data open. Services to support FAIR data are commonly not available for a general repository -- or at least not yet.

Working together on a solution that makes repositories easier to find, and researcher services more available is complex and essential for ensuring data are open and FAIR. We must also

implement attribution and credit for sharing data, and of equal importance, reusing data, demonstrating the value data has in furthering science.

Bio:

Shelley Stall, Director, AGU Data Programs

Shelley Stall manages AGU's Data Programs including the [Enabling FAIR Data](#) project funded by the Laura and John Arnold Foundation and the Data Management Assessment Program. Shelley has more than two decades of experience in high-volume, complex data management environments. She has helped organizations in not-for-profit, commercial, defense, and federal civilian communities address implementation of regulation, interoperability, data governance, metadata management, master data management, and organizational change management. Shelley has diverse experience as a program and project manager, software architect, database architect, performance and optimization analyst, data product provider, and data integration analyst. Shelley is one of five certified Enterprise Data Management Experts (EDME) through the CMMI Institute's Data Management Maturity (DMM) program and a Certified Data Management Professional (CDMP) through DAMA International's certification program. Shelley has a degree in Mathematics and a Master of Business Administration (management) with a focus on technical operations. She is a Project Management Professional (PMP) through the Project Management Institute. Since coming to AGU, Shelley has worked extensively with the Earth and space science repository community, publishers, and related organizations on data management. Shelley's publications, talks, and posters are in her [ORCID](#) profile.