





Data Sharing Practices in Agricultural Research: Findings from a Systematized Review

Isabella Baxter, University of Maryland, College Park, MD, USA, ibaxter@umd.edu 

Erin Antognoli, USDA National Agricultural Library, Beltsville, MD, USA 

Paria Aria, Texas Woman's University, Denton, TX, USA

Maggie Albro, University of Tennessee, Knoxville, TN, USA 

Jocelyn Boice, Colorado State University, Fort Collins, CO, USA

Michal McCullough, USDA National Agricultural Library, Beltsville, MD, USA 

Carolyn Jackson, Kansas State University, Manhattan, KS, USA

Abstract

Objective: Agricultural researchers who follow data sharing best practices advance the state of research in a variety of critical areas including plant breeding, cropping systems, and climate change adaptation. Data sharing makes research more reliable and reproducible, therefore, data sharing practices of researchers are integral to advancing science. To assess how agricultural researchers adhere to these practices, we conducted a systematized review of their published output and examined different ways data were shared.

Methods: Our study focused on corn and soybean production research published from 2017 to 2022 by authors at our institutions. We searched five databases, retrieved 8,271 articles, and created a randomized sample of 1,250 papers that contained an equal number of examples from each year. Following a rigorous set of criteria, we screened each article for inclusion and recorded the characteristics of the data, funder information, and whether the researchers shared data.

Received: July 19, 2024 **Accepted:** November 13, 2024 **Published:** December 4, 2024

Keywords: data sharing practices, research synthesis, agriculture and life sciences

Citation: Baxter, Isabella, Erin Antognoli, Paria Aria, Maggie Albro, Jocelyn Boice, Michal McCullough, and Carolyn Jackson. 2024. "Data Sharing Practices in Agricultural Research: Findings from a Systematized Review." *Journal of eScience Librarianship* 13 (3): e967. <https://doi.org/10.7191/jeslib.967>.

Data Availability: The data generated by this study are archived in the Open Science Framework and are available at <https://doi.org/10.17605/OSF.IO/PGHTU>. This protocol was registered with the Open Science Framework Registries and is available at <https://osf.io/amw6e>.

The *Journal of eScience Librarianship* is a peer-reviewed open access journal. © 2024 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

See <https://creativecommons.org/licenses/by/4.0>.

OPEN ACCESS

Abstract Continued

Results: Of the articles that met the inclusion criteria, less than 15% shared the full dataset associated with the research. The rate of articles sharing data did not change appreciably over time and was low regardless of funding source. Methods for sharing data varied widely, both in data availability statements and in storage options.

Conclusions: These results indicate a need for improved agricultural data sharing and suggest an important role for librarians and data professionals in promoting best data practices to meet increasingly strict funder requirements.

Introduction

Data sharing is integral to successful research, and research sponsors increasingly require researchers to share their data and findings in open-access repositories. Shared agricultural data has many applications including crop and climate change models, informing cover crop rotations, or enhancing plant breeding lines. Agricultural researchers who follow data sharing best practices advance the state of agricultural research since data sharing makes the research results more reliable and reproducible. However, many researchers are still concerned about sharing and making their data publicly available (Fecher et al. 2015).

To examine the data sharing practices of agricultural researchers at their institutions, the Agriculture Network Information Collaborative (AgNIC) Data Working Group conducted a systematized review, as defined by Grant and Booth (2009). AgNIC is a partnership of institutions with a mission of advancing agricultural information practices, including data related practices. The findings from this systematized review discovered trends in data sharing behaviors that may be useful for librarians or data professionals interested in helping agricultural researchers at their institutions share datasets and promote best data practices.

Research Questions

This study investigates what the published research shows about the data sharing practices of agricultural researchers and seeks to answer the following questions:

1. What are agricultural researchers' data sharing practices over time?
2. In which journals and with which publishers are agricultural researchers publishing?
3. Where are the researchers depositing their data sets?
4. In which section(s) of the published article do researchers indicate access to their data?
5. Which funding sources are used the most by agricultural researchers who shared their data?

Literature Review

Benefits and Challenges of Data Sharing

Open science aims to make science reliable, replicable, and reproducible and sharing data is one action that scientists can take to facilitate this movement (Allen and Mehler 2019). Data sharing increases transparency in research results, improves the integrity of research, and, in some cases, may provide time efficiency for data re-users (Pronk 2019). The ability to replicate published results adds credibility to the research while allowing researchers to build upon previous research to enhance science. In addition, data sharing impacts and benefits the community, as it can influence policy (Williams et al. 2019).

Researchers often reuse data to compare with their new data, jump start a new project, or relook at data using new research questions (Pasquetto et al. 2019). By doing so, they can imagine and implement their project more efficiently and reduce their project time. Depending on their field of study, scientists reuse data to discover the trends in a field site, compare real-world data with their experimental data and plan their data collections accordingly (Pasquetto et al. 2019). In a survey of researchers worldwide, Tenopir et al. (2020) found about a third of survey respondents in the agriculture and natural resources fields regularly used secondary data.

Barriers to sharing data often discourage researchers from putting more effort into finding ways to make their data open and accessible. A best practice for agricultural data includes creating minimal datasets with metadata and data dictionaries to make data useful and interoperable (Moore et al. 2022). However, metadata creation for datasets often depends on the external repository's guidelines (Cooper 2021) and researchers who indicated needing help with metadata quality proposed having automated metadata creation and better searchability of data in repositories (Donaldson and Koepke 2022).

Lack of incentive to share data, concern about misuse of data, and stability and security of data in repositories are some of the other concerns of researchers (Allen and Mehler 2019; Borgman and Bourne 2022; Donaldson and Koepke 2022; Perrier et al. 2020). These barriers are often caused by a lack of trust and compounded by the burden of cleaning and preparing the data for sharing (Williams et al. 2019). Incorporating more robust data review into the peer review process and creating appraisal processes to ensure high standards of data quality can mitigate some of the mentioned challenges (Moore et al. 2022).

As the data ecosystem has grown more diverse, locating data repositories and reusing primary data has become increasingly challenging for researchers (Wilkinson et al. 2016). Hedgepeth et al. (2023) found overwhelming evidence that the United States Department of Agriculture (USDA) researchers would like more help complying with federal data sharing policies, including suggestions for acceptable repositories to store their data for the sharing aspect of these mandates. In another study, Tenopir et al. (2020) found that a third of respondents in the fields of agriculture and natural resources experienced difficulty locating external

repositories to store their data. Concerns also exist about long-term maintenance of external repositories and the potential future losses in funding (Cooper 2021).

Along with sharing data and locating appropriate repositories, finding links to data shared in published articles may be challenging. Researchers, such as crop science faculty at the University of Illinois or agriculture researchers at The Ohio State University, often shared their data in PDF format as supplementary material or in summary form in the article without sharing the raw data (Williams et al. 2019; Williams 2016; Diekmann 2012). Additionally, Williams (2016) found that the majority of agricultural journals sampled did not supply individual DOIs for supplementary files. While Williams (2016) did not examine the prevalence of inaccessible supplements due to inoperative hyperlinks, they did find that approximately 1% of the examined articles referred to supplemental materials hosted by the journal which were unable to be located on the publisher website. As seen, benefits of sharing data are significant; however, the barriers in data sharing create many obstacles for researchers and impact their data sharing practices.

Data Sharing Practices

In an effort to remove some of the aforementioned barriers, Wilkinson et al. (2016) proposed Findability, Accessibility, Interoperability, and Reusability (FAIR) Guiding Principles. These guidelines have gained traction over the years and are practical and easy-to-use strategies for researchers to manage their data. Researchers are gradually adopting the practice of sharing their data in open repositories. Although the benefits of sharing data are known and support for data sharing exists, researchers often hesitate to make their data available (Fecher et al. 2015).

Researchers use data management plans and institutional repositories; however, their practices are inconsistent across projects (Donaldson and Koepke 2022). In some papers, data are made available for readers, while in other papers, data sharing practices are lacking. This trend is seen across disciplines. While 334 of 487 clinical trials published in *JAMA (The Journal of the American Medical Association)*, *The Lancet*, and *NEJM (The New England Journal of Medicine)* declared their intention to share their data, only two articles made their deidentified data publicly available (Danchev et al. 2021).

Researchers who share their data expressed that they request to retain rights and receive attributions for their data (Pasquetto et al. 2019). In other cases, researchers were not sharing data that were associated with private funders, or because they felt the data's value was time-bound and thus not worth sharing (Perrier et al. 2020; Cooper 2021). Although researchers acknowledge the value of reusing agricultural data, they felt that usefulness of small-plot research is short-lived (Diekmann 2012).

As seen, researchers struggle with various requirements of data sharing practices, yet many do not consider libraries and librarians as experts in data management who can help them with their various data sharing needs (Donaldson and Koepke 2022; Pouchard and Bracke 2016). The data sharing practices of researchers

highlights the need for educational outreach regarding benefits of sharing data and its impact on advancing science.

Methodology

Search Strategy

We systematically searched multiple databases to retrieve all published articles related to the research questions. Databases selected for this search were the Web of Science Core Collection (Clarivate Analytics), Scopus (Elsevier), CAB Abstracts, CAB Abstracts Archive, and CABI Full Text (all searched on CABI website) to identify peer-reviewed publications. We did not include grey literature in this search strategy as our research questions focused on peer-reviewed literature. The search string included the terms soybean, soybeans, and corn, which were selected as these are common crops researched across all the institutions included in this study (see Appendix A).

The search was limited to articles published by authors at Kansas State University; North Dakota State University, Fargo; Texas A&M University; Virginia Tech; University of Maryland, College Park; University of Tennessee, Knoxville; and the United States Department of Agriculture. Institutional affiliation fields were used in the Web of Science Core Collection and Scopus databases, and institutional names were entered as keywords in the CABI databases. These institutions were selected as they corresponded with the affiliations of our research team at the time the search was performed.

We limited searches to peer-reviewed articles published between 2017 and 2022. We selected 2017 as our lower date cut-off because it was four years past the White House Office of Science and Technology Policy (OSTP) “Memorandum on Increasing Access to the Results of Federally Funded Research” (Holdren 2013). The Holdren memo provided time for grant funders and researchers to adapt to the new data sharing guidelines and we hoped to see an increasing number of articles sharing their data. The search was performed in August 2022.

Due to the extreme volume of articles retrieved, a statistician generated a random sample of articles for the team to review. A stratified random sample was created using Statistical Analysis Software (SAS). We piloted our screening process and recorded the time reviewers took to review five articles. Based on the pilot, we determined that our full-text review should take six months with a one-to-two-hour time commitment per week for each reviewer. We calculated 1,250 articles as the ideal sample size to meet these parameters. This random sample represents the same distribution across the publication years as what appears in the full set and excludes the titles that did not have a year listed in the metadata. Afterward, articles were uploaded to Covidence, an online tool for managing systematic review processes, to follow the article screening and selection process.

Eligibility Screening

The review team consisted of seven librarians from the AgNIC Data Working Group. Each article was reviewed in full text by two members to determine whether the article needed to be included in the study based on the inclusion or exclusion criteria. Consensus for conflicting reviewer decisions was reached in bi-weekly meetings with the full research team. Articles were included in the study if they were peer-reviewed articles focused on corn or soybean production, produced primary data, were published between 2017 and 2022, and had at least one author from one of our selected institutions or had received funding from the USDA. Articles were excluded from the study for the following reasons, in ranked order:

- Foreign language - Article presented in a language other than English.
- Out of scope - Articles that were not about the production of corn or soybeans as a central topic.
- Not data - Articles that did not feature or use data. Examples: literature reviews, opinion papers, theoretical papers, plant registrations, etc.
- Unclear data source - Articles where no clear data source was described. In other words, it was unclear if the data were collected by the authors or if it was a secondary data source.
- Models - Articles that used primary or secondary datasets to input into a model to predict certain phenomena. We considered this for any type of model (e.g., machine learning, phenotyping, etc.).
- Secondary data - Articles that utilized datasets from other studies, collected by other authors. If data were collected by the same author but used in other articles, we would consider it primary data.
- Meeting abstract - Abstracts for conferences or meetings. Not a peer-reviewed primary journal article.
- Methods - Methodology papers that detailed the methods used to collect data, rather than the data results.

The team applied exclusion tags to each excluded article in Covidence, evaluating in the order shown above. After completing the screening process, all articles meeting the inclusion criteria were moved to the data extraction phase.

Extraction Steps

Extraction took place during both the full-text review and as a separate step. In both stages, two reviewers extracted information from each article, and conflicting reviewer decisions were resolved by the full research team. Initially, during the full-text review, tags were applied to included articles to record sharing status and sharing location. Reviewers captured funding agency and external repository names in a Google Sheets spreadsheet.

It was then determined that the research team wished to capture more robust data from included articles, so a formal extraction stage was performed in Covidence. An extraction form was used to collect the following information: the article's title and URL/DOI; year of publication; the journal's name; the journal publisher's name; author affiliations (matching our selected institutions); funding agency (if any); the primary topic of the article; if the article utilized secondary data in addition to the data they produced; if the article shared full data files or only summary data; the section of the article specifying access to the shared data; where the article authors shared their data; external repository name (if any); if data were available only by request; and if a DOI was linked to the shared data. During extraction, we defined sharing status and sharing location as such:

- Did not share: while the article generated original data, it did not share any data, or it shared partial data through PDF or Word files.
- Summary data: the article only shared summarized data used to create charts or tables, either in CSV or other ready-to-use spreadsheet format without sharing the raw or minimally processed data.
- Full dataset: the article shared raw or minimally processed data through CSV or other spreadsheet files either directly with the journal or an external repository.
- Shared with journal: the article shared either summary data or the full dataset directly through the journal website.
- External repository: the article shared data through an external repository.

The data collected by the research team in this phase were utilized for our data analysis.

Data Analysis

Data were exported from Covidence in CSV files for analysis. These files were first analyzed in RStudio Version 4.2.3. This platform allows for coding, debugging, editing, saving, and viewing results of statistical computing in one tool. For this article, RStudio was used to obtain descriptive statistics of the results reported in the extraction form. Functions such as frequencies and percentage calculations were used to categorize and describe the dataset. Functions such as delimitation were used to turn fields with multiple values (such as unique funders) into individual items that could be counted and described.

While RStudio has the capacity to generate graphics, including charts, the complexity of the data and the multiple variations of analysis (e.g., by funder type and year or by access specification location) led us to create graphics in Google Sheets by consolidating the CSV and R-analyzed summary values into smaller tables to feed charts and graphs.

Results

Selection of Studies

The search strategy retrieved 8,918 articles, and 4,532 articles remained after deduplication. A random sample of 1,250 articles used for eligibility screening was created from this set. A total of 267 articles matched the scope of the study and were included for analysis. Appendix B shows the results of the search and eligibility screening process. The data generated by this study are archived in the Open Science Framework and are available at <https://osf.io/pghtu>.

Ninety-eight different sub-topic focus areas within corn and soybean science were examined within the included studies. The most common sub-topics were soybean genetics (62 articles), cropping systems (15 articles), corn genetics (15 articles), corn cropping systems (12 articles), corn pests (7 articles), soybean production (6 articles), soybean growing conditions (6 articles), soybean genomics (6 articles), and soybean cropping systems (6 articles).

Data Sharing and Locations

In assessing data sharing practices, it was important to determine if the data being used for a study were primary or a combination of primary and secondary to assess the extent of data sharing for specific works. Sixty-nine (25.8%) studies included in the analysis produced primary data while also using secondary data from other authors. Most studies (206 or 77%) did not share their data, while 26 (9.7%) provided summary data only. Thirty-five (13.1%) of the included studies provided access to their full dataset. Twelve of the 206 articles that did not share data (4.5% of the total included articles) specified that access to the data were available by request only. Eleven articles (4.1%) provided a DOI for their data, making access easy for readers. Twenty-eight (10.5%) articles shared their data via the journal they published in (through supplementary materials/files section or similar) while twenty (7.5%) shared in an external repository. Thirteen (4.9%) articles shared data in both the journal and an external repository. Of those that used repositories, eight (6.7%) used more than one repository. The frequency of specific repositories used can be seen in Table 1.

As seen in Figure 1, data sharing practices remained consistent over the time period examined. A large majority of studies in any given year did not share their data, while either full or summary data sharing generally accounted for less than a quarter of articles published in a year.

Sharing locations of the data, as demonstrated in Figure 2, generally did not see large shifts during the five-year inclusion period. Again, the majority of articles did not share their data, while fairly equal amounts of sharing occurred in journals and repositories, respectively. The number of articles that shared data in both journals and repositories started small but increased from 2019 to 2021 before dropping off again in 2022.

Table 1: Frequency of Repositories used by Studies that Shared Data.

Repository	Frequency
NCBI*	15
Figshare	15
GitHub	4
Dryad	2
SoyBase	2
Ag Data Commons	1
European Nucleotide Archive	1
Purdue University Research Repository	1
Texas A&M University OAKTrust Digital Repository	1
USDA-ARS Agricultural Collaborative Research Outcomes System (AgCROS)	1

* NCBI includes GenBank, Gene Expressions Omnibus, BioProjects, Information Sequence Read, etc.

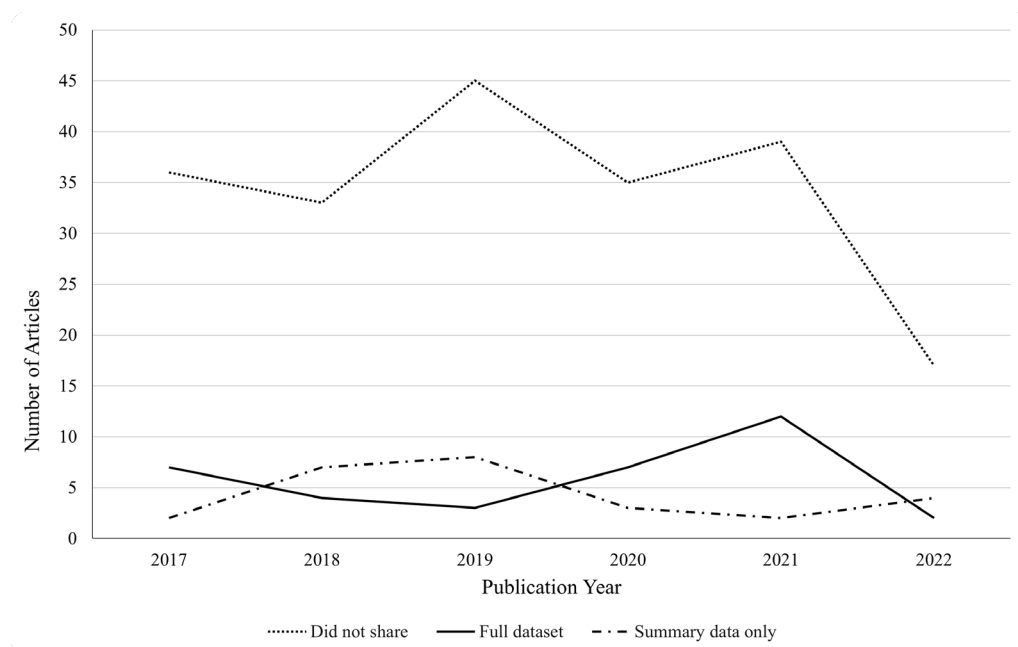


Figure 1: Data Sharing Over Time.

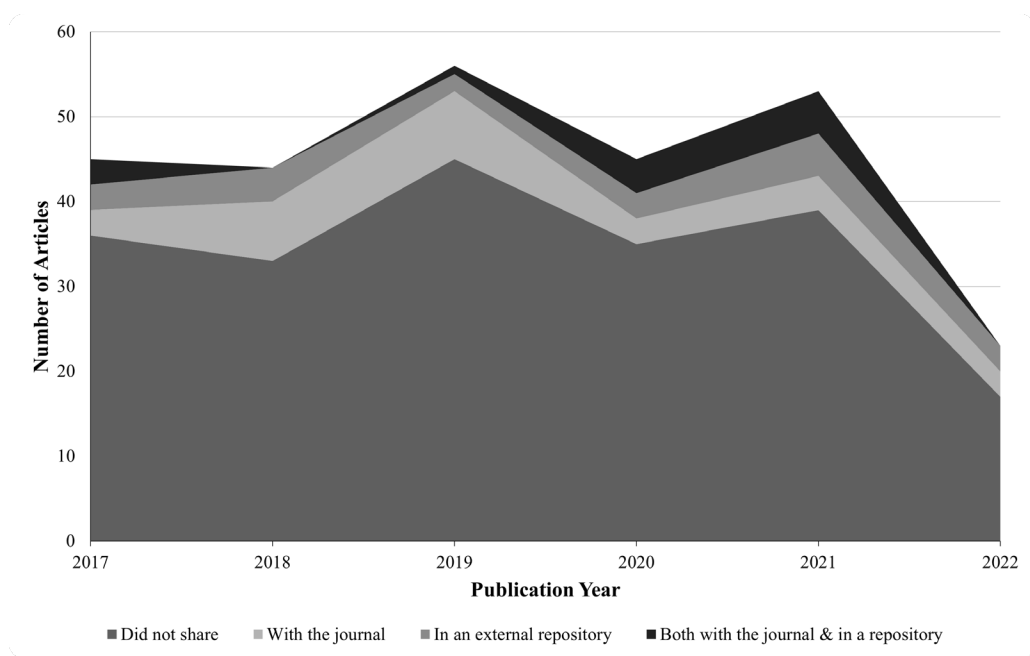


Figure 2: Sharing Locations Over Time.

Table 2: The Ten Most Commonly Used Funders for Included Studies. The majority of these were funded by the USDA, which is composed of multiple agencies geared toward supporting both intramural and extramural agricultural research.

Funder	Number of articles
USDA Agricultural Research Service	54
USDA National Institute of Food and Agriculture	46
United Soybean Board	33
USDA*	18
United States National Science Foundation	14
Tennessee Soybean Promotion Board	11
North Central Soybean Research Program	9
Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)	8
Conselho Nacional de Desenvolvimento Científico e Tecnológico	7
Kansas State University	7
* USDA is entered separately here for any articles that do not specify a particular funding agency within the department.	

Funders

The most common funders can be seen in Table 2. We identified 153 unique funders acknowledged across the 267 included studies. Sixty-one of the included studies reported no funding sources, while 131 included multiple sources of funding. Eight of these funders were private companies.

Over the included time period, studies that did not list funding sources decreased, and alongside this change there was an increase in studies with combined funding from U.S. federal sources and non-U.S. federal sources. These trends can be seen in Figure 3. U.S. federal funding, on its own, decreased in 2020, but increased in 2021.

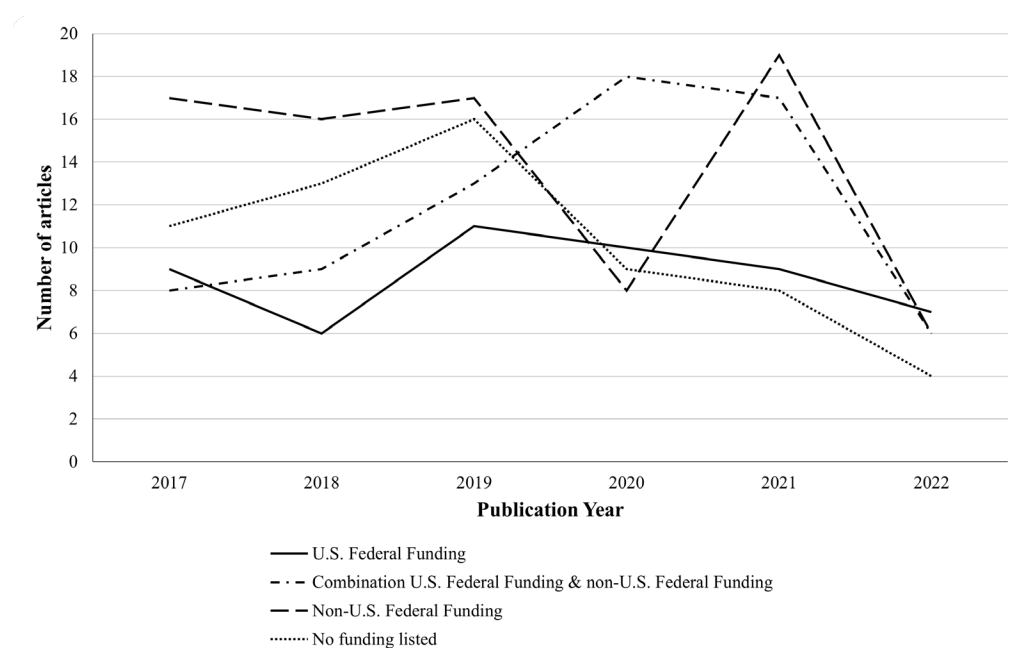


Figure 3: Funding Type Over Time.

Most of the studies included in this analysis did not share their data. This remained the case when the scope was narrowed to those with U.S. federal funding. Data sharing was low each year for articles with federal funding whether sharing data in an external repository, sharing with the journal, or sharing both with the journal and in a repository. Figure 4 shows the sharing location trends over time.

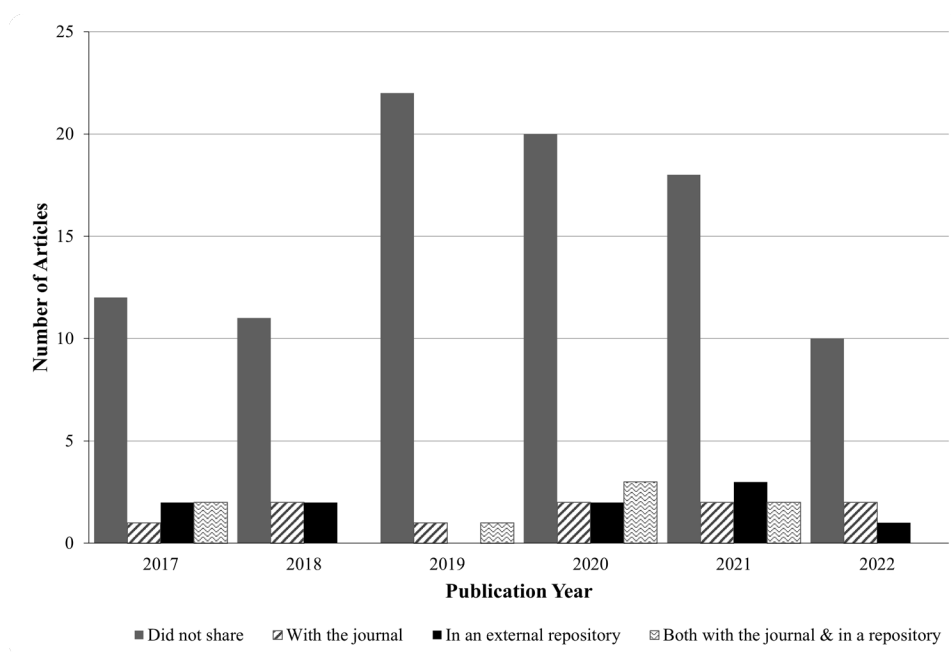


Figure 4: Data Sharing of Federally Funded Research Over Time.

Journals and Access Specifications

The included studies spanned 120 different journals and 31 different publishers. Nearly 70% of all included studies used five common publishers: Wiley (81 articles), Elsevier (34 articles), Springer Nature (31 articles), MDPI (21 articles), and Taylor & Francis (19 articles). The most frequently used journals were *Agronomy Journal* (26 articles), *Crop Science* (10 articles), *Frontiers in Plant Science* (9 articles), *Journal of Plant Nutrition* (7 articles), and *The Plant Genome* (7 articles). Many journals appeared less than five times among the included publications.

Data availability was indicated in different sections of the 61 articles that shared either summary data or the full dataset. Twenty-four articles contained statements about data availability in multiple locations. Of these, eighteen articles used the supplementary materials section, and seventeen articles used the data availability section to specify data availability, with the majority utilizing both of these sections. Six articles solely used the data availability section to state availability. Seven articles stated availability in the results section and four in the methods section.

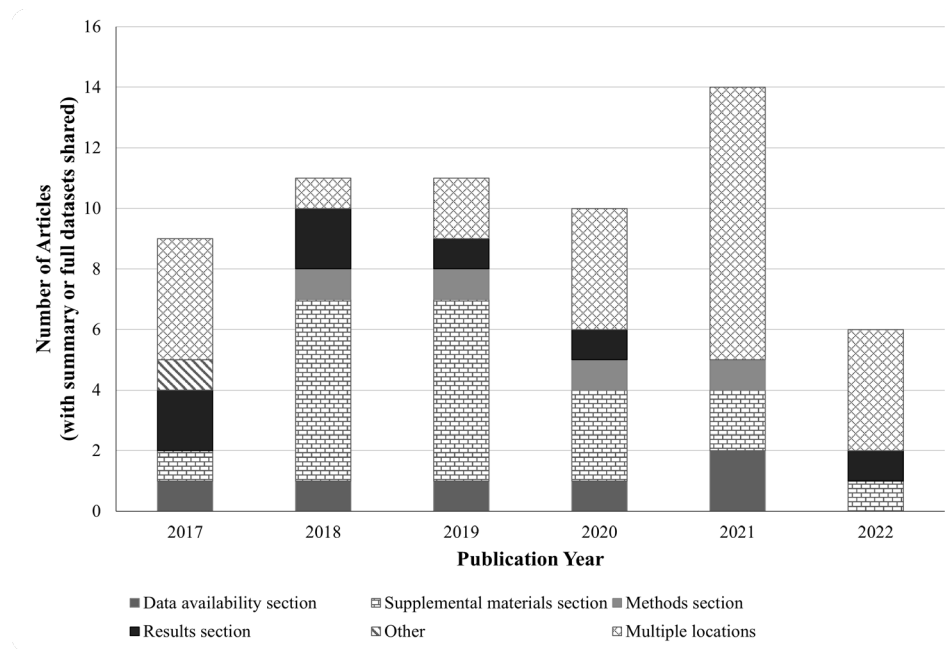


Figure 5: Access Specifications Over Time.

The section of the article where authors specify access to their shared data did shift over time. Figure 5 visualizes these trends for articles that shared either summary data or the full dataset. In 2017, four out of nine articles specified access to data in multiple article sections, with another two articles specifying access in the results section. In both 2018 and 2019, six out of eleven articles specified access to data in the supplemental materials section. In 2020, the trend shifted with four articles specifying access in multiple locations and three other articles specifying access in the supplemental materials section. Afterward, in both 2021 and 2022, the multiple location approach was the primary method for specification, with nine articles in 2021 and four articles in 2022.

Discussion

Data Sharing Practices

This review did not find a notable difference between the number of researchers who chose to share data through the journal or with an external repository, nor was there a standard repository for agricultural data. This may be due to the variance of data types collected by agricultural researchers, even in the narrowed sub-discipline of crop science, so that an obvious or single choice for storage was not apparent, such as a relevant disciplinary repository. Our findings align with the study conducted by Hedgepeth et al. (2023) who assessed more than 200 locations containing open agricultural data, found an educational need for identifying suitable repositories, and compiled 20 recommended data sharing options for a subset of USDA research.

Wiley, Elsevier, Springer Nature, MDPI, and Taylor & Francis were the five top publishers of the agricultural research papers in this review. These five publishers are wide-reaching and produce a large portfolio of journals. All five publishers have a data sharing policy which either encourages the authors to publish their data or gives guidelines on how authors can share their data (Elsevier 2024; MDPI 2024; Springer Nature 2024; Taylor & Francis 2024; Wiley 2024). The impact of these policies on data sharing practices is not clear as this impact was not within the scope of this review.

The included articles in this review represent 120 different journals, suggesting crop science researchers publish in a variety of places, despite a small number of journals being particularly prevalent. This variety may also contribute to the lack of standard data sharing practices, as each journal may have its own standards for creating supplementary materials or indicating where data are shared. Most articles in this review received their funding from USDA Agricultural Research Service, although it should be noted that USDA authors may not list funding because the work is part of their employment and therefore this number could be more than what was found in this research.

The majority of the agricultural researchers in our sample, including those who received U.S. federal funding for their research, did not share the location of their full dataset. While the OSTP 2013 memo (Holdren 2013) stipulated federally funded research must make its outcomes available and called for data sharing, the actual practices five to ten years out show that this stipulation is often overlooked. When researchers did share data, they did not always do so in a way that enabled others to easily find or reuse the data. Despite the relatively high use of supplementary materials sections and data availability statements, there was no standard section within published articles where authors specified data access. This requires readers to go through the full article to determine whether data are shared, adding an extra barrier to retrieving shared data. This highlights the interconnectedness of the FAIR principles of findability, accessibility, interoperability, and reusability as outlined by Wilkinson et al. (2016). If data are findable, they will be accessed, and if they are accessed, they will be reused.

Future Research

This study is not a comprehensive review of all the articles retrieved by our search. We performed a broad search to capture a wide range of articles, and we only examined articles focused on corn and soybean production. Data sharing trends may be different across other agricultural sub-disciplines, such as animal science, forestry, entomology, etc. Additionally, we performed our search in mid-2022, therefore, we retrieved fewer articles from 2022 than other years, and as seen in several charts, our results were always lower for 2022.

Looking ahead, changes in funding agency requirements due to the OSTP memorandum on “Ensuring Free, Immediate, and Equitable Access to Federally Funded Research,” also known as the Nelson memo, may influence the status quo for agricultural research data sharing, as it prompts federal agencies to revise their

policies on sharing research outputs (Nelson 2022). Once these changes are implemented by December 31, 2025, researchers receiving federal funding are required to make peer-reviewed publications and associated data available to the public at no charge, with certain exceptions.

The memo requires the agencies to follow specific expectations such as uploading the data in acceptable repositories, following machine-readable formats, or adhering to compatibility standards with assistive technology for users with disabilities (Nelson 2022). As a result, researchers may face more stringent parameters than in previous guidelines, which could act as barriers. On the other hand, these changes may also lead to increased standardization in agricultural data sharing with potential impacts on data discoverability and reuse.

Questions surrounding the impact of changing funder mandates on agricultural data are possible areas for future research. In particular, once the USDA revises and implements its policies in accordance with the Nelson memo, an exploration of data sharing pre- and post- implementation could be enlightening. Will there be an increase in data sharing, standardization of practices, or adherence to the FAIR principles (Wilkinson et al. 2016)? Likewise, the impact of journal requirements and institutional policies on agricultural data sharing deserves investigation. Research into the prevalence and characteristics of data sharing in segments of agricultural research other than corn and soybean production may also yield important insights into factors that influence the data sharing practices of agricultural researchers.

Implications for Practice

Given the marked lack of agricultural research data sharing that our study found in recent publications, and since the impetus for data sharing may increase in the future, how can librarians and data specialists assist agricultural researchers to improve in this area? One way is to stay up to date with relevant funder, publisher, and institutional data sharing mandates and their interpretations. Another way is to partner with institutional offices and individuals who support and handle data (Association of American Universities (AAU) and Association of Public and Land-grant Universities (APLU) 2021; Bryant et al. 2020). The library's data specialist, institutional office of research, or office of sponsored programs can be valuable partners to agriculture librarians in coordinating and publicizing available sources of data sharing assistance for researchers.

Agriculture librarians may also be ideally placed to leverage their relationships with research faculty and graduate students to begin conversations about what "open" means for agricultural research and publishing, why open data matters, and what limitations or special considerations may exist in agricultural research. Finally, librarians serving these areas should prepare for a potential increase in requests for help with metadata creation, identifying repositories, and related data sharing and curation tasks. This may involve learning new skills, identifying currently available services for referrals, or coordinating with other colleagues and units to best utilize existing capabilities and efficiently provide the necessary assistance.

Conclusion

Data sharing enhances the integrity of research data and provides other researchers with secondary data they can build upon to advance research. By storing research data in open or institutional repositories, researchers can track usage to assess the impact of their studies. However, our research showed an obvious lack of data sharing in many articles about corn and soybeans published from 2017 to 2022.

Through a systematized review of agricultural articles authored by researchers at our institutions, we discovered that most studies (77%) did not share their data. When sharing data, researchers opted to share their data via journals and repositories. Data hosting by the journal was the most popular option, while NCBI and Figshare were the most frequently used repositories. Data availability statements, when included, were in various locations of the published papers such as the supplementary materials section or the results section, increasing the difficulty of accessing the data.

Many articles in this review received funding from the USDA or its agencies. Despite existing guidelines for sharing federally funded research data, this review shows that agricultural data sharing practices have not changed appreciably over the time period represented by this study and a large portion of articles are still not sharing their data. There is an ongoing need for data sharing education.

With the implementation of the Nelson memo in 2025, changes might occur in data sharing practices and future research can investigate the impact of this mandate on agricultural research. This review highlights the need for promoting these data sharing best practices among researchers and points out the integral role of librarians in introducing and encouraging these practices at their institutions.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgements

We would like to acknowledge and thank Nicole Juve and Cozette Comer for their advice and support in the early stages of this research, and Curt Doetkott for his assistance in creating the random sample of articles.

The content of this article is based on the presentation entitled "Data Sharing Practices in Agricultural Research: Findings from a Systematized Review" originally presented at [RDAP Summit 2024](#), available from Open Science Framework: <https://osf.io/pghtu>.

Data Availability

The data generated by this study are archived in the Open Science Framework and are available at <https://doi.org/10.17605/OSF.IO/PGHTU>.

This protocol was registered with the Open Science Framework Registries and is available at <https://osf.io/amw6e>.

Database Search Strings and PRISMA 2020 Flow Diagram are available under the article [Supplementary Files](#):

Appendix A: Search Strings

Appendix B: PRISMA 2020 Flow Diagram

References

- Allen, Christopher, and David M. A. Mehler. 2019. "Open Science Challenges, Benefits and Tips in Early Career and Beyond." *PLOS Biology* 17 (5): e3000246. <https://doi.org/10.1371/journal.pbio.3000246>.
- Association of American Universities (AAU) and Association of Public and Land-grant Universities (APLU). 2021. "Guide to Accelerate Public Access to Research Data." Association of Public and Land-grant Universities (APLU). <https://www.aplu.org/library/guide-to-accelerate-access-to-public-data/file>.
- Borgman, Christine L., and Philip E. Bourne. 2022. "Why It Takes a Village to Manage and Share Data." *Harvard Data Science Review* 4 (3). <https://doi.org/10.1162/99608f92.42eec111>.
- Bryant, Rebecca, Annette Dortmund, and Brian Lavoie. 2020. *Social Interoperability in Research Support: Cross-Campus Partnerships and the University Research Enterprise*. Dublin, OH: OCLC Research. <https://doi.org/10.25333/WYRD-N586>.
- Cooper, Kristen A. 2021. "Data Sharing Attitudes and Practices in the Plant Sciences: Results from a Mixed Method Study." *Journal of Agricultural & Food Information* 22 (1-2): 37-58. <https://doi.org/10.1080/10496505.2021.1891923>.
- Danchev, Valentin, Yan Min, John Borghi, Mike Baiocchi, and John PA Ioannidis. 2021. "Evaluation of Data Sharing after Implementation of the International Committee of Medical Journal Editors Data Sharing Statement Requirement." *JAMA Network Open* 4 (1): e2033972. <https://doi.org/10.1001/jamanetworkopen.2020.33972>.
- Diekmann, Florian. 2012. "Data Practices of Agricultural Scientists: Results from an Exploratory Study." *Journal of Agricultural & Food Information* 13 (1): 14-34. <https://doi.org/10.1080/10496505.2012.636005>.
- Donaldson, Devan Ray, and Joshua Wolfgang Koepke. 2022. "A Focus Groups Study on Data Sharing and Research Data Management." *Scientific Data* 9 (2022): 345. <https://doi.org/10.1038/s41597-022-01428-w>.
- Elsevier. 2024. "Research Data." Elsevier. <https://www.elsevier.com/about/policies-and-standards/research-data>.
- Fecher, Benedikt, Sascha Friesike, and Marcel Hebing. 2015. "What Drives Academic Data Sharing?" *PLOS ONE* 10 (2): e0118053. <https://doi.org/10.1371/journal.pone.0118053>.
- Grant, Maria J., and Andrew Booth. 2009. "A Typology of Reviews: An Analysis of 14 Review Types and Associated Methodologies." *Health Information & Libraries Journal* 26 (2): 91-108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>.

- Hedgepeth, Shavonne, Erin Antognoli, Sara Duke, and Alison Rehfus. 2023. "What About Your Friends?": How a Collaborative Transdisciplinary Training Approach Supports FAIR Data Sharing Principles in Federally Funded Research." *Journal of eScience Librarianship* 12 (3): e755. <https://doi.org/10.7191/jeslib.755>.
- Holdren, John P. 2013. "Increasing Access to the Results of Federally Funded Scientific Research." Office of Science and Technology Policy. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf.
- MDPI. 2024. "MDPI Research Data Policies." MDPI. https://www.mdpi.com/ethics#_bookmark21.
- Moore, Eli K., Adam Kriesberg, Steven Schroeder, Kerrie Geil, Inga Haugen, Carol Barford, Erica M. Johns, et al. 2022. "Agricultural Data Management and Sharing: Best Practices and Case Study." *Agronomy Journal* 114 (5): 2624-2634. <https://doi.org/10.1002/agj2.20639>.
- Nelson, Alondra. 2022. "Ensuring Free, Immediate, and Equitable Access to Federally Funded Research." Office of Science and Technology Policy. <https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf>.
- Pasquetto, Irene V., Christine L. Borgman, and Morgan F. Wofford. 2019. "Uses and Reuses of Scientific Data: The Data Creators' Advantage." *Harvard Data Science Review* 1 (2). <https://doi.org/10.1162/99608f92.fc14bf2d>.
- Perrier, L., E. Blondal, and H. MacDonald. 2020. "The Views, Perspectives, and Experiences of Academic Researchers with Data Sharing and Reuse: A Meta-Synthesis." *PLOS ONE* 15 (6): e0234275. <https://doi.org/10.1371/journal.pone.0229182>.
- Pouchard, Line, and Marianne Stowell Bracke. 2016. "An Analysis of Selected Data Practices: A Case Study of the Purdue College of Agriculture." *Issues in Science and Technology Librarianship* 85 (December). <https://doi.org/10.29173/istl1691>.
- Pronk, Tessa E. 2019. "The Time Efficiency Gain in Sharing and Reuse of Research Data." *Data Science Journal* 18 (2019): 10. <https://doi.org/10.5334/dsj-2019-010>.
- Springer Nature. 2024. "Research Data Policy." Springer Nature. <https://www.springernature.com/gp/authors/research-data-policy>.
- Taylor & Francis. 2024. "Data Sharing." Informa plc. <https://taylorandfrancis.com/our-policies/data-sharing>.
- Tenopir, Carol, Natalie M. Rice, Suzie Allard, Lynn Baird, Josh Borycz, Lisa Christian, Bruce Grant, Robert Olenendorf, and Robert J. Sandusky. 2020. "Data Sharing, Management, Use, and Reuse: Practices and Perceptions of Scientists Worldwide." *PLOS ONE* 15 (3): e0229003. <https://doi.org/10.1371/journal.pone.0229003>.
- Wiley Author Services. 2024. "Wiley's Data Sharing Policies." John Wiley & Sons, Inc. <https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/data-sharing-citation/data-sharing-policy.html>.
- Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3 (2016): 160018. <https://doi.org/10.1038/sdata.2016.18>.

Williams, Sarah C. 2016. "Practices, Policies, and Persistence: A Study of Supplementary Materials in Crop Science Journals." *Journal of Agricultural & Food Information* 17 (1): 11-22.

<https://doi.org/10.1080/10496505.2015.1120213>.

Williams, Sarah C., Shannon L. Farrell, Erin E. Kerby, and Megan Kocher. 2019. "Agricultural Researchers' Attitudes Toward Open Access and Data Sharing." *Issues in Science and Technology Librarianship* 91 (June).

<https://doi.org/10.29173/istl4>.