



CEPLAS

Cluster of Excellence on Plant Sciences

Appendix to the CEPLAS Research Data Policy v 2.0

—

Practical Guidance for Researchers

July 17th, 2026

Introduction

This document provides practical guidance for complying with the CEPLAS Research Data Policy version 2. While the policy defines the requirements and responsibilities for managing research data, this guide explains in simpler terms what researchers need to do, when they need to do it, and how to use DataPLANT tools and services to meet those requirements.

This document may be updated more frequently than the CEPLAS Research Data Policy to align with the ongoing developments of the employed research data management system.

Key Concepts

What is an ARC?

An ARC (Annotated Research Context) is the project container used to organize research data, metadata, protocols, workflows, and documentation throughout the lifecycle of a project. All CEPLAS projects have to be organized into ARCs and shared with the community. An ARC can be created at the beginning of a project and updated continuously as research progresses. At a minimum, an initial ARC should contain descriptive metadata about the project. As the project develops, it should also contain data, protocols, workflows, and other relevant research outputs.

What is the PLANTdataHUB?

The PLANTdataHUB (<https://git.nfdi4plants.org/>) is the platform provided by DataPLANT for managing and sharing ARCs. Built on GitLab, it combines version control, access management, and collaboration features to support reproducible research workflows.

- Store, version, and manage ARCs in a central location
- Share ARCs with individual collaborators or entire groups
- Organize research groups and members (e.g. CEPLAS, individual labs, or consortia)
- Manage permissions and access through group- and project-level roles
- Use project management features such as issues, discussions, and task tracking alongside your data

Namespaces

A namespace determines who owns and manages an ARC within the PLANTdataHUB. There are two common types of namespaces:

- Personal namespace – Every user has their own personal namespace where ARCs are created by default.

- Group namespace – Groups and subgroups also have namespaces that can be used to share ARCs with a consortium, department, or project team.

Type	URL	Display Name
Personal namespace	https://git.nfdi4plants.org/brilator/	Dominik Brilhaus
Group namespace	https://git.nfdi4plants.org/ceplas/	CEPLAS

Most researchers will initially create an ARC in their personal namespace and later share it with the consortium through the CEPLAS PLANTdataHUB group.

More details

<https://nfdi4plants.org/nfdi4plants.knowledgebase/datahub/working-together/>

Visibility

The PLANTdataHUB visibility options are:

- Private – Only explicitly invited users or groups can access the ARC.
- Internal – Any authenticated PLANTdataHUB user can access the ARC.
- Public – The ARC can be accessed by anyone.

By default, newly created ARCs are private. Researchers may therefore work privately during the early stages of a project before sharing data with collaborators or making it publicly available. Visibility settings can be configured both for ARCs and for groups. These settings are independent of each other. For example, a group such as CEPLAS may be publicly visible so that users can discover the organization and its members, while the ARCs contained within the group remain private and accessible only to authorized collaborators.

More details

<https://nfdi4plants.org/nfdi4plants.knowledgebase/datahub/working-together/>

Roles and Permissions

Permissions determine who can access an ARC and what actions they are allowed to perform on the data, metadata and settings. When creating an ARC, the creator is automatically assigned the role of “Owner” and has full control over the ARC. Additional users can be invited to collaborate on the ARC with different roles, which define their level of access and capabilities.

The PLANTdataHUB provides several predefined roles:

Role	Typical permissions
Reporter	View and download ARC contents, but not modify them.
Developer	View, upload, edit, and contribute changes to the ARC.
Maintainer	Manage the ARC, including settings, members, and collaboration workflows.
Owner	Full control over the ARC or group, including permissions and administrative settings.

A common setup in research projects is:

- ARC creator – Owner
- Principal Investigator (PI) – Owner
- Active collaborators and data contributors – Developer
- Colleagues who only need to inspect the data – Reporter

Forks

A fork is a copy of an ARC that remains linked to the original ARC. Forking allows researchers to share a specific version of an ARC with a group while continuing development in their own copy. The original ARC and the fork can then evolve independently.

For compliance with this policy, researchers should:

1. Maintain their working ARC in their personal or group namespace.
2. Create a fork into the CEPLAS group when sharing with the consortium is required.
3. Update the consortium fork according to the timelines described below.

More details

<https://nfdi4plants.org/nfdi4plants.knowledgebase/datahub/arc-features/datahub-arc-fork/>

CEPLAS Data Delivery

Several phases and events between CEPLAS project initiation, completion, and beyond require varying tasks for data delivery as summarized in Figure 1.

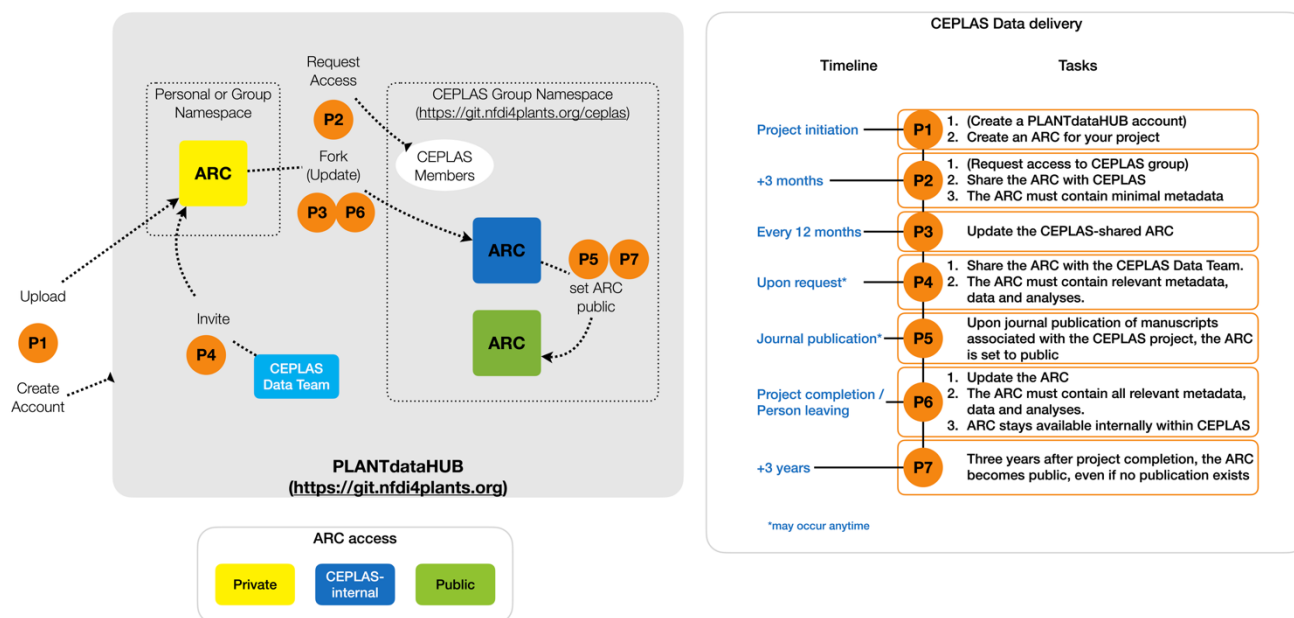


Figure 1: CEPLAS Data Delivery timeline and tasks

P1 – Project Initiation

At the start of every CEPLAS project:

1. Create a PLANTdataHUB account via <https://register.nfdi4plants.org> using your university email.
2. Create an ARC for the project.

At this stage

- the ARC may remain private
- metadata may be incomplete and no sharing of raw data or data analysis is required

The ARC should be maintained throughout the project lifetime.

P2 – Three Months After Project Initiation

Within three months of project start:

1. Request membership access in the CEPLAS PLANTdataHUB group (<https://git.nfdi4plants.org/ceplas/>), if you are not already a member. Send an informal email to the CEPLAS Data Team to notify about your request; this is simply a second security layer.
2. Once you are a member of the CEPLAS group, share the ARC with the other CEPLAS

members by creating a fork of your ARC into the CEPLAS PLANTdataHUB group.

3. The ARC must contain minimal metadata

Minimum requirement

The forked ARC must contain project metadata sufficient to identify and describe the project, including information such as:

- Project title
- Project description
- Responsible researchers
- Relevant project information

At this stage, uploading experimental data is not yet required.

P3 – Every twelve Months

The CEPLAS-shared ARC (i.e. fork) must be updated at least once every twelve months.

At this stage, researchers should ensure that:

- Metadata remains current.
- Project information reflects the current state of the work.
- Relevant documentation is maintained.

Regular updates ensure that other CEPLAS members have an accurate overview of ongoing research activities.

P4 – Upon Request

Researchers may be asked to provide the ARC, including data, before publication or project completion.

Examples include:

- TAC reviews
- Internal quality assurance
- Research data management audits

When requested, the ARC

1. must be shared with the CEPLAS Data Team by member invitation via the PLANTdataHUB.
2. must contain relevant metadata, data and analyses reflecting the current state of the CEPLAS project.

P5 – Upon Journal Publication

When results from the project are published,

1. The ARC must be made public.
2. Metadata and data should be openly available unless legal, ethical, contractual, or intellectual property restrictions apply.

The public ARC serves as the research record supporting the publication.

P6 – Project Completion or Researcher Departure

When a project is completed, or when the responsible researcher leaves CEPLAS:

1. Update the ARC one final time.
2. Include the project data in addition to metadata.
3. The ARC stays available internally within CEPLAS.

At this stage, ensure the ARC accurately reflects the completed state of the project. The ARC does not need to be publicly accessible.

Minimum requirement

The ARC should contain:

- Metadata
- Experimental data
- Documentation
- Protocols and workflows where applicable

P7 – Three Years After Project Completion

If a project has not resulted in a publication, the ARC must still be made public no later than three years after project completion.

The ARC may be published either by:

- The researchers (if still present), or
- The CEPLAS Data Team.

At this stage, both metadata and data become publicly accessible. An exception from making the data public are intellectual property constraints, or external ownership. The PI needs to apply to the steering committee to grant such an exception.

CEPLAS Data Publication

Before an ARC can be published, it goes through a publication workflow to ensure quality and reproducibility. This workflow is comprised of both automated and manual steps for validation and revision, as summarized in the figure below.

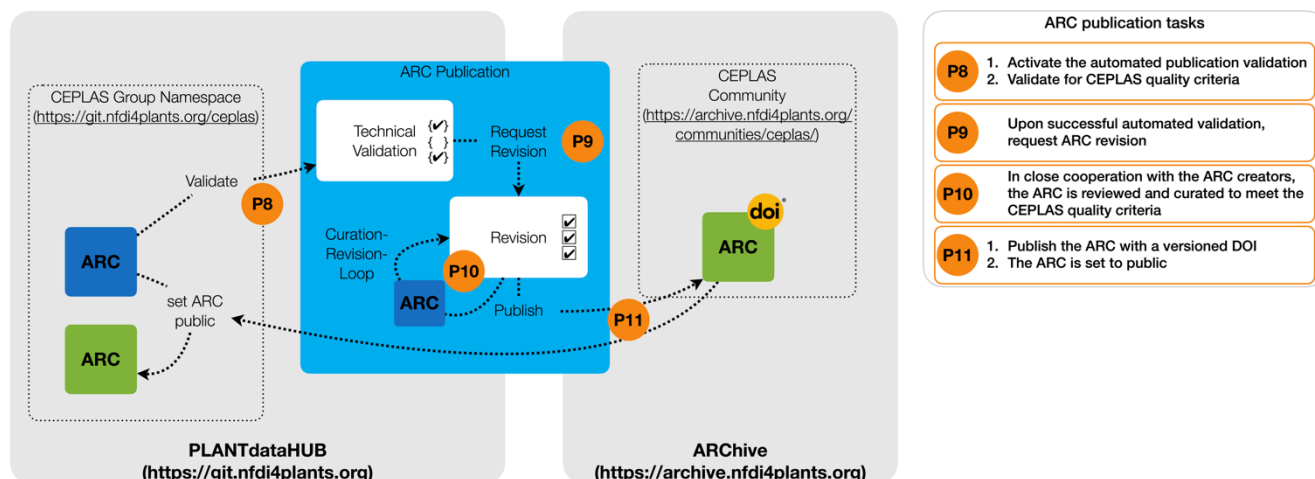


Figure 2: CEPLAS Data Publication tasks

P8 – Validation

This is an automated step that checks the ARC validity against both publication and CEPLAS quality criteria. The validation process is initiated by the ARC creator. The validation checks for existence and coherence of metadata and annotated datasets.

1. Activate the automated publication validation
2. Validate for CEPLAS quality criteria

More details

<https://nfdi4plants.org/nfdi4plants.knowledgebase/arc-publications>

P9 – Revision

Upon successful automated validation, request ARC revision. This is a manual step where the CEPLAS Data Team reviews the ARC content and provides feedback to the creators. This step requires disclosure of both the ARC content and the associated publication manuscript to ensure that all the information is transparent and accurate. The manual revision can take a variable amount of time depending on the complexity of the ARC and the volume of requests. It is recommended to initiate the revision process well in advance of the intended publication date to allow sufficient time for feedback and necessary adjustments.

P10 – Curation

In close cooperation with the ARC creators, the ARC is reviewed and curated to meet the CEPLAS quality criteria. This may involve updating metadata, adding missing documentation, or ensuring that data formats are standardized. The curation process ensures that the ARC is complete, accurate, and ready for publication.

P11 – Publication

Once the ARC has passed validation and revision, it is ready for publication. The final step involves making the ARC publicly accessible and assigning a versioned DOI (Digital Object Identifier) to ensure that it can be cited in future research.

Resources

CEPLAS Data Team

- Contacts: Vittorio Tracanna, Dominik Brillhaus
- Website: <https://www.ceplas.eu/en/research/ceplas-data>

PLANTdataHUB

- Main site: <https://git.nfdi4plants.org>
- Account registration: <https://register.nfdi4plants.org>
- CEPLAS group: <https://git.nfdi4plants.org/ceplas/>
- CEPLAS group members: https://git.nfdi4plants.org/groups/ceplas/-/group_members
- The following CEPLAS members are assigned the role “Owner”
 - Maria von Korff-Schmising
 - Björn Usadel
 - Céline Hönl
 - Vittorio Tracanna
 - Dominik Brillhaus

ARChive

- Main site: <https://archive.nfdi4plants.org>
- CEPLAS community: <https://archive.nfdi4plants.org/communities/ceplas>

Additional DataPLANT Resources

- Main site: <https://nfdi4plants.org>

- Knowledge Base: <https://nfdi4plants.org/nfdi4plants.knowledgebase/>
 - includes **CEPLAS-tailored** documentation
- Events page: <https://nfdi4plants.org/events/>
 - includes **CEPLAS-hosted** training events
- Support: <https://nfdi4plants.org/support/>

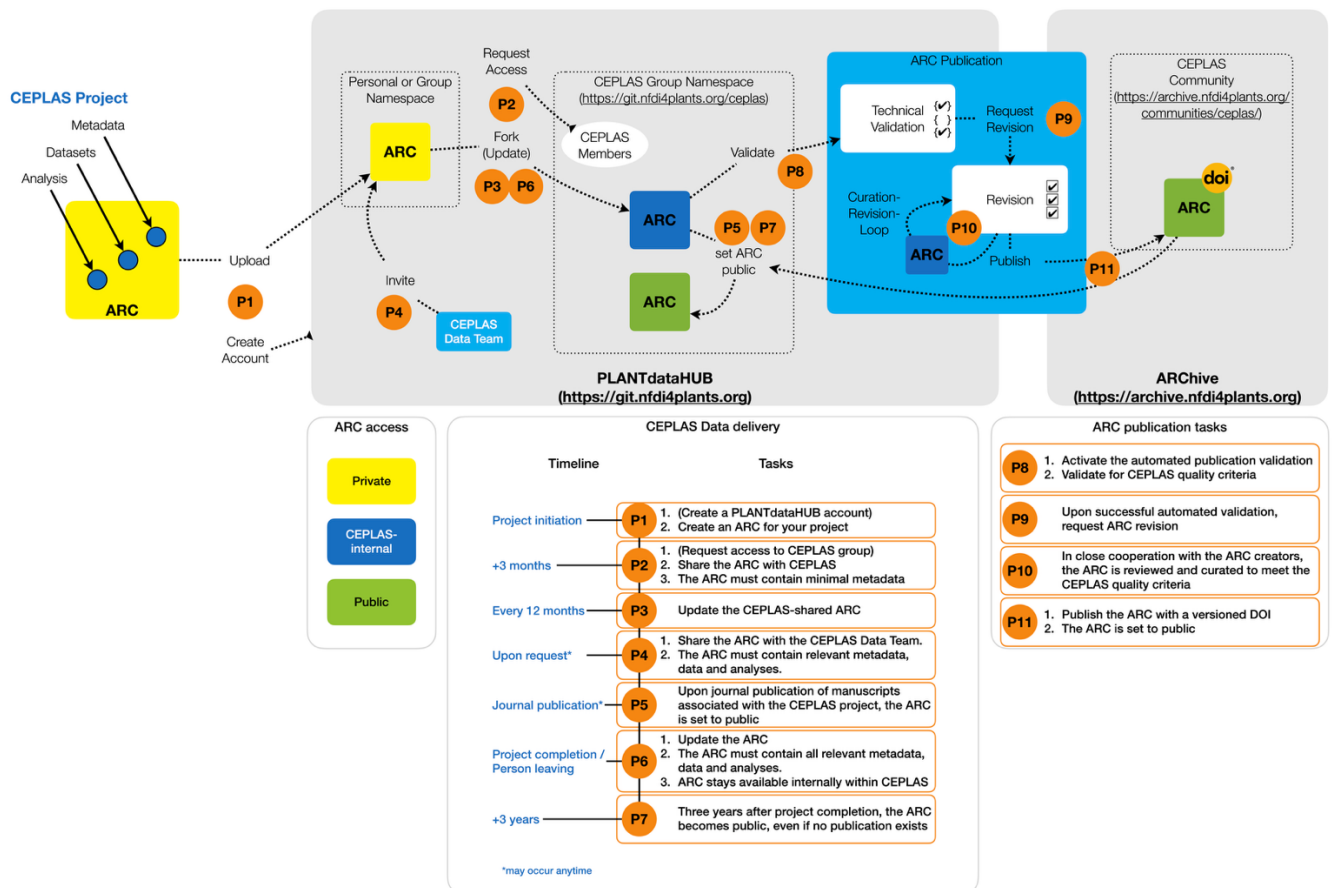


Figure 3: Complete overview of CEPLAS data delivery and data publication