



**Implementation Guide for Public Access to Scholarly Publications
and Digital Scientific Research Data
US Department of Agriculture – 25 January 2023**

1. How to Use This Guide

This guide is a resource to assist USDA agencies in implementing the USDA Departmental Regulation [Public Access to Scholarly Publications and Digital Scientific Research Data](#). It can also assist researchers and their organizations in understanding how they can comply with the related policies or procedures of their funding or home agency. It offers current interpretation of the Departmental Regulation but does not constitute official policy; please consult with agency officials if there are questions. When referring to the Departmental Regulation, we use “Policy” with a capital P, and “policy” or “policies” for any other formal or informal policies or procedures.

This guide provides general implementation assistance for the Department and its stakeholders and should be used in conjunction with any additional policies, processes, and implementation resources developed by USDA agencies or staff offices.

This guide will evolve as implementation progresses and more resources become available. Please provide comments or suggestions to <https://www.nal.usda.gov/ask-question>.

2. General Information and Policy Scope

a. Who must comply with the Policy?

All USDA mission areas, agencies, and staff offices, all USDA employees who engage in scientific research during official duties, and all awardees and contractors from non-USDA organizations who are engaged in USDA-supported scientific research must comply with the Policy. Non-USDA organizations are responsible for ensuring compliance by their researchers if they receive:

- Any direct funding from a USDA grant or cooperative agreement active in Fiscal Year 2023 or beyond
- Any direct funding from a USDA contract signed on or after October 1, 2022
- Any direct funding or arrangement, including in-kind support, from any USDA intramural research

For the definition of scientific research, and of intramural and extramural research, see the Policy Appendix B. Agencies may further clarify, using these definitions, what constitutes scientific research supported by their agency.

b. What research products are in scope of the Policy?

- Peer-reviewed journal articles
- Digital scientific research data that are available for public access (see DR Section 5a).

Software code and computer models, lab notebooks, and educational curricula are not covered by this Departmental Regulation, but making this type of information publicly available is encouraged,

especially if the information is associated with peer-reviewed publications and publicly available data.

c. Are USDA-funded research products that were developed or published prior to the Policy subject to the same requirements?

Peer-reviewed publications and digital scientific research data created prior to October 1, 2022, are not required to be made publicly available by this Policy but may be subject to other requirements (such as terms and conditions of awards) and are encouraged to be made publicly available.

d. How does this Policy align with other USDA public access policies?

Some USDA agencies or staff offices have already developed specific policies and procedures that overlap with the content of this Policy. This Policy supersedes any policies released prior to October 1, 2022. USDA agencies and staff offices are required to update their specific policies, procedures, and resources by July 20, 2023. Until those agency-specific policies or procedures are updated they should still be followed.

3. Requirements for Scholarly Publications

[PubAg](#) is the USDA public access archive system for scholarly publications maintained by the National Agricultural Library.

a. Publications in Scope of the Policy

(1) Are scholarly publications arising from USDA-funded scientific research prior to October 1, 2022, in scope of the Policy?

The Policy only applies to peer-reviewed journal articles released after October 1, 2022. Older publications are not in scope of the Policy but may be covered under other agency-specific requirements, such as the Agricultural Research Service's P&P 152.1 v7 or terms and conditions of existing NIFA awards.

(2) How does USDA define a peer-reviewed journal article or scholarly publication'?

See definitions in DR Appendix B. To further explain, a publication is peer-reviewed if it uses an assessment process widely used by scholarly journals to evaluate the quality of manuscripts submitted for publication. The peer-review process may vary among journals, but it generally entails an editor-in-chief who oversees evaluation of the submitted manuscript. It commonly includes reviews by subject matter experts who evaluate the methodology, assumptions, results, and interpretations underlying the manuscript. In concert with reviewers, the editor-in-chief decides whether to accept or reject the manuscript for publication or request further modifications. Pre-prints (e.g., pre-peer-review manuscripts posted publicly to a server such as [bioRxiv.org](#)) are not currently in scope for the Policy as they are not yet peer reviewed.

(3) Are conference proceedings in scope of the Policy?

Conference proceedings that are peer-reviewed full papers and published are within the scope of the Policy.

Conference abstracts alone (many of which do not undergo peer review) are not in scope of the Policy.

(4) Does USDA expect other categories of research publications produced with USDA funding to be made publicly accessible?

Other categories of research publications such as USDA-published reports or book chapters are not in scope of the Policy, unless they are linked to datasets that ARE in scope for the regulation. (See DR Section 5c(5) on requirements for research publications connected to digital scientific research assets covered by the Policy.) USDA encourages other categories of research publications to be made publicly available, but under current collection policies, they should not be submitted to PubAg (see below).

b. If a journal requires a subscription to read, does this mean their articles can't be made publicly accessible?

Although the published version of an article in a subscription-only journal may not be freely accessible on the publisher's website, the content of that article (e.g. the final, accepted manuscript) can still be shared through the PubAg repository (see below).

c. Does the Policy require a researcher to pay an article processing charge so their peer-reviewed journal article can be Open Access?

No, the Policy does not require researchers to publish in open access journals or to pay Open Access article processing charges. If a peer-reviewed publication is released with an open access license, then the published version may be submitted to PubAg, as sharing will not violate copyright. If a peer-reviewed publication is not released as open access, then the author's last accepted manuscript (just prior to publication) must be submitted to PubAg instead. Some journals make it easy for an author to download this version upon acceptance.

d. Can USDA-funded researchers include Open Access article processing charges in extramural funding budgets?

Yes.

e. How does a researcher submit a peer-reviewed publication to PubAg?

(1) How are scholarly publications submitted to PubAg, and whose responsibility is it to ensure that submission has occurred?

The PubAg submission website is available for all USDA-employed authors or their proxies (e.g. assistants): <https://submit.nal.usda.gov>. It will soon be available for all USDA-funded researchers outside USDA with an [eAuth account](#). Agencies may provide different requirements for who is responsible for submitting the articles and how they make their way to PubAg.

(2) If a USDA-funded scholarly publication is already publicly accessible through another federal catalog (e.g., PubMedCentral) is it in compliance with the Policy?

No. The publication must still be submitted to PubAg. NAL is working with PubMedCentral to streamline this process.

- (3) If a USDA-funded scholarly publication is already publicly accessible in a location other than a federally maintained catalog (e.g., posted on laboratory or university website), is it in compliance with the Policy?**

No, a full text version of the publication must be in PubAg to be in compliance.

- (4) What is the difference between the final peer-reviewed, accepted manuscript and the final published article? Which should be submitted to PubAg?**

USDA follows the Department of Energy's Office of Scientific and Technical Information definition for an accepted manuscript which is "the version of a journal article that has been peer-reviewed and accepted for publication, and includes changes made by the author during the peer-review process. The content is the same as the published article, but it does not have the final formatting and copy-editing done by the publisher in the final published version, and it is not a reproduction copy of what appears or will appear in the journal; it is not a reprint of the published article." (<https://www.osti.gov/what-accepted-manuscript>, 2015).

If the final published article is Open Access (e.g., the author has paid for the article to be made available under a Creative Commons Attribution license, such as <https://creativecommons.org/licenses/by/4.0/>) then that version may be submitted to PubAg. Otherwise, the author's version of the accepted manuscript prior to publication must be submitted.

f. Digital Persistent Identifiers for Scholarly Publications

- (1) What digital persistent identifiers are acceptable for scholarly publications, and how are they assigned?**

A Digital Object Identifier (DOI) is the most common digital persistent identifier assigned by journal publishers to their scholarly articles. Other identifiers such as Handles may be used instead if the journal issues those and a DOI is not available. A digital persistent identifier is unique, persists forever, and resolves (provides information) in a standard, publicly accessible way such as putting it into a web browser.

- (2) For scholarly publications that are not assigned a digital persistent identifier by the journal publisher, how will these publications be assigned an identifier to be compliant with the Policy?**

Full text submissions without an assigned DOI will be given a persistent URL through the NAL handle server.

4. Requirements for Digital Scientific Research Data Assets

a. Data Assets in Scope of the Policy

- (1) What types of USDA-funded data assets are in scope of the Policy?**

All digital scientific research data assets that arise from unclassified intramural or extramural scientific research supported wholly or in part by USDA are in scope of the Policy (DR Section 5a). While all USDA-funded scientific research is subject to the data management plan requirement under the Policy (DR Section 5c(1)), some digital scientific research data assets are

exempt from requirement to make data publicly accessible (see below). Note that the research still needs data management plans, and scholarly publications based on these data still need to comply with the Policy.

Data assets that are not collected in support of intramural or extramural scientific research funded by USDA, including those arising from routine analysis in support of program operations, are not considered digital scientific research data assets. These non-research data assets are outside the scope of the Policy.

The term “asset” can be understood to be that collection of data, or that dataset, that has some value. So one might interpret the Policy to apply the versions of data that are valuable for the research enterprise. Data management plans can make this clear in specific cases, but often the most raw data and the most cleaned up data have more value than intermediate steps in an analysis pipeline.

(2) Are digital scientific research data assets arising from scientific research funded by USDA prior to October 1, 2022 in scope of the Policy?

The Policy applies to digital scientific research data assets arising from scientific research funded on or after October 1, 2022, as defined in DR Section 5a.

USDA encourages data authors to make digital scientific research assets produced prior to the effective date of the Policy publicly accessible when appropriate. Additionally, some digital scientific research data assets arising from scientific research funded prior to October 1, 2022 are subject to data management and public access requirements from agency-specific policies. For example, prior agency-specific policies and procedures, such as ARS P&P 630.0 and NIFA requirements for data management plans in some competitive awards, should still be followed for digital scientific research data assets that are outside the scope of this DR.

b. Data Management Plans (DMPs)

The Policy requires all scientific research funded by USDA to be accompanied by an agency-approved Data Management Plan (DR section 5c(1))

(1) What are the required components of a DMP, and how should it be formatted?

DMPs for scientific research covered by the Policy should be consistent with [guidance maintained by the National Agricultural Library](#). Some USDA agencies and offices also may have specific policies and procedures that describe required DMP components and format, how many and at what level data management plans are required, and the process by which these plans will be submitted and reviewed.

(2) At what level should a DMP be prepared (e.g., project level, unit-level, or other)?

This depends on the implementation of the specific USDA agency.

(3) How are DMPs reviewed and approved?

USDA Agency Administrators and Staff Office Directors are responsible for ensuring that a policy or procedure exists for reviewing and approving data management plans, which may include

agency- or office-specific training and resources for reviewers. NAL offers a [checklist for DMP review](#). For example, NIFA proposal and ARS project plan peer reviewers are among those responsible for review of DMPs.

(4) Can researchers include data management and sharing costs in grant budgets?

Yes.

(5) Are other USDA resources and services available to support the development and review of DMPs?

The National Agricultural Library maintains a variety of resources to help researchers comply with the Policy and to support best data management practices. The [Data Management Planning Services page](#) contains information about the [expected DMP format](#), a [DMP review checklist](#), information about how to [find a suitable subject-specific data repository](#) to fulfill public access requirements, and sample DMPs for researchers to use as a springboard to create their own. We also provide [webinar recordings](#) that walk through the process of creating a DMP and depositing data into the Ag Data Commons.

Some research organizations, including ARS and some universities, offer support to develop and review DMPs. Peers doing similar research can often help make sure that a DMP adheres to community best practices.

c. Scope of Public Access Requirements for Digital Scientific Research Data Assets

(1) If scientific research is not expected to produce digital scientific research data assets that must be made publicly accessible (i.e., all expected data assets meet the exemption criteria listed in DR Section 5c(2)), does the research require a data management plan?

Yes, all scientific research expected to produce digital scientific research data assets must be accompanied by a reviewed and approved data management plan (DR Section 5c(1)). The data management plan must indicate any categories of digital scientific research data assets that are exempt from public access requirements.

(2) Do the public access exemptions listed in DR Section 5c(2) apply to the whole data asset or specific elements of a data asset?

If a digital scientific research data asset meets public access exemption criteria under DR Section 5c(2), the entire data asset is currently exempt from public access requirements under the Policy. That said, if particular elements can be omitted, aggregated, or anonymized and the rest of the dataset can be publicly available, this can provide transparency and value, and avoid unnecessary requests for access.

(3) Computational models are listed as exempt from public access requirements under DR Section 5c(2a(1)). If computational models are central to a research project, are they still exempt from public access requirements?

Computational models are not required to be made publicly accessible under the Policy, though researchers are encouraged to make these assets (including code, parameters, and outputs) publicly accessible when appropriate, where they can then be linked to model input data and

resulting scholarly publications. Model input data and resulting scholarly publications are required to be made publicly accessible if these research products are in scope of the Policy.

The Ag Data Commons accepts all USDA-supported research data and data products, including code, software tools, and models. Note that executable files cannot be accepted into the Ag Data Commons as a direct upload, and source code may be best maintained in a source code repository. Researchers may create records in the Ag Data Commons that point to executable files or source code contained in a more suitable repository.

d. Selecting a Data Repository

(1) How would a researcher choose an appropriate data repository for publishing their digital scientific research data assets?

NAL and many academic libraries offer [reference or research services](#) to assist in finding repositories. NAL's [Find a Data Repository](#) page includes resources to help researchers locate and choose appropriate subject-specific repositories for their data. Several strategies outlined include re3data.org, FAIRsharing.org, Repository Finder, DataSeer, OpenDOAR, and journal publisher repository recommendations. A downloadable Word document provides screen captures and more detailed instructions for using these resources. Researchers can consider the [Ag Data Commons](#), which meets requirements of the Policy and provides a generalist web submission form containing information required for a variety of USDA and scientific applications.

Can researchers use the Ag Data Commons as a data repository to comply with this policy? Yes.

The Ag Data Commons accepts most USDA-supported research data to allow researchers to comply with the Policy and meets requirements for public access, storage, and persistent identifiers. NAL encourages researchers to use subject-specific repositories according to community best practice if they meet requirements for public access, storage, and persistent identifiers. In this case, researchers should create an Ag Data Commons catalog record that points to the subject-specific repository. Provision for storage and release of large datasets should be worked out in agency-approved data management plans; the Ag Data Commons team should be included in development of these plans if ADC is the planned repository.

di. Machine readability requirements

(1) What is machine readability and why is it important?

Machine readability means that an asset such as a dataset is formatted in a way that it can be easily interpreted by a machine, such as a computer program, without human intervention. Examples of machine-readable formats are CSV (comma separated values), XML (extensible markup language), or RDF (resource description format). This is important because a machine-readable asset can usually be transformed into something human readable, so the same asset can be used many ways. The reverse is rarely true – an asset in a human-readable format such as a PDF or a typical chart image often requires many human decisions to ensure that information isn't lost or garbled when transforming to a machine-readable format.

(2) How do researchers ensure that their data are machine readable?

NAL provides [basic guidance](#) on how to format datasets as CSV (comma separated value) files. Researchers can find a webinar on [making data machine-readable](#), but specific implementations will vary by subject area. Note that machine-readable versions of data are to be provided but there are often good reasons to also make other versions available, too (e.g. Excel Workbooks in addition to CSV files).

f. Timeline for public access to digital scientific research data assets

(1) When must digital research data assets be made publicly accessible?

The DR allows a 12-month delay on public accessibility after the publication of a related scholarly publication, or the end of the performance or funding period if there is no related scholarly publication. Ideally, a data asset would be submitted to a repository at the earliest opportunity and public release could be delayed if the repository allows that (Ag Data Commons does).

(2) What does the ‘end of the period of performance or funding period’ mean?

This should be defined by the relevant USDA agency or office-specific policies.

g. Persistent Identifiers for Digital Scientific Research Data Assets and Associated Research Publications

(1) How do researchers obtain persistent identifiers (e.g., DOI) for USDA-published scientific research articles and reports connected to digital scientific research data asset (as required by DR Section 5c(5))?

Please contact the PubAg team at pubag_curator@usda.gov to discuss arrangements for this.

h. Submitting Metadata to the Ag Data Commons

(1) Submission process and requirements for metadata catalog entry

The Ag Data Commons requests a handful of essential information (metadata) and provides a variety of other information fields to add value to research data. In addition to storing data, the Ag Data Commons can link data to scholarly publications and related content for greater context. Researchers can read documentation about the submission process as well as essential and other available information fields in the [Ag Data Commons Data Submission Manual](#). The webinar [Submitting Data to the Ag Data Commons](#) walks viewers through the submission process in real time.

(2) Is a manual submission of metadata to Ag Data Commons always required?

Currently, yes. There are plans to harvest metadata from other repositories. If a researcher is not sure if there will be an ingestion process for their preferred repository, please contact Ag Data Commons at NAL-ADC-Curator@usda.gov.

(3) Metadata requirements for geospatial data

Geospatial data have specific metadata requirements covered by other USDA policies ([Enterprise Geospatial Data Management | USDA](#)). If they are digital scientific research data assets, they will also need a metadata entry in Ag Data Commons. While [policy](#) established that all USDA geospatial metadata must follow the ISO 19115 metadata standard, details about additional fields or information required as well as options for creating and validating these records are still in progress.

5. Digital Persistent Identifiers for Researchers

a. What is an individual digital persistent identifier?

An individual digital persistent identifier, or DPI, is an identification number for a person that is unique, persists forever, and resolves (provides information) in a standard, publicly accessible way such as putting it into a web browser. For example, the Open Research and Contributor ID or ORCID iD is a type of digital persistent identifier for researchers, and the one preferred by most journals, universities, and agencies, including USDA.

b. Who needs an individual digital persistent identifier?

All authors of federally-funded research products subject to this Policy are required to use a DPI such as an ORCID iD when they publish.

c. How do researchers obtain a DPI?

A researcher keeps the same ORCID iD for the duration of an entire career, even if their name or institutional affiliation changes. The National Agricultural Library maintains the Departmental membership in ORCID. Individuals can obtain a free ORCID iD at <https://www.orcid.org>. Many research organizations provide a special link for their staff to obtain and/or connect their ORCID iD. USDA staff are encouraged to connect their ORCID iD to USDA by going to <https://digitop.nal.usda.gov/orcid>. If USDA staff do not connect their ORCID iD to USDA, compliance tracking may be difficult.

d. Not all journals or data repositories collect DPIs such as an ORCID iD when a manuscript or dataset is published. How can researchers comply with the Policy when a journal or repository does not accommodate DPIs?

When researchers submit their manuscripts to PubAg they will be asked to provide ORCID iDs if these are not already associated with the publication. Similarly, if data are published elsewhere without DPIs such as ORCID iDs, they can be provided to Ag Data Commons when the catalog record is submitted.

e. Do non-USDA funded authors of scholarly publications need to have and use an ORCID iD?

The Policy currently requires that all authors of USDA-funded publications must use an identifier such as an ORCID iD when publishing USDA-funded research. Agency policies do not yet indicate who is responsible for enforcing compliance by non-USDA-funded authors. It is generally a good idea

to use them, however, as they make it easier for all of a researcher's publications to be associated with them.