12-13-2018

Research data sharing: Practices and attitudes of geophysicists

Carol Tenopir
University of Tennessee - Knoxville, ctenopir@utk.edu

Lisa Christian
University of Tennessee, Knoxville

Suzie Allard
University of Tennessee, Knoxville, sallard@utk.edu

Josh Borycz
University of Tennessee, Knoxville

Recommended Citation
Tenopir, C., Christian, L., Allard, S., & Borycz, J. (2018). Research data sharing: Practices and attitudes of geophysicists. Earth and Space Science, 5(12), 891-902.

AGU100 ADVANCING EARTH AND SPACE SCIENCE

Earth and Space Science

RESEARCH ARTICLE
10.1029/2018EA000461

Key Points:
- Survey of geophysicists from the American Geophysical Union on data management attitudes and practices
- Scientists have positive attitudes toward sharing and reusing data in general but are have important concerns about sharing their own data
- Assuring scientists that their data will be properly cited and providing metadata assistance could help improve sharing behavior

Correspondence to:
C. Tenopir,
ctenopir@utk.edu

Citation:
Tenopir, C., Christian, L., Allard, S., & Borycz, J. (2018). Research data sharing: Practices and attitudes of geophysicists. *Earth and Space Science*, *5*, 891–902. https://doi.org/10.1029/2018EA000461

Received 1 SEP 2018
Accepted 8 NOV 2018
Accepted article online 19 NOV 2018
Published online 13 DEC 2018

Research Data Sharing: Practices and Attitudes of Geophysicists

C. Tenopir[1], L. Christian[1], S. Allard[1], and J. Borycz[1]

[1]Univeristy of Tennessee, School of Information Science, Knoxville, TN, USA

Abstract Open data policies have been introduced by governments, funders, and publishers over the past decade. Previous research showed a growing recognition by scientists of the benefits of data-sharing and reuse, but actual practices lag and are not always compliant with new regulations. The goal of this study is to investigate motives, attitudes, and data practices of the community of Earth and planetary geophysicists, a discipline believed to have accepting attitudes toward data sharing and reuse. A better understanding of the attitudes and current data-sharing practices of this scientific community could enable funders, publishers, data managers, and librarians to design systems and services that help scientists understand and adhere to mandates and to create practices, tools, and services that are scientist-focused. An online survey was distributed to the members of the American Geophysical Union, producing 1,372 responses from 116 countries. The attitudes of researchers to data sharing and reuse were generally positive, but in practice, scientists had concerns about sharing their own research data. These concerns include the possibility of potential data misuse and the need for assurance of proper citation and acknowledgement. Training and assistance in good data management practices are lacking in many scientific fields and might help to alleviate these doubts.

1. Introduction

Interest in sharing and reusing data has increased over the last decade as funders, publishers, and governments have begun to implement more open data policies or mandates (Putri et al., 2015; van den Van Den Eynden et al., 2016; Zuiderwijk & Janssen, 2014), reproducibility of science has become a growing concern (McNutt, 2014; Yaffe & Koch, 2015), and scientists increasingly recognize that there are real benefits to open data (Lowndes et al., 2017; McKiernan et al., 2016). Sharing research data in open repositories is now required for research funded by the European Commission and the Welcome Trust and is recommended by U.S. Federal agencies such as the National Institute of Health (n.d.) and National Science Foundation (n.d.). However, many scientists are still not yet entirely compliant with local, regional, or international data-sharing requirements. Scientists also have some concerns about open data that negatively impact compliance (Chatfield & Reddick, 2018; Putri et al., 2015; Tenopir, Dalton, et al., 2015).

There is evidence that attitudes toward data sharing and data-sharing practices vary by subject discipline. Open data tend to be more accepted in disciplines that (1) do not deal with human subjects, (2) involve large-scale instrumentation that is shared by many to collect data, (3) have established metadata standards for data description, or (4) have a long history of data sharing and openness (Akers & Doty, 2013; Herold, 2015; Kim & Burns, 2016). Earth and planetary geophysics therefore are expected to be among the leaders in data sharing, although there may be variations between subdisciplines.

Official data policy of American Geophysical Union (AGU) requests that "all data necessary to understand, evaluate, replicate, and build upon the reported research must be made available and accessible whenever possible (AGU Publication Data Policy, 2016)." This policy, according to AGU, "is grounded in the value of full and open sharing of such data and associated documentation for research and education. Adherence to this policy will foster scientific advances, yield economic benefits, improve decision-making, enhance public safety and well-being, contribute to national and global security, and lead to a more informed public." Strict open data standards and proper data management, according to AGU, is crucial for the Earth and Space science because "the state of natural systems is never repeated, data losses, or missed data collection opportunities can never be corrected. Consequently, the value of data grows with time, placing a premium on very long-term data curation… For some issues, such as responding to natural hazards, access to real-time data is critical (AGU Publication Data Policy, 2016)."

Understanding the motivations, attitudes, and current data-sharing practices of scientists in this community should enable funders, publishers, data managers, and librarians to design systems and services that help scientists understand and adhere to relevant mandates and to create practices, tools, and assistance that are research focused.

The goal of this study was to answer questions that relate to the research data-sharing practices and attitudes of researchers that work closely with the American Geophysical Union (AGU) and was performed in collaboration with both the AGU (AGU; agu.org) and the DataONE distributed data framework (dataone.org). AGU members come from all over the confidential manuscript submitted to Earth and Space Science world and include scientists in the atmospheric and ocean sciences, solid-Earth sciences, hydrologic sciences, and space sciences (agu.org). The primary research questions are as follows:

Do AGU scientists …

- … currently share their research data?
- … currently use data management practices, including creating data management plans, providing metadata and provenance information?
- … currently reuse data collected by others in their work?
- … have positive or negative attitudes toward research data sharing and data reuse?
- … differ in their practices and attitudes by subdiscipline of science?

Research data are defined here as "the recorded factual material commonly accepted in the scientific community as necessary to validate research findings" (Department of Energy, 2018); open data are defined as "publicly available data structured in a way that enables the data to be fully discoverable and usable by end users" (Burwell et al., 2013). Data may be deposited in subject repositories such as DataBasin, IRIS (Incorporated Research Institutions for Seismology), and BCO-DMO (Biological and Chemical Oceanography Data Management Office), or general repositories such as Dryad, FigShare, or PANGAEA (Publishing Network for Geoscientific & Environmental Data).

1.1. Related Work

Previous research on data-sharing and reuse practices have demonstrated that while both the scientific and the publishing communities see these practices in an increasingly favorable light, there are several significant barriers to widespread adoption of data sharing. Tenopir, Dalton, et al. (2015) and Tenopir et al. (2011) discovered that the majority of scientists were "willing to share at least some of their data and reuse others' data pending certain conditions or restrictions on use (Tenopir et al., 2011)." They also demonstrated that scientists may be more willing to share data if they are guaranteed to receive formal citation for their work and are provided knowledge of and access to the research that uses the data; however, funder mandates may be the most important motivator (Schmidt et al., 2016). Some of the obstacles that prevent data sharing include insufficient time, the need to publish first, and lack of funding (Tenopir et al., 2011; Tenopir, Dalton, et al., 2015). Wallis et al. (2013) learned that scientists were generally willing to share their data as long as their rights as researchers and authors were protected. For example, scientists generally want to be able to publish their data findings first, for their data to be cited properly and interpreted correctly, and for the process of sharing data to be simple and convenient. Convenience is often a factor in scientists' decisions on where and when to share their data sets (Yoon, 2017). Scientists are more likely to share data if the processes are standardized, simple, and they are given assistance (van Den Van Den Eynden et al., 2016).

Other research has suggested that geographic location also influences data-sharing practices and attitudes, at least in shaping the economic environments that affect data-sharing activities (for example, according to Bezuidenhout and Chakauya (2018), low/middle-income countries have limitations in resource provision, research support, and extralaboratory infrastructures that shape data-sharing practices).

Many data managers have argued that lack of metadata standardization prevents researchers from openly sharing their data. Improving these standards may very well increase the level trust that scientists have for other members of the scientific community (Yoon, 2017). This is because rich metadata increases the discoverability, accessibility, and validity of data sets and ensures that scientists can reuse data without later discovering that it contains mistakes or was falsified (Faniel et al., 2016).

Some researchers point to the need for a joint effort by the scientific, publishing, and policy-making communities to work together to increase the data-sharing and reuse practices of the scientific community. Wallis et al. (2013) stressed that "collaborative effort is needed to address data sharing and data reuse, one that supports the needs of scientists, researchers, funding agencies, and the public." An editorial from Science (2011) pointed out that researchers and publishers should treat data as more than just a supplement in science, "We must all accept that science is data and that data are science, and thus provide for, and justify the need for the support, of much-improved data curation (Wallis et al. (2013)."

The editors of the New England Journal of Medicine's *Sounding Board* recently argued that a more standardized method of organizing the authorship of data sets would provide greater incentives to researchers to share data by ensuring that their work was properly cited (Bierer et al., 2017). They argued that realigning the incentives of data publication by ensuring the proper citation of data sets could help remove some barriers to a more open scientific environment (Bierer et al., 2017; Kim & Stanton, 2016). Pampel and Dallmeier-Tiessen (2014) was also able to confirm that data citation and data sharing can lead to more citations. Tenopir, Dalton, et al. (2015) and Aleixandre-Benavent et al. (2016) argued that this should motivate younger academics seeking promotion and tenure to share more of their data. Multiple large, federally funded research projects from countries in the European Union, Canada, and the United Kingdom have demonstrated a positive impact on research when certain data-sharing practices are mandated (Douglass et al., 2014; Kim & Burns, 2016; Nugroho et al., 2015).

2. Methods

Researchers from the DataONE Usability & Assessment Working Group designed the survey instrument used in this work to capture scientists' perceptions about data sharing, information about current data-sharing practices, satisfaction with data tools, and perceptions of the organizational support provided for research processes. The questionnaire was modeled after previous work that was also completed DataONE (Tenopir et al., 2011; Tenopir, Dalton, et al., 2015).

In March 2017 the American Geophysical Union (AGU) first distributed the online survey to all 62,000 of its members. The e-mail invitation included an embedded link to the online survey hosted on the University of Tennessee server. After an e-mail reminder in August 2017, the survey closed in March 2018 with 1,372 responses from 116 countries for a response rate of 2.2%.

The study was approved by the University of Tennessee Institutional Review Board as an anonymous online survey. Findings are reported in aggregate with no personally identifiable information. In compliance with the Institutional Review Board approval for work with human subjects, respondents could skip any question or withdraw from the study at any time. The first page of the survey included an informed consent, in which respondents indicated that they understood the terms and were over 18 years of age.

Survey data were collected in Qualtrics and housed on a secure server at the University of Tennessee. Researchers exported data to IBM SPSS 25 Statistical Analysis Software Package for analysis. For all correlations, the significance is at the.05 level unless otherwise stated.

3. Results

3.1. Survey Population and Demographics

Respondents match the AGU membership fairly closely. They are distributed relatively evenly between age groups, with 43.1% under 40 years of age. Over two thirds (68.8%) of respondents were identified as male. This imbalance is reflective of the larger AGU membership, in which just 28% identify as female (American Geophysical Union, 2017). This distribution also reflects the broader gender imbalance in the sciences (American Geophysical Union, 2017). Just over 50% of survey respondents were from the United States (51.5%), with an additional 23.6% being from Europe and Russia; 3.6% from Central and South America; 11.9% from Asia, Africa, or the Middle East; and 9.7% from elsewhere. This distribution is consistent with AGU membership, of which 61% come from the United States, 19% from Europe, and 2% from Central and South America (American Geophysical Union, 2014).

The survey provided 18 subject disciplines from which respondents could choose, including a choice of *other*, which allowed them to specify their exact subject discipline. For the purposes of analysis, we collapsed the subject disciplines into broad but related categories (Tables 1 and 2). Most respondents come from Geology/Earth Science (26.4%) and the Life Sciences (20.3%). Other disciplines include education,

Table 1
Subject Disciplines of AGU Respondents[a]

Subject Discipline	Survey Respondents	AGU Membership
Life Sciences	20.3%	25%
Atmospheric Science	17.2%	28%
Engineering/Information Science/ Computer Science	6.8%	7%
Geology/Earth Science	26.4%	14%
Hydrology	8.3%	12%
Physical Sciences	14.7%	13%
Other	6.3%	10%

[a]See Appendix A for detailed explanation of how the disciplines were grouped by the authors.

psychology, social sciences, and unspecified multidisciplinary respondents. This distribution was similar to AGU membership, of which 25% were in the life sciences, 13% were in the physical sciences, and 12% were in hydrology (American Geophysical Union, 2014).

Field Research (37.5%) and Modeling (38%) were the most common research activities of the survey respondents. Lab Research (19.2%) and Geographic Information Systems (6.5%) were also performed by a significant proportion of the researchers in this survey. The majority of respondents come from academia (68.7%), with government a distant second (20.2%), followed by nonprofit (4.2%), commercial (4.0%), and other (2.8%). This distribution is also representative of AGU membership, of which 55% are from academia, 17% are governmental, 14% are students, 13% are nonspecified/other, and 7% are from industry (American Geophysical Union, 2014).

3.2. Current Data-Sharing and Data Management Practices

Despite the fact that only slightly over one third of respondents report that their organizations require a data management plan during the life of the project (the short term; 36.3%) or beyond the life of the project (long term; 39.5%), Figure 1 shows that most researchers are satisfied with the processes that they use to store their data short term (74.8%) and approximately half are satisfied with the processes that they use to store their data long term (51.2%).The levels of satisfaction with metadata tools, provenance tracking, and ability to locate suitable data storage repositories may be lower primarily because many respondents are either unsure of the meanings of these terms or the resources provided by their institutions that relate to these terms (Ahonen-Rainio & Kraak, 2005).

There are some disciplinary differences in metadata preparation (χ^2 = 28.001, d.f. = 12, p = 0.006). A slight majority of respondents in the life sciences (43%), hydrology (37%), and other disciplines (36%) are not satisfied with their tools for preparing metadata, while those in engineering/information science/computer science (47%), geology/Earth science (42%), and physical science (42%) have no opinion on metadata preparation. Those in the atmospheric sciences (37%) are the most satisfied. Part of this may be due to the fact that atmospheric scientists tend to be more familiar with metadata tools as a result of widespread promulgation by federal agencies such as the National Aeronautics and Space Administration through their Distributed Active Archive Center program (https://nssdc.gsfc.nasa.gov/earth/daacs.html).

Most respondents report that their organizations do not provide training or assistance in data management practices (Figure 2). Again, there are some disciplinary differences between data management services, with more respondents in geology/Earth science saying that their organizations *do not* provide training for data management or for creating data management plans. More respondents in the life sciences and hydrology seem to receive assistance with their data management.

Whether they receive training in data management plans or not, respondents are generally confident in their abilities to generate plans related to data collection and management (Figure 3). Researchers also frequently

Table 2
Groups of Subject Disciplines

Subject Discipline Group	Individual Subject Disciplines
Life Sciences	Ocean sciences, bio geoscience, global environmental change, paleoceanography
Atmospheric Science	Atmospheric sciences, space physics and aeronomy, atmosphere and space electricity
Engineering/Information Science/ Computer Science	Technophysics, Earth and space science informatics
Geology/Earth Science	Planetary science, Earth and planetary science processes, geomagnetism, geodesy, Earth's deep interior, near surface geophysics, mineral and rock physics
Hydrology	Hydrology
Physical Sciences	Volcanology/geochemistry, seismology, nonlinear geophysics
Other	No data, natural hazards, cryosphere science, societal impacts

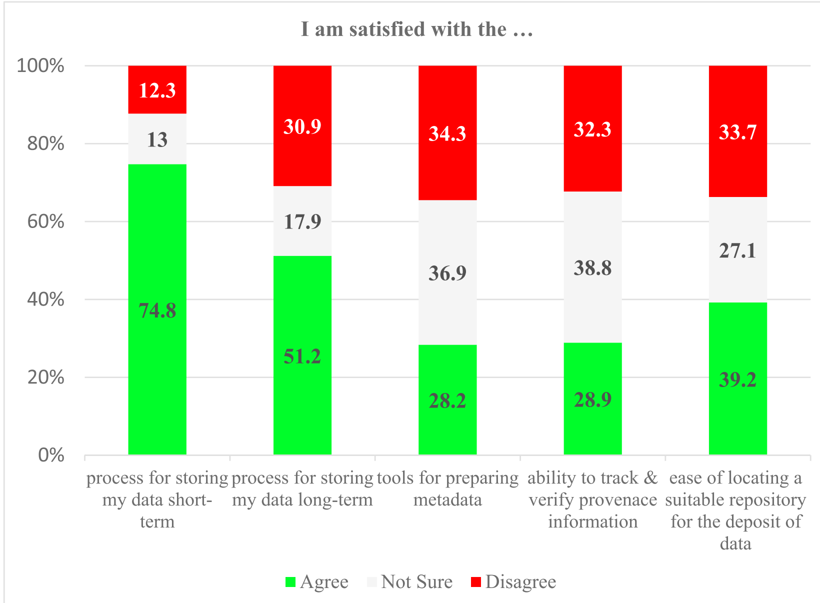

Figure 1. Satisfaction by AGU respondents on their ability to store and manage their data.

go to colleagues if there is a data need that they cannot satisfy for themselves. They are significantly less likely to ask for help from a librarian or to consult a data management expert.

3.3. Data Sharing and Data Reuse

Trustworthiness is an important consideration for researchers when they are deciding which data to use and to share. For instance, respondents recognize the importance of data access to scientific progress (Figure 4). Approximately three quarters of researchers report that the lack of access to others' data impedes scientific progress, and half of them also state that lack of access restricts their own ability to conduct research. These low levels of sharing could be attributable to the concerns that researchers express when they share their

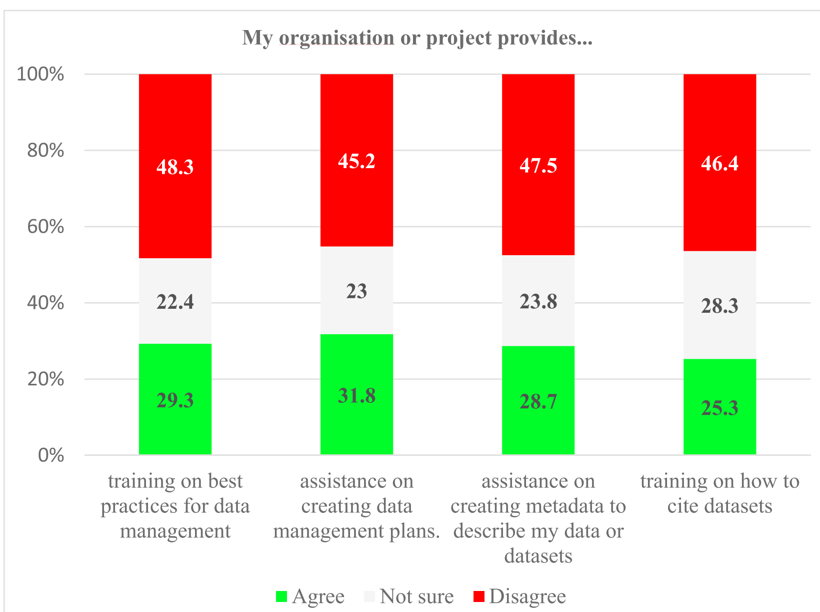

Figure 2. Agreement by AGU respondents on whether their organizations assist with data management.

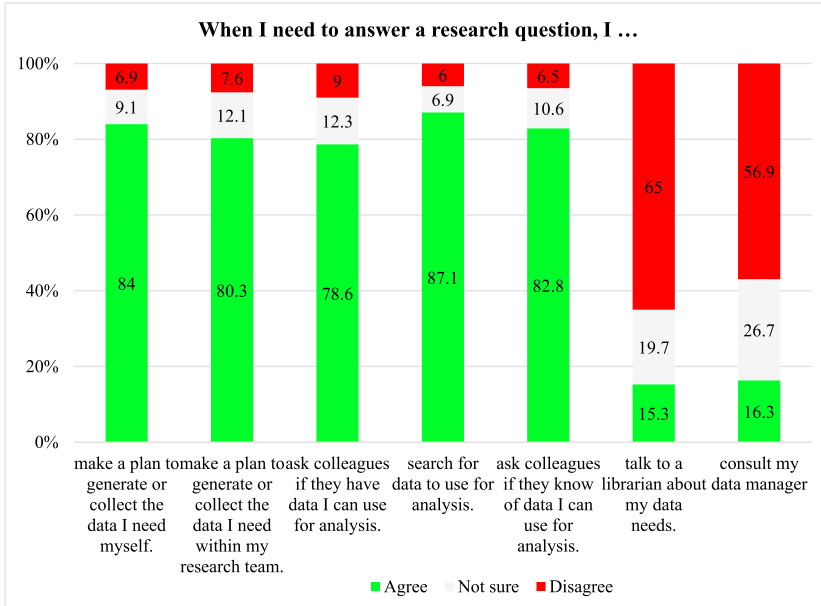

Figure 3. Agreement by AGU respondents on where they turn when they need data.

own data. Respondents fear that data can be misinterpreted because of complexity (79%) or poor quality (78%). They also fear that data may be used in other ways than intended (76.5%).

The quality of metadata can have a strong impact on the confidence of researchers when using other's data (Figure 5). Including metadata (72.2%), details about the collection procedure (84.3%) and provenance information (60.7%) will increase confidence and would make them more likely to use others' data.

Attitudes toward data sharing and data reuse tend to be quite positive (Figure 6) as long as there are assurances that data are cited properly and that researchers can specify which data they would like to share. Willingness to use other researchers' data sets if they are easily available (91.3% agree), willingness to share data across a broad group of researchers (89.4% agree), and willingness to place at least some data in a

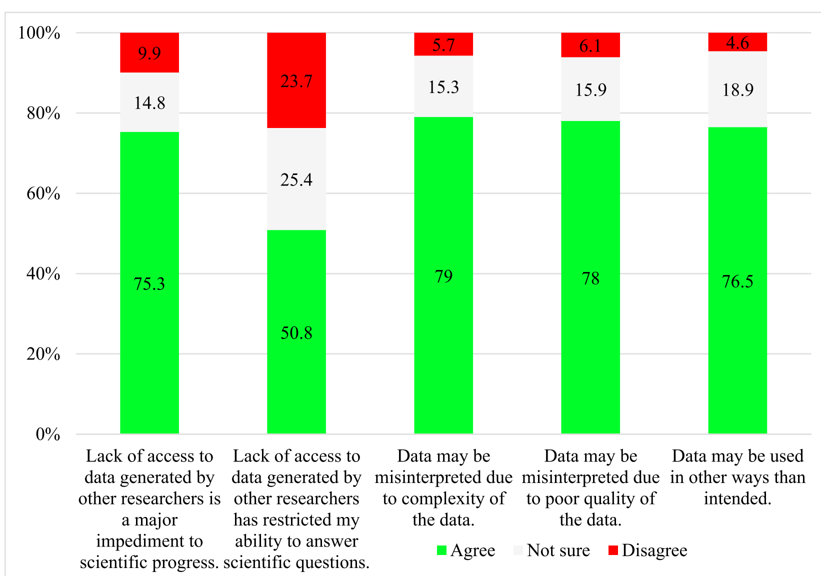

Figure 4. Agreement by AGU respondents on the use of scientific data.

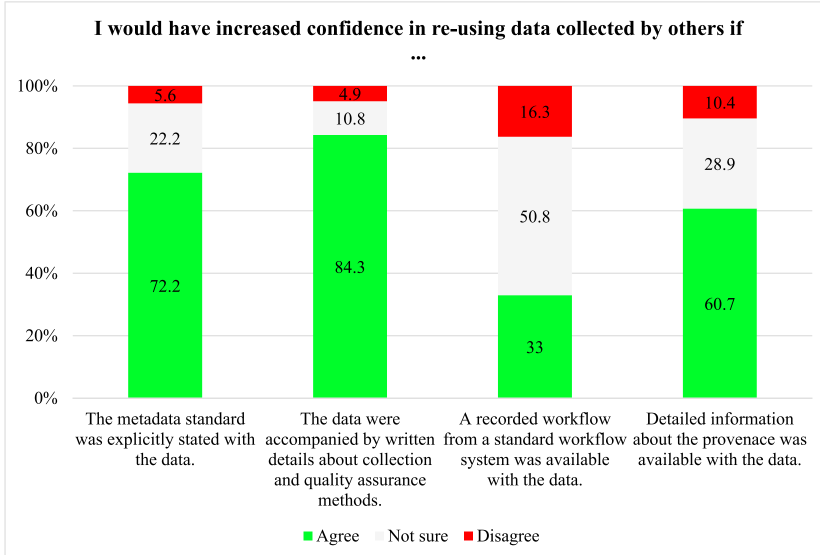

Figure 5. Agreement by AGU respondents on the reuse of scientific research data.

central repository (82.3% agree) are received positively by a strong majority. This willingness is dependent, however, on assurances that their data are cited (92.2%) and less than half (49.1%) are willing to place all of their data in a central repository with no restrictions.

Regardless of discipline, most respondents are willing to share their data. However, more respondents in the life sciences (95%), hydrology (91%), and atmospheric science (90%) agree that they would share their data across a broad group of researchers. This level of agreement drops slightly for geology/Earth science (88%), physical science (86%), and engineering/information science/computer science (82%).

Figure 6. Agreement by AGU respondents on sharing scientific data.

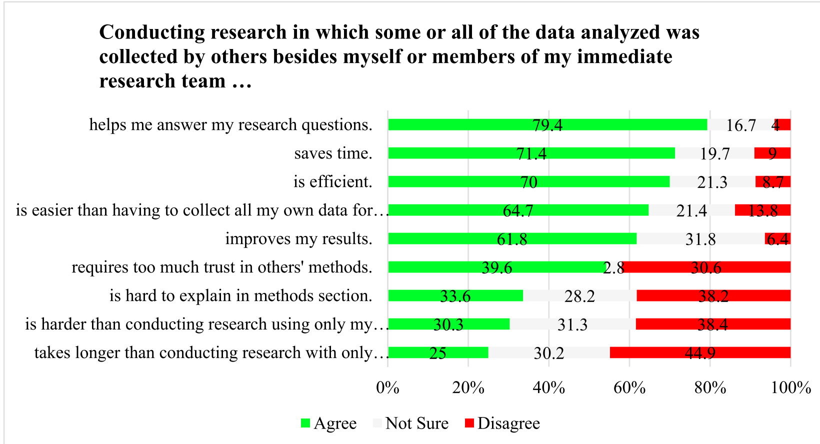

Figure 7. Agreement by AGU respondents on using data collected by others.

Hesitancy in sharing data is more prevalent among those in the physical sciences (11%) and engineering/information science/computer science (15%). By contrast, only 9% of those in atmospheric science, 7% of those in hydrology and geology/Earth science, 6% of *others* and just 3% of those in life sciences are unsure about the benefits of sharing their data.

Most respondents felt that using others' data sets helps them answer research questions (79.4%), saves time (71.4%), and is efficient (70%; Figure 7), and is not more difficult than research using only their data.

Respondents had a much more mixed reaction with regard to the *trustworthiness* of others' data (Table 3). The chi-squared result (χ^2 = 22.835, d.f. = 12, p = 0.029) showed that there was a significant difference by discipline. Disciplines that share instrumentation and rely on communally collected data are expected to trust shared data more. More respondents in atmospheric science (77.2%), physical sciences (71.8%), and life sciences (70%) agree that using data collected by others is more efficient. The level of respondent agreement dips to 65–69% for respondents in other disciplines and 29.1% of respondents in hydrology report being unsure of the efficiency of using others' data.

A significant portion of respondents were unsure about the trustworthiness of others' data collection methods. Almost half of respondents in hydrology and engineering/information science/computer science, 42.9% of those in *other* disciplines, 42.4% of those in geology/Earth science, and 41.1% in life sciences agree that using data collected by others requires too much trust in others' methods. This agreement drops to 35.0% for those in the physical science and 29.6% for those in atmospheric science.

3.4. Reciprocity of Data Sharing

The vast majority of AGU respondents would allow other researchers to use their data provided they receive an acknowledgement (90.4%) and/or citations (87.5%; Figure 8). A slight majority (53.5%) would also like the opportunity to collaborate. Most do not require final approval or review before dissemination nor even legal permission or co-authorship.

Respondents in life sciences (32%), engineering/information sciences/computer sciences (39%), and geology/Earth sciences (34%) want legal permission to use their data as well (χ^2 = 24.403, d.f. = 12, p = 0.018). Only 26% of those in the physical sciences, 25% of other disciplines, 25% of hydrology respondents, and 21% of those in atmospheric sciences required legal permission.

4. Discussion, Limitations, and Conclusions

4.1. Discussion

This research indicates that scientists in general may be unsure of the use or meaning of common data management terms and tools such as metadata, provenance, and public repositories. Furthermore, there is a perceived low level of assistance with various data management tasks (Figure 2). Scientific researchers indicated a lack of awareness that there are data management or information experts in their institutional

Table 3
Discipline Differences by Agreement of the AGU Respondents on Using Data Collected by Others

Conducting research in which some or all of the data analyzed was collected by others besides myself or members of my immediate research team

	Discipline	Disagree	Not Sure/Neutral/ No Opinion	Agree	Total
Is efficient (χ^2 = 22.476, d.f. = 12, p = 0.033)	Life sciences	(n = 20)	(n = 43)	(n = 147)	(n = 210)
		9.5%	20.5%	70.0%	100.0%
	Atmospheric science	(n = 9)	(n = 32)	(n = 139)	(n = 180)
		5.0%	17.8%	77.2%	100.0%
	Engineering/Information science/	(n = 7)	(n = 13)	(n = 39)	(n = 59)
	Computer science	11.9%	22.0%	66.1%	100.0%
	Geology/Earth science	(n = 36)	(n = 58)	(n = 192)	(n = 286)
		12.6%	20.3%	67.1%	100.0%
	Hydrology	(n = 1)	(n = 25)	(n = 60)	(n = 86)
		1.2%	29.1%	69.8%	100.0%
	Physical science	(n = 12)	(n = 32)	(n = 112)	(n = 156)
		7.7%	20.5%	71.8%	100.0%
	Other	(n = 4)	(n = 18)	(n = 41)	(n = 63)
		6.3%	28.6%	65.1%	100.0%
Requires too much trust in others' methods (χ^2 = 22.835, d.f. = 12, p = 0.029)	Life sciences	(n = 70)	(n = 53)	(n = 86)	(n = 209)
		33.5%	25.4%	41.1%	100.0%
	Atmospheric science	(n = 75)	(n = 51)	(n = 53)	(n = 17)
		41.9%	28.5%	29.6%	100.0%
	Engineering/Information science/	(n = 15)	(n = 17)	(n = 27)	(n = 59)
	Computer science	25.4%	28.8%	45.8%	100.0%
	Geology/Earth science	(n = 76)	(n = 90)	(n = 122)	(n = 288)
		26.4%	31.3%	42.4%	100.0%
	Hydrology	(n = 21)	(n = 26)	(n = 40)	(n = 87)
		24.1%	29.9%	46.0%	100.0%
	Physical science	(n = 47)	(n = 55)	(n = 55)	(n = 157)
		29.9%	35.0%	35.0%	100.0%
	Other	(n = 16)	(n = 20)	(n = 27)	(n = 63)
		25.4%	31.7%	42.9%	100.0%
Improves my results (χ^2 = 23.620, d.f. = 12, p = 0.023)	Life sciences	(n = 9)	(n = 57)	(n = 144)	(n = 210)
		4.3%	27.1%	68.6%	100.0%
	Atmospheric science	(n = 8)	(n = 56)	(n = 115)	(n = 179)
		4.5%	31.3%	64.2%	100.0%
	Engineering/Information science/	(n = 7)	(n = 16)	(n = 35)	(n = 58)
	Computer science	12.1%	27.6%	60.3%	100.0%
	Geology/Earth science	(n = 27)	(n = 85)	(n = 175)	(n = 287)
		9.4%	29.6%	61.0%	100.0%
	Hydrology	(n = 4)	(n = 37)	(n = 46)	(n = 87)
		4.6%	42.5%	52.9%	100.0%
	Physical science	(n = 10)	(n = 50)	(n = 94)	(n = 154)
		6.5%	32.5%	61.0%	100.0%
	Other	(n = 2)	(n = 28)	(n = 32)	(n = 62)
		3.2%	45.2%	51.6%	100.0%

libraries (Figure 3). (See Yoon and Schultz (2017) and Tenopir, Hughes, et al. (2015) for recent studies of library data management assistance.) This combination of perceived lack of assistance and lack of common data management knowledge may not negatively impact researcher's satisfaction of their data storage practices, however (Figure 4), with three quarters (75.3%) indicating that they are satisfied with their short-term data storage practices and half (50.8%) indicating that they are satisfied with their long-term data storage practices.

Scientists have positive attitudes toward data sharing and reuse in general. Scientists acknowledge that sharing scientific data can have a positive impact on scientific progress with regard to time savings and research efficiency, but when it comes to sharing their own research data, scientists have concerns, including worries that it being misused or misinterpreted (Figure 8). These concerns may be alleviated if more metadata for published research was included, proper acknowledgement and citation of the utilized data was assured,

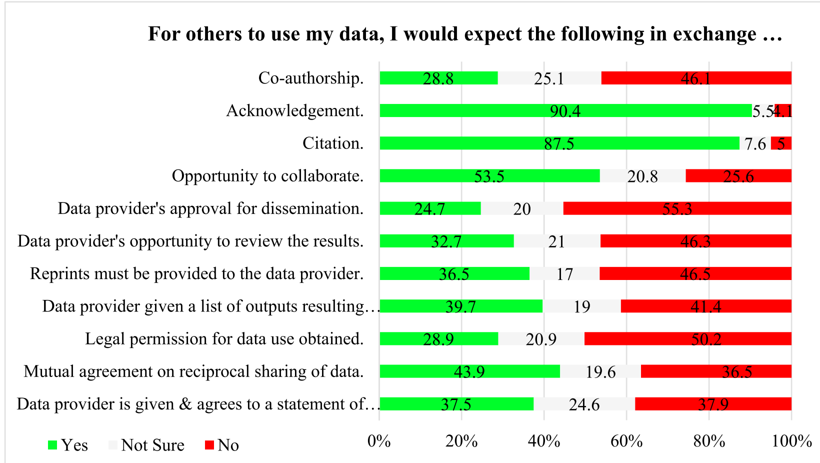

Figure 8. Conditions by AGU respondents for other researchers to use their data.

and some control over the use of the data was given to the original researchers (Figures 6 and 8). Assisting researchers with metadata creation, adding provenance information to data sets, and searching for adequate public data repositories may also help researchers to adopt better data-sharing practices.

In general, metadata practices and use of standards needs to be improved, training in or assistance with data management tasks are perceived to be lacking, and many are unaware they need or can ask for help. Scientists in particular fields, such as the life sciences and hydrology, receive significantly more assistance with metadata, creating data management plans, and training for best practices in data management than their counterparts in atmospheric science, physical sciences, and geology/Earth science. This tendency to receive more assistance and training may also explain in part the willingness of those in the life sciences and hydrology to share their data with other researchers. Researchers in these fields tend to be more informed about best practices in data sharing more often, and perhaps, as a result, feel more confident about the ability of other researchers' to use their data responsibly.

Generalizing the principles that have been adopted by these fields would go a long way to improving data management standards and data-sharing behaviors in all scientific fields. Our data indicate that some of the key first steps to improving behaviors would be to assure proper acknowledgement and citation of data used by all researchers and to advertise the data management expertise of data librarians or research data managers.

4.2. Limitations

Since respondents were allowed to skip any question or to exit the survey at any time and they were automatically timed out of the survey if they did not complete it within a week, response rates for individual questions may differ from the overall response rate of 2.2%. For the purposes of analysis, each question is considered to have a response rate of 100%; that is, if 987 respondents answered a certain question, then that question has a response rate of 100% or 987 of 987. With a response rate of only 2.2% at a maximum, we must assume that geophysicists who are knowledgeable about open data issues are more likely to respond. The respondents do, however, approximate AGU membership regarding age, gender, and geographic distribution. Although the survey was distributed internationally and had responses from 73 countries, only 3 countries had over 50 responses, we did not have enough responses to make claims about the differences in data sharing and reuse based on individual country.

4.3. Conclusions

The survey of the AGU statistics showed that the attitudes of researchers to data sharing and reuse were generally positive. At the same time, scientists still had concerns about sharing their own research data. The concerns include potential misuse of their data and needing assurance of adequate citation and acknowledgement. These concerns can be potentially addressed by assuring that proper citation, acknowledgement, and metadata are provided.

Assistance and training in good data management practices are lacking in many scientific fields. Examples of several fields, including the life sciences and hydrology, that receive significantly more assistance, show that assistance and training can eliminate some of the concerns about sharing research data.

While many scientists do not have appropriate metadata standards and are unaware of where to ask for help, it provides the data librarians and research data managers with an opportunity to proactively reach out and assist scientists with their data needs.

Acknowledgments
This project, Data Observation Network for Earth (DataONE), is funded by the National Science Foundation (NSF), award 080944, under a cooperative agreement, William Michener, Principal Investigator. NSF had no role in the research design, data collection, data analysis, nor in the writing of this paper or the decision where to publish. The authors of this article have no conflicts of interest to declare. We would like to thank the American Geophysical Union for their assistance, in particular Brooks Hanson, AGU Executive Vice President of Science. We would also like to thank Natalie Rice of the Center for Information and Communication Studies (University of Tennessee), and graduate research assistants Paris Whalon and Hannah Blanco, also of the University of Tennessee. The data for this study can be found in the Dryad data repository at the time of publication.

References

Ahonen-Rainio, P., & Kraak, M.-J. (2005). Deciding on fitness for use: Evaluating the utility of sample maps as an element of geospatial metadata. *Cartography and Geographic Information Science*, *32*(2), 101–112. https://doi.org/10.1559/1523040053722114

Akers, K. G., & Doty, J. (2013). Disciplinary differences in faculty research data management practices and perspectives. *The International Journal of Digital Curation*, *8*(2), 5–26. https://doi.org/10.2218/ijdc.v8i2.263

Aleixandre-Benavent, R., Moreno-Solano, L. M., Ferrer Sapena, A., & Sánchez Pérez, E. A. (2016). Correlation between impact factor and public availability of published research data in information science and library science journals. *Scientometrics*, *107*(1), 1–13. https://doi.org/10.1007/s11192-016-1868-7

American Geophysical Union (2014). AGU demographics. [online Image]. Retrieved from URL https://sites.agu.org/media-kits/files/2014/12/AGU-Demographics.pdf

American Geophysical Union (2017). Demographics for 2017 honors cycle. Author. Retrieved from URL https://honors.agu.org/files/2014/11/2017-Honors-Cycle-Demographics.pdf

Bezuidenhout, L., & Chakauya, E. (2018). Hidden concerns of sharing research data by low/middle-income country scientists. *Global Bioethics*, *29*(1), 39–54. https://doi.org/10.1080/11287462.2018.1441780

Bierer, B. E., Crosas, M., & Pierce, H. H. (2017). Data authorship as an incentive to data sharing. *The New England Journal of Medicine*, *376*(17), 1684–1687. https://doi.org/10.1056/NEJMsb1616595

Burwell, S. M., VanRoekel, S., Park, T., & Mancini, D. J. (2013). M-13-13—Memorandum for the heads of executive departments and agencies: Open data policy. Washington D.C. Retrieved from https://project-open-data.cio.gov/policy-memo/

Chatfield, A. T., & Reddick, C. G. (2018). The role of policy entrepreneurs in open government data policy innovation diffusion: An analysis of Australian federal and state governments. *Government Information Quarterly*, *35*(1), 123–134. https://doi.org/10.1016/j.giq.2017.10.004

Department of Energy (2018). DOE policy for digital research data management: Glossary. Retrieved May 12, 2018 from https://www.energy.gov/datamanagement/doe-policy-digital-research-data-management-glossary

Douglass, K., Allard, S., Tenopir, C., Wu, L., & Frame, M. (2014). Managing scientific data as public assets: Data sharing practices and policies among full-time government employees. *Journal of the Association for Information Science and Technology*, *65*(2), 251–262. https://doi.org/10.1002/asi.22988

Faniel, I. M., Kriesberg, A., & Yakel, E. (2016). Social scientists' satisfaction with data reuse. *Journal of the Association for Information Science and Technology*, *67*(6), 1404–1416. https://doi.org/10.1002/asi.23480

Herold, P. (2015). Data sharing among ecology, evolution, and natural resources scientists: An analysis of selected publications. *Journal of Librarianship & Scholarly Communication*, *3*(2), 1–23. https://doi.org/10.7710/2162-3309.1244

Kim, Y., & Burns, C. S. (2016). Norms of data sharing in biological sciences: The roles of metadata, data repository, and journal and funding requirements. *Journal of Information Science*, *42*(2), 230–245. https://doi.org/10.1177/0165551515592098

Kim, Y., & Stanton, J. M. (2016). Institutional and individual factors affecting scientists' data-sharing behaviors: A multilevel analysis. *Journal of the Association for Information Science and Technology*, *67*(4), 776–799. https://doi.org/10.1002/asi.23424

Lowndes, J. S. S., Best, B. D., Scarborough, C., Afflerbach, J. C., Frazier, M. R., O'Hara, C. C., et al. (2017). Our path to better science in less time using open data science tools. *Nature Ecology & Evolution*, *1*(6), 0160. https://doi.org/10.1038/s41559-017-0160

McKiernan, E. C., Bourne, P. E., Brown, C. T., Buck, S., Kenall, A., Lin, J., et al. (2016). How open science helps researchers succeed. *eLife*, *5*, e16800. https://doi.org/10.7554/eLife.16800

McNutt, M. (2014). Reproducibility. *Science*, *343*(6168), 229. https://doi.org/10.1126/science.1250475

National Institute of Health (n.d.). NIH data sharing policy and implementation guidance. Retrieved May 9, 2018, from https://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm

National Science Foundation (n.d.). Dissemination and sharing of research results. Retrieved May 9, 2018, from http://www.nsf.gov/bfa/dias/policy/dmp.jsp

Nugroho, R. P., Zuiderwijk, A., Janssen, M., & de Jong, M. (2015). A comparison of national open data policies: Lessons learned. *Transforming Government: People, Process and Policy*, *9*(3), 286–308.

Pampel, H., & Dallmeier-Tiessen, S. (2014). Open research data: From vision to practice. In *Opening science* (pp. 213–224). Cham: Springer.

Putri, N. R., Anneke, Z., Marijn, J., & Martin, d. J. (2015). A comparison of national open data policies: Lessons learned. *Transforming Government: People, Process, and Policy*, *9*(3), 286–308. https://doi.org/10.1108/TG-03-2014-0008

Schmidt, B., Gemeinholzer, B., & Treloa, A. (2016). Open data in global environmental research: The Belmont Forum's open data survey. *PLoS One*, *11*(1), e0146695. https://doi.org/10.1371/journal.pone.0146695

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., et al. (2011). Data sharing by scientists: Practices and perceptions. *PLoS One*, *6*(6), e21101. https://doi.org/10.1371/journal.pone.0021101

Tenopir, C., Dalton, E. D., Allard, S., Frame, M., Pjesivac, I., Birch, B., et al. (2015). Changes in data sharing and data reuse practices and perceptions among scientists worldwide. *PLoS One*, *10*(8), e0134826. https://doi.org/10.1371/journal.pone.0134826

Tenopir, C., Hughes, D., Allard, S., Frame, M., Birch, B., Baird, L., et al. (2015). Research data services in academic libraries: Data intensive roles for the future. *Journal of eScience Librarianship.*, *4*(2). https://doi.org/10.7191/jeslib.2015.1085b

Van Den Eynden, V., Knight, G., Vlad, A., Radler, B., Tenopir, C., Leon, D., et al. (2016). Survey of Wellcome researchers and their attitudes to open research. *Technical Report*. https://doi.org/10.6084/m9.figshare.4055448.v1

Wallis, J. C., Rolando, E., & Borgman, C. L. (2013). If we share data, will anyone use them? Data sharing and reuse in the long tail of science and technology. *PLoS One*, *8*(7), e67332. https://doi.org/10.1371/journal.pone.0067332

Yaffe, M. B., & Koch, D. H. (2015). Reproducibility in science. *Science Signaling*, *8*(371), eg5456. https://doi.org/10.1126/scisignal.aaa5764

10.1029/2018EA000461

Yoon, A. (2017). Data reusers' trust development. *Journal of the Association for Information Science and Technology*, *68*(4), 946–956. https://doi.org/10.1002/asi.23730

Yoon, A., & Schultz, T. (2017). Research data management services in academic libraries in the US: A content analysis of libraries' websites. College & Research Libraries, [S.l.], v. 78, n. 7, p. 920, Oct. 2017. ISSN 2150–6701. Retrieved from https://crl.acrl.org/index.php/crl/article/view/16788/18346. Date accessed: 29 aug. 2018. https://doi.org/10.5860/crl.78.7.920

Zuiderwijk, A., & Janssen, M. (2014). Open data policies, their implementation, and impact: A framework for comparison. *Government Information Quarterly*, *31*(1), 17–29. https://doi.org/10.1016/j.giq.2013.04.003