

April 2011 Curation Manual

Notifications of new submissions arrive via email to all Dryad users with curator privileges.

Claiming submissions

1. **Log in** to Dryad and go to **My Submissions** page.
2. New submissions should be in the **Tasks in the pool** list.
3. Click on **Accept/Reject task** to view a package summary, where you may click the **Take task** button. This package will now appear in the Tasks you own list on the My Submissions page.

Evaluating submissions

1. Is this submission from an **integrated journal**?
(April 2011 fully integrated journals: The American Naturalist, Biological Journal of the Linnean Society, Evolution, Evolutionary Applications, Journal of Evolutionary Biology, Journal of Heredity, Molecular Ecology, Molecular Ecology Resources, Systematic Biology)
 - a. If **yes**, go to the dryad.journal.submit@gmail.com account and **look up the corresponding metadata email**.
 - i. **If you find a match**, proceed to **Editing submissions before approval**. If the matching email was sent more than a couple weeks ago, do a google search on the article title or search the journal website to check that it hasn't been published in the meantime. If the article has been published, you will add citation information as described below.
 - ii. **If you do not find a match**, first google the article title to see if it has been published some time in the past. If the **submission corresponds to a published article, proceed as in 2.b**. If there is **no matching email and no matching published article**, the manuscript has probably not been accepted by the journal yet (or, for certain journals, the author neglected to give permission that the metadata email be sent to Dryad). **Reject the submission** and use the boilerplate explanation given at the end of this document and on the Templates for Correspondence wiki page.
 - b. If **no**, proceed to **step 2**.
2. Has the **article associated with this submission already been published**?
View the full metadata of the package and look for an article DOI in dc:relation.isreferencedby or volume/issue/year in dc:relation.ispartofseries. If you don't have those depositor-supplied article metadata, google the article title or search the journal website.
 - a. If **no**, you are going to take the depositor's word that the manuscript has been accepted (this is for non-integrated journals; integrated journals were covered in step 1), but you are going to **make sure that the files are**

- embargoed at least until article publication** (though files associated with manuscripts that you know are accepted because you have a metadata email can be open if that was the depositor's embargo choice).
- b. If **yes**, you will **edit the package metadata** to correctly record article citation information (described below), since the submission system doesn't accommodate this yet.
3. Is the **content appropriate for the repository?** (and other data file issues)
Download the submitted data files and readme files, and look for the following problems.
- a. The **data file is the article manuscript itself** or a portion thereof: **reject the submission** and use the boilerplate explanation.
 - b. You suspect that one or all of the **data files are the figures from the manuscript** itself (not supplementary material or appendices): **email the depositor** using the boilerplate. Leave the submission as a "task you own" until you hear from the depositor. If the depositor hasn't responded after 48 hours, proceed with editing and approval.
 - c. The **data files do not contain data ready for analysis** (especially when the submitted files are clearly based upon files that are more appropriate for the repository, as in a submitted PDF summarizing results in a table): **email the depositor** using the boilerplate. Leave the submission as a "task you own" until you hear from the depositor. If the depositor hasn't responded after 48 hours, proceed with editing and approval.
 - d. The **files submitted as readme files contain data or files submitted as data are readme files**: **proceed with editing and approval**, creating new file records in the package, deleting file records and adding readmes to other files, etc., as needed. If you can't figure out what is supposed to be going on, you may need to email the depositor.
 - e. **Other issues with data files**: **check with Ryan** if it seems like the issue might be on our side, then/or **email the depositor**, and finally **reject the submission** (with explanation) if it is clearly inappropriate for the repository.

Editing submissions before approval

If you have a corresponding integrated journal metadata email, keep it handy while editing.

While you can edit items before approval from within the submission system, it is best to navigate out of the submission system to do so. Visit the item page while logged in as a curator, then click "Edit this item" on the left sidebar. This way, you can access all metadata fields and view the actual metadata being stored (rather than the depositor-friendly interface, which obscures some metadata) and also preview how the item will look to end users.

1. In all metadata fields, look for spelling and other errors. Generally **check that the metadata looks reasonable**. For example, look for duplicated metadata (like two complete matching sets of dc:subject terms), make sure author names are

formatted as Lastname, Firstname M, article title is in sentence case, etc. If something seems like a system error, look for a related bug ticket in Fogbugz, and file one if needed.

2. In all metadata fields, look for **special characters, punctuation marks, or formatting** that isn't displaying correctly. Formatting tags need to be removed and special characters may need to be copied and pasted from the notification email, unicode.org, or elsewhere. (This is rarely a problem any more.)
3. If needed, add **taxon names** that are prominent in the title or abstract in dwc:ScientificName. Likewise, possibly add **geographic names** or coordinates as dc:coverage.spatial and **geologic timespans** as dc:coverage.temporal. Do this at the package level.
4. **Look for any errors the depositor or submission system may have introduced**, such as metadata going into the wrong fields.
5. If the **submission is associated with a published article, edit the metadata** accordingly.
 - a. On the package record, article DOI is stored in dc:relation.isreferencedby with value formatted as "doi:####"
 - b. On the package record, article citation is stored in dc:identifier.citation with value formatted as "LastName FM, LastName FM (Year) Article title. Journal Title Volume(Issue): FirstPage–LastPage." List out all authors and do not abbreviate the journal title. If the article is published online and will later be published in print, include "Journal Title, online in advance of print." in place of the volume, issue, and pages.
 - c. On the file record, if the depositor has chosen to embargo until article publication, change the dc:type.embargo value to "none". If the depositor has chosen to embargo until one year after article publication, and that date is still in the future, you will change the dc:date:embargoedUntil value to match one year after the article publication date, but this can't be edited until after approval.
 - d. You will also change dc:date.issued on the package record to match the article publication date, but this can't be edited until after approval.
6. If the submission is from an integrated journal, **make sure that the appropriate managing editor and/or publisher email addresses appear** in the package metadata in repeats of workflow:archive.mailUsers. If they are missing, and the article has not already been published some time in the past, add these email addresses.
7. If the submission was not completed via the integrated process, **check the journal name**. Search to see if this journal is already represented in Dryad (Browse by Journal Title is useful, but watch out for initial adjectives and also do a search), and, if so, make sure there is an exact match in the form of title used in prism:publicationName on the existing records and new submission. If this journal title is new to Dryad, choose the form of title using the TRLN catalog and/or Worldcat as a guide.
8. If time permits, search for existing content in Dryad from authors of the new

submission, and compare forms of name. **Edit names for consistency** across the repository, if needed, preferring the fullest form of the name. [Note: name authority control is a much larger project than can be accommodated in our current practices. Even at this level, research is often required to determine author identity and disambiguate names.]

Approval and DOI registration

1. Navigate back to the summary page for the item you have edited. You can do this by going to My Submissions page and clicking Awaiting curator review next to an item in the Tasks you own list, or, if you had been on the Submission overview page for the package, by clicking Return. Click the **Approve** button on the summary page.
2. Make sure that all the **DOIs were registered properly**, keeping in mind that there can be a delay before they resolve. In general, after clicking Approve you will be returned to My Submissions page if all DOIs were successfully registered and will resolve to a 502 proxy error page if there was a problem with DOI registration. When in doubt, check DOIs using the EZID interface. If needed, manually create any missing DOIs using EZID.

Editing submissions after approval

1. **Update any date metadata** that you couldn't edit before approval, usually only necessary if the associated article has already been published.
 - a. The **date that the article was first published** (take the online publication date if that is earlier than the print publication date) is stored in the **package record dc:date.issued** with value formatted as YYYY-MM-DD.
 - b. Edit **embargo date for one year embargoes** to match one year after article was first published in **file record dc:date.embargoedUntil** with value formatted as YYYY-MM-DD. For all embargoed items before article publication, the embargoed until date should be 9999-01-01. For one year embargoes, the embargoed until date will be updated when the article is published.

Notifications and documentation of approved submissions

1. For integrated journals, **check that the notification email was sent**.
2. Enter information about the data package and any submission errors or other notes in the **JournalSubmissionTracking googledocs spreadsheet**.
3. Significant submission errors should also be recorded on the wiki:
https://www.nescent.org/wg_dryad/Submission_Integration_Issues. And deposits with unique characteristics may be recorded on the wiki:
https://www.nescent.org/wg_dryad/Sample_Dryad_Content.

Weekly Reporting

These processes are completed every Friday.

1. **Collect weekly summary data:** the number of article notifications can be determined by counting the metadata emails received by dryad.journal.submit@gmail.com (filters are applied to sort by journal), and the new submissions for each journal can be found in the JournalSubmissionTracking googledocs spreadsheet.
2. **Update the Integration Details pages on the wiki:**
https://www.nescent.org/wg_dryad/American_Naturalist_Integration_Details
https://www.nescent.org/wg_dryad/BJLS_Integration_Details
https://www.nescent.org/wg_dryad/Evolution_Integration_Details
https://www.nescent.org/wg_dryad/Evolutionary_Applications_Integration_Details
https://www.nescent.org/wg_dryad/JEB_Integration_Details
https://www.nescent.org/wg_dryad/Journal_of_Heredity_Integration_Details
https://www.nescent.org/wg_dryad/Molecular_Ecology_Integration_Details
https://www.nescent.org/wg_dryad/Molecular_Ecology_Resources_Integration_Details
https://www.nescent.org/wg_dryad/Systematic_Biology_Integration_Details
3. **Generate an integration report email for each integrated journal** using the template. The recipient list for each journal is stored in the Contacts of the dryad.journal.submit@gmail.com account. Set the sender address to curator@datadryad.org.
4. **Generate a curation summary email** using the template and send to the public [dryad-dev](#) Google group.

Updating article citation information

The curator uses **RSS feeds and/or eTOCs emails to monitor article publication for all integrated journals**. As these notices arrive, they are searched manually for matches in Dryad. If a newly published article is found to have data in Dryad, the Dryad records are updated as follows.

1. Package records
 - a. The **article DOI is stored in dc.relation.isreferencedby** formatted as “doi:####”. For articles without DOIs, look for another identifier to store in the same field, such as a PubMed ID, formatted as “pmid:####”.
 - b. The **article citation is stored in dc:identifier.citation** with value formatted as “LastName FM, LastName FM (Year) Article title. Journal Title Volume(Issue): FirstPage–LastPage.” List all authors (no “et al.”) and do not abbreviate the journal title. If the article is published online and will later be published in print, include “Journal Title, online in advance of print” in place of the volume, issue, and pages.
 - c. Update **dc:date.issued to match the earliest article publication date**. If the article is online in advance of print, use the electronic publication date.
 - d. **Compare Dryad metadata to what appears with the published article**, such as minor changes to the title and abstract that may have occurred between acceptance and publication, and especially author names (the addition of middle initials, accents and diacritics, etc.), and update as

needed.

2. File records
 - a. Update **embargoes**, if needed. For items embargoed until article publication, lift the embargo when you are on the Edit Item page by clicking on Item Embargo and selecting the radio button labeled Disabled. For items with one year embargoes, set the embargoed until date to one year after the earliest article publication date (Edit Item, then Edit Metadata for each file), formatted as “YYYY-MM-DD”.
 - b. Double check that **rights information is present in the metadata** and that the Open Data and CC0 logos display once any embargoes are lifted.
 - c. If you edited author names or other metadata at the package level that is reproduced on file records, **edit each file record to match the package**.
3. **Update the google spreadsheet** to indicate that an article citation has been added. Values for this column: N (no citation), online (for online in advance of print), Y (for final version of the citation). If the final citation is not yet added, this cell should be highlighted in light yellow.
4. View the **full text of the article and search the Dryad DOI** (searching / finding in page with the term “dryad” will help find the data reference even if it is incorrectly presented). If the DOI is incorrect or presented in a way that will not allow readers to find the data, you may need to email the corresponding author (and managing editor for integrated journals) to let them know a correction is needed. If there is time, and particularly if this is a noteworthy example, add an entry to the DOI Examples Google spreadsheet.

Templates for Correspondence

Rejecting submission because manuscript has not yet been accepted

Dear [],

It appears that you may have submitted data files associated with your [Journal Title] manuscript ahead of the acceptance of that manuscript. At this time, Dryad archives data once the article has been accepted for publication by [Journal Title]. For this journal (and most others) that still leaves adequate time to assign the data DOI and include it in the published article. If I am mistaken about the status of your manuscript, please let me know. We look forward seeing your submission again once the article has been accepted.

Thanks,

[]

Rejecting submission because files are inappropriate for the repository

Dear [],

This submission is being sent back to you because [describe the problem and name the problem file if there is more than one in the package]. We look forward seeing your submission again with an alternative data file. Please see the following page for information about what to deposit in Dryad:
<http://www.datadryad.org/depositing>

Let us know if you have any questions.

Thanks,
[]

Asking depositor about potentially inappropriate files: figures or other content from the article

Dear [],

I'm ready to approve your Dryad submission and assign a DOI, but wanted to ask about the following file that you uploaded:
[file name]

It appears that this file may contain content from the manuscript itself [or more specific phrasing], which would be inappropriate to release under Creative Commons Zero.

Please see the following page for information about what to deposit in Dryad:
<http://www.datadryad.org/depositing>

and the following page for information about the Creative Commons Zero terms:
<http://creativecommons.org/about/cc0>

Please let me know if you would like to provide alternative data files, or if you have any questions about the type of files to submit, why Dryad asks you to release your files from copyright restrictions, and any other concerns you may have.

Thanks,
[]

Asking depositor about potentially inappropriate files: data not ready for analysis

Dear [],

I'm ready to approve your Dryad submission and assign a DOI, but wanted to ask if you would be able to provide your data files in an alternate format, such as [give appropriate examples based on their file, like CSV, tab delimited, plain text, even Excel]. Data in Dryad should be submitted in a state that is ready for analysis and optimal for reuse, so

[the file format they submitted] is not ideal. Thanks for considering this request.

[]

Notification of submission approval and DOI (automated)

Subject: Dryad submission approved and assigned a DOI doi:10.5061/dryad.#####

Body:

Thank you for your recent submission to the Dryad repository titled "[package title]". Your data package has been archived in Dryad and assigned a unique identifier, called a DOI. Including the data DOI in the published article will make readers aware that the data files are available, and enable their access and citation.

The DOI may be presented as follows:

Data deposited at Dryad: doi:10.5061/dryad.#####

Dryad links your data to your published article; it is up to authors and journals to link the published article to its supporting data. Please work with your journals editorial and production staff to ensure that the data DOI is included.

Many journals specify a particular location for such links to data in repositories, or have a section on Data Accessibility. You can also provide your data DOI in the text, just before the References. If your article has already been published, use this DOI to refer others to your data.

Please let us know if you have any questions or concerns.

The Dryad Team

Weekly integration report

In the week since the last report (YYYY-MM-DD), Dryad has received the following.

New notifications of accepted publications in this journal: #

New deposits associated with publications in this journal: #

Please find the data DOIs for new deposits below, for inclusion in the published articles.

CorrespondingAuthorLastName doi:10.5061/dryad.#####

CorrespondingAuthorLastName doi:10.5061/dryad.#####

The DOI may be presented in either of these 2 forms:

Data deposited at Dryad: <http://dx.doi.org/10.5061/dryad.#####>

Data deposited at Dryad: doi:10.5061/dryad.#####

[Omit if no new deposits]

Dryad links the data files to the published articles; it is up to journals to link published articles to their supporting data. Journal staff and authors collaborate to ensure that data DOIs are included in the appropriate locations for the print and online versions of published articles.

Summary statistics may be found here:

[URL of this journal's integration summary page on wiki]

Thank you!

- []

Weekly curation summary

Subject: weekly dryad curation summary

Body:

In the past week (since YYYY-MM-DD):

New deposits from integrated journals: #

New deposits from non-integrated journals: #

Of the [] deposits from non-integrated journals [] were submitted ahead of article publication. They are associated with articles to appear in the following journals: [].

Dryad accepted data associated with its [] journal: [].

[Other notes]

- []