

Usability Study: Research Data Repository

Recommendations

Deposit Data Interface

1. Provide a roadmap/decision tree at the start of the submission portal, so users will be aware of what information they will need to have prepared for a successful submission. A lot of publication submission portals will provide a checklist so that submitters are aware of all the content they will need to have ready throughout the submission process.
2. Provide more clarity on how aspects of the dataset submission timeline can be adjusted depending on the publication release date.

Browse Data Interface

1. Provide an option to download a single folder of files within the dataset.
2. Rethink the fileset page by removing confusing information like the fixity check and ensuring useful content like the file preview is available.
3. Redo thumbnails to better match file type, especially for directories.
4. Consider displaying README content on the dataset homepage and de-emphasizing more traditional, publication-oriented metadata like long author lists and article abstracts.

Background

This study investigated user experience of the Research Data Repository (RDR), an online digital repository intended for researchers to both download published datasets and upload datasets from their own publications. Therefore, the study was split into two sub-studies, one for the upload interface and one for the download interface. The upload interface study investigated how well the submission interface supports a participant's ability to successfully fill out metadata related to a publication dataset. The download interface study investigated how well the search interface supports a participant's ability to search for datasets and comprehend specific content within a dataset's homepage.

Participants were recruited into two participant cohorts. One cohort was for "expert" users who had used the RDR in the past to deposit data from one or more publications. Another cohort was for "novice" users who have academic backgrounds but were completely unfamiliar with the RDR. The expert participants were recruited by a member of the RDR Product Council, and the novice participants were recruited from a list of volunteer graduate students available to be contacted about library user

studies. Novices completed a screener survey to establish that they had no prior experience with the RDR.

All study sessions were conducted and recorded remotely via Zoom. Yasha Saxena moderated the sessions, while Shelley Dong took notes.

Questions we hoped to answer through this study:

Deposit Data Interface

1. [Are the instructions for filling out dataset metadata clear and unambiguous?](#)
2. [Do users have a good understanding of how to respond to the restricted information and Creative Commons license sections?](#)
3. [What are potential areas where information may be entered incorrectly or insufficiently?](#)
4. [What are users' interpretations of a README requirement?](#)

Browse Data Interface

1. [Does the dataset homepage provide users with a reasonable outline of what the dataset consists of?](#)
2. [Do people understand how to navigate through the dataset?](#)
3. [Is the usage stats page useful for dataset owners?](#)
4. [Does the fileset page that lists details about individual files provide useful information?](#)
5. [How do users respond to datasets that must be accessed through Globus?](#)

Participants

“Novice” Participants

Affiliation

- 3 PhD students (2 Biomedical Engineering, 1 Toxicology)
- 1 Masters student (Masters in Engineering Management)

Prior Experience with RDR Webpages

None had prior experience with RDR webpages, all had some experience with publishing research or conducting research

“Expert” Participants

Affiliation

- 2 Postdoctoral Fellows (1 Biomedical Engineering, 1 Physics)
- 1 Senior Policy Associate (Nicholas School of the Environment)
- 1 Faculty member (Professor of Pathology)

Prior Experience with RDR Webpages

All have had experience using the RDR to upload dataset from their publications. None had used the “Browse datasets” section of the website.

Key findings

Deposit Data Interface

1. Are the instructions for filling out dataset metadata clear and unambiguous?

Yes. All participants were able to correctly interpret all the fields listed out in the metadata form. There were only a few potential ambiguities identified:

- Information/DOI for a publication that is still in review: A couple of the expert participants pointed out the back and forth discussion they had with the RDR staff related to data being under embargo until the affiliated publication was officially published. In some cases, data should be published only after the publication review process, but the publication review process may require datasets to already have a DOI. For both participants, the process was not made clear by the submission portal itself and the situation was addressed through asking a lot of questions. Some clarity and dialogue on the sequence of events and expectations for this specific issue would have been nice to have here.
- Value of filling out unrequired fields: Because these fields were not required, there was a tendency to not fill them unless the information was readily available. During the submission process, those fields do not mean much to the uploader unless their relevance in the “Browse Datasets” interface is made more clear. The “Funding” and “Grant Number” categories would be important to include based on the affiliated publication. If some fields could improve the usability or clarity in the download interface, it might be good to encourage uploaders to enter those fields. Including a mock up of what the metadata would look like in the “Browse Datasets” interface might be a good way to encourage the uploader to fill out the metadata the way they would want it viewed by a downloader.
- Sensitivity of author lists, contributor lists: This was mainly brought up by one novice user. The novice user felt that the gravity of the author list order and importance of including contributors was not explicitly stated in the field descriptions. In general, author order and contributor inclusion is a sensitive consideration for publications, so it may be advisable to encourage the same practice in the data repository. Another aspect here is that data downloaders may want to contact authors based on what they contributed to the dataset. If these specifics are included like they are in a publication, the right people would be contacted from the start. All expert users did adhere to the way author and contributor lists were handled in their respective publications.

- Description section as README or abstract: Most expert users defaulted to pasting in the publication abstract into the description section of the form. One expert used it to actually describe her dataset and dataset hierarchy. In general, the inclusion of a README file provides a lot of clarity on the dataset, and all experts agreed it is a good thing to have, but this is not a required element in the upload interface. There are also no guidelines on how to set up the README, so all experts came up with a template on their own. Including README requirements up front may provide useful information for the dataset and guide the uploader on what to include. A quality README may provide better content for the dataset homepage than the abstract of a related publication.

2. Do users have a good understanding of how to respond to the restricted information and Creative Commons license sections?

- Restricted Information: There was no confusion, and the default assumption is data without personal identifiers or that does not involve humans is fine. If there are ambiguous scenarios, it might be good to list them because all uploaders assume their data is not restricted.
- Creative Commons: In general, participants mentioned that some definitions for each of the options on the licensing page would be helpful for uploaders. The faculty member was familiar with and preferred CC0 licenses, but the other expert participants needed to clarify this part with their advisors or through further research. This section in general creates the most pause for uploaders because of the lack of awareness of licensing implications.

3. What are potential areas where information may be entered incorrectly or insufficiently?

- Every participant felt that the information entry process was easy if they had knowledge of what to enter. There are a few fields like the ones mentioned above where users may need to check with supervisors or think about what to enter. A few participants mentioned that having knowledge of what they would need to enter in advance would be helpful or having the option to save the form might be useful to be able to prepare information in advance of submission.

4. What are users' interpretations of a README requirement?

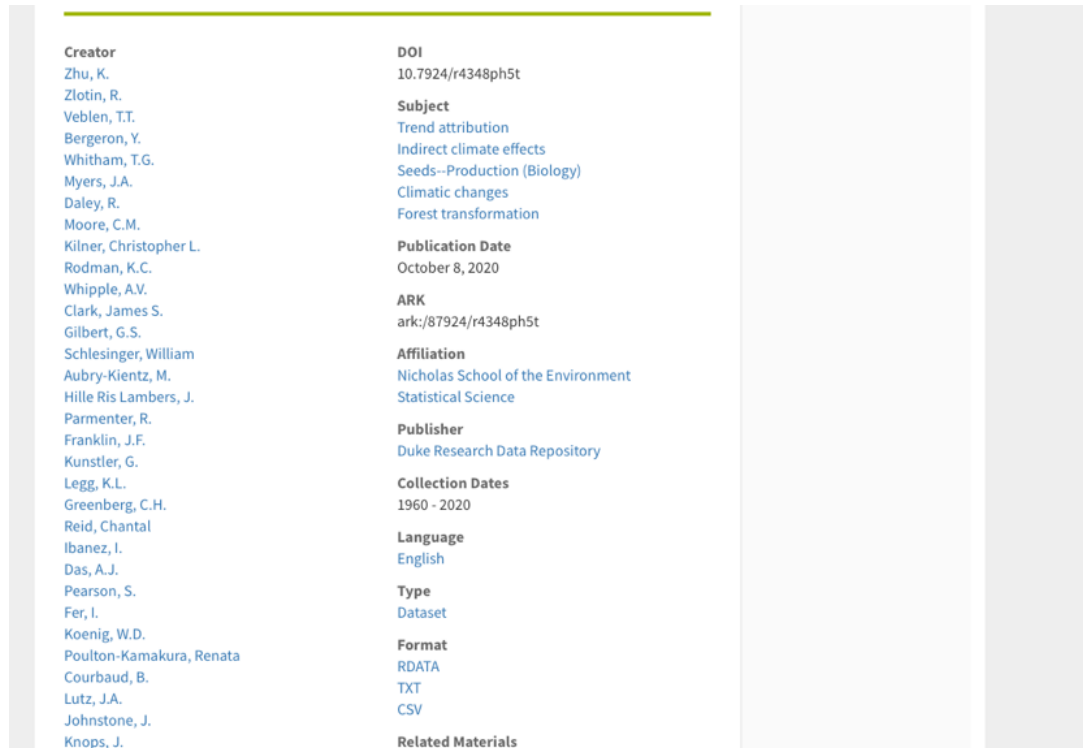
- The expert participants included README files, but they were not aware of requirements and set them up as they best saw fit or from guidelines from other sources.

Browse Data Interface

1. Does the dataset homepage provide users with a reasonable outline of what the dataset consists of?

Users were generally unable to quickly get a good grasp of what the dataset contained based on some usability hindrances:

- Long author lists: The long author lists required users to scroll quite a bit, and in some cases the moderator needed to ask the user to scroll further to be able to view the list of downloadable files at the bottom of the page.



The screenshot shows a dataset homepage with a long list of authors on the left and metadata on the right. The authors list is scrollable, and the metadata is organized into sections.

Creator	DOI
Zhu, K.	10.7924/r4348ph5t
Zlotin, R.	
Veblen, T.T.	Subject
Bergeron, Y.	Trend attribution
Whitham, T.G.	Indirect climate effects
Myers, J.A.	Seeds--Production (Biology)
Daley, R.	Climatic changes
Moore, C.M.	Forest transformation
Kilner, Christopher L.	
Rodman, K.C.	Publication Date
Whipple, A.V.	October 8, 2020
Clark, James S.	
Gilbert, G.S.	ARK
Schlesinger, William	ark:/87924/r4348ph5t
Aubry-Kientz, M.	
Hille Ris Lambers, J.	Affiliation
Parmenter, R.	Nicholas School of the Environment
Franklin, J.F.	Statistical Science
Kunstler, G.	
Legg, K.L.	Publisher
Greenberg, C.H.	Duke Research Data Repository
Reid, Chantal	
Ibanez, I.	Collection Dates
Das, A.J.	1960 - 2020
Pearson, S.	
Fer, I.	Language
Koenig, W.D.	English
Poulton-Kamakura, Renata	
Courbaud, B.	Type
Lutz, J.A.	Dataset
Johnstone, J.	
Knops, J.	Format
	RDATA
	TXT
	CSV
	Related Materials

- Lack of README type content displayed on the home page: Participants mentioned that to actually understand the dataset they need to download the README. Having it included somewhere on the homepage would allow them to bypass a potentially unnecessary download.
- Requests for visual content: A couple participants mentioned that inclusion of a graphical abstract or figures from the publication might provide faster guidance on what the dataset contains.

2. Do people understand how to navigate through the dataset?

- Search page: In general, participants were able to search for the example datasets with the search bar. Participants that tried to use the filters did not have much success with the task to find a dataset related to “climate change on trees.”

The screenshot shows the search results for 'climate change trees' on the Duke University Digital Repository Research Data platform. The search bar at the top contains the text 'climate change trees'. The results are filtered by this search term. On the left, there are three filter menus: 'Subject', 'Creator', and 'Format'. The 'Subject' menu shows 'Blue carbon' (4), 'Salt marshes' (3), 'Sea level' (3), 'Seagrasses' (3), and 'Climatic changes' (2). The 'Creator' menu shows 'Olander, Lydia' (4), 'Warnell, Katie' (4), 'Currin, Carolyn' (3), 'Clark, James S.' (2), and 'Swenson, Jennifer J.' (2). The 'Format' menu shows 'CSV' (8), 'TXT' (4), and 'PDF' (3). The main results area shows three items, each with a document icon and a 'Data Citation' link. The first item is 'Data from: Continent-wide tree fecundity driven by indirect climate effects' by Clark, J.S., et al. (2020). The second is 'Code from: Analysis of observed warming trends at U.S. basic combat training installations and consideration of implications for future recruit training' by Patton, E., & Doyle, M. (2023). The third is 'Data from: Mechanistic modeling of climate effects on redistribution and population growth in a community of fish species' by Tang, B., et al. (2023). The page also includes navigation links for 'Previous', 'Next', 'Sort by relevance', and '10 per page'.

DUKE UNIVERSITY LIBRARIES DIGITAL REPOSITORY RESEARCH DATA

climate change trees

Yasha Saxena

Limit your search

Filtering by: climate change trees

Start Over

« Previous | 1 - 10 of 14 | Next »

Sort by relevance

10 per page

Subject

- Blue carbon 4
- Salt marshes 3
- Sea level 3
- Seagrasses 3
- Climatic changes 2
- more »

Creator

- Olander, Lydia 4
- Warnell, Katie 4
- Currin, Carolyn 3
- Clark, James S. 2
- Swenson, Jennifer J. 2
- more »

Format

- CSV 8
- TXT 4
- PDF 3

Data from: Continent-wide tree fecundity driven by indirect climate effects

Data Citation: Clark, J.S., Andrus, R., Aubry-Kientz, M., Bergeron, Y., Bogdziewicz, M., Bragg, D.C., Brockway, D., Cleavitt, N.L., Cohen, S., Courbaud, B., Daley, R., Das, A.J., Dietze, M., Fahey, T.J., Fer, I., Franklin, J.F., Gehring, C.A., Gilbert, G.S., Greenberg, C.H., ... Zlotin, R. (2020). Data from: Continent-wide tree fecundity driven by indirect climate effects. Duke Research Data Repository. <https://doi.org/10.7924/r4348ph5t>

Code from: Analysis of observed warming trends at U.S. basic combat training installations and consideration of implications for future recruit training

Data Citation: Patton, E., & Doyle, M. (2023). Code from: Analysis of observed warming trends at U.S. basic combat training installations and consideration of implications for future recruit training. Duke Research Data Repository. <https://doi.org/10.7924/r4057r100>

Data from: Mechanistic modeling of climate effects on redistribution and population growth in a community of fish species

Data Citation: Tang, B., Roberts, S. M., Clark, J. S. & Gelfand, A. E. (2023). Data from: Mechanistic modeling of climate effects on redistribution and population growth in a community of fish species. Duke Research Data Repository. <https://doi.org/10.7924/r4tq63p9k>

- Unclear visuals for directories versus files: Every participant was thrown off by the “files” that did not have download buttons. The interface requires clicking into directories to enter another level. All files need to be downloaded either as part of the complete dataset or at the individual file level. Every participant at first assumed there was something about the directory-level data that they could not download. The moderator had to ask the users to click into a list item with no download button to figure out that it was actually a folder of files. Generally, the visuals and widgets in this section of the interface do not make it clear how the data is organized and how a user can navigate to download a particular file or subset of files. Furthermore, the long author lists are still included when you click into a directory of the data, so users have to scroll all the way down before they see the individual files in that directory.

Thumbnail	Title	Date Uploaded	Actions
	readme.txt	2020-10-08	Download
	mastif-BONA	2020-10-08	
	mastif-DukeSeeds	2020-10-08	
	mastif-DukeTrees	2020-10-08	
	mastif-EPENN	2020-10-08	
	mastif-MORA	2020-10-08	
	mastif-NEONSeeds	2020-10-08	
	mastif-NEONtrees	2020-10-08	
	mastif-NIWO-Ve	2020-10-08	
	mastif-SEQU_YOSE	2020-10-08	

mastif-BONA

Creator
 Redmond, M.D.
 Keenig, W.D.
 Zlotin, R.
 Wion, A.P.
 Stephenson, N.L.
 Johnstone, J.
 Whipple, A.V.
 Cleavitt, N.L.
 Brockway, D.
 Fer, I.
 Moran, E.
 Messaoud, Y.
 Daley, R.
 Schlesinger, William
 Witham, T.G.
 Kline, Christopher L.
 Swift, Maggie
 Franklin, J.F.
 Sutton, Samantha
 Cohen, S.
 Schwantes, Amanada
 Dietze, M.
 Guo, Q.
 Shanahan, E.
 Bergeron, Y.
 LaMontagne, J.M.
 Parmenter, R.
 Reid, Chantal
 Steele, M.

Publication Date
October 7, 2020

ARK
ark:/87924/r4x63j65c

Publisher
Duke Research Data Repository

Format
CSV
TXT

License
Creative Commons CC0 1.0 Universal

Home
↓
Data from: Continent-wide tree
fecundity driven by indirect...
↓
This Dataset

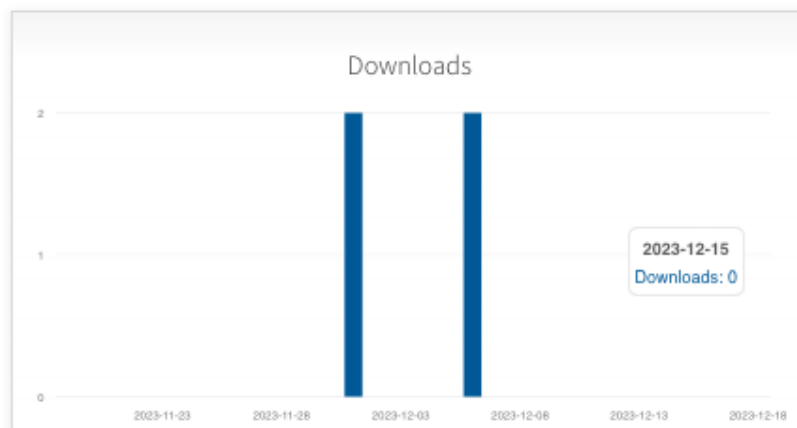
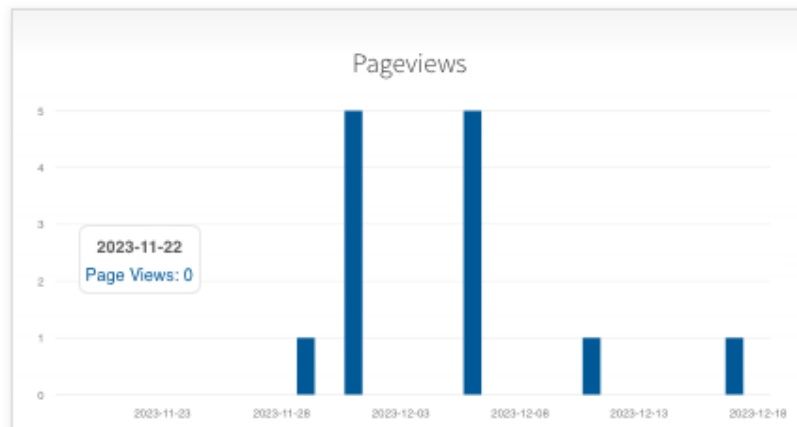
[USAGE STATS](#)

3. Is the usage stats page useful for dataset owners?

- Most dataset owners (i.e., the experts) did not find this page useful and were not aware that it existed. A couple of them mentioned having an interest in being able to see if there were views on their dataset after attending a conference or giving a talk, but the current level of information would not be able to provide that level of detail. In general, this page did not seem important to dataset owners.

Work Analytics

This work has 280 total views and 83 downloads



4. Does the fileset page that lists details about individual files provide useful information?

- This page was universally found to not provide any useful information. Some users mentioned that if the page had a file preview, that would be useful, but in the files that we explored through the study this was not the case. The fixity check was also extremely confusing to all participants. In the case of a fixity check failure, a lot of users mentioned that if they had seen this they would generally be very confused about the quality of the file and may not download it.

No preview available

BNZ_SeedCounts.csv Public

File Details

Depositor: James Clark, Ph.D.	UTC; expected_result: urn:sha1:6092a02d4e13ace46519759fd22662089c305abf
Date Uploaded: 2020-10-08	
Date Modified: 2020-10-08	
Fixity Check: FAIL 1 File with 2 total versions checked between 2020-10-08 20:53:26 UTC and 2023-11-12 11:41:02 UTC Failed checks: ChecksumAuditLog id: 339276; file; checked_uri; date: 2023-11-12 11:41:02	Characterization: File Format: csv (Comma-Separated Values (CSV)) File Size: 144555 Original Checksum: d040aa26feb7d946d5242af2d39f69ac Mime Type: text/csv

User Activity	Date
---------------	------

5. How do users respond to datasets that must be accessed through Globus?

- Generally, once participants figured out why the Globus interface was the only download option through the hover text, there was an understanding of needing to go through that interface to get the data.

Data from: Nested circ

[Get Data from Globus](#) [More About Gl](#)

Given the size of this dataset (347 GB) you will need to use and social affiliation the Globus to transfer all files. Individual files may be downloaded depending on size.

the C. elegans depends on brain circuits mice produce ultrasonic vocalizations (USVs) during midbrain periaqueductal gray (PAG-USV neurons) and that both PAG-USV neurons and USVs can be s

[More\]](#)

Appendix A: Research Data Repository Usability Study Script

Expert Users

Part 1. Pre-Study Discussion

Question Type	Question
Ice Breaker	Thanks so much for participating in our study! How are you today?
Introductions	<p>We'll start by introducing ourselves. My name is Yasha Saxena, and I'm a 2nd year PhD student in Biomedical Engineering. [Pass to Shelley]</p> <p>My name is Shelley Dong, and I am a junior majoring in Psychology and minoring in Computer Science. [Pass to Yasha]</p> <p>Shelly and I work for the Assessment & User Experience Strategy department in the Libraries. Today we'll be talking in depth about your experiences with research data repositories, both as a depositor and as a user.</p>
Consent to record	<p>Before we get started, we would like to ask to record this session to review it and ensure our notes are as accurate as possible. This will only be used by AUXS staff and project team members to make improvements to the tool. If that sounds okay to you, I'll start the recording, and then I'll ask you again to confirm that you're okay being recorded.</p> <p>[Begin recording]</p> <p>Just to reconfirm, do you consent to being recorded through this session? (Yes or no)</p>
Context	Thanks so much. To get started, we'd like to hear a bit more about your background.
Background-1	What is your current title?
Background-2	What is your educational background?
Domain Knowledge-1	Have you submitted a research publication in the last 3 years?
Domain Knowledge-2	Would you briefly describe your experiences with using and/or publishing research data?
Domain Knowledge-3	Can you describe your previous interactions with the Duke Research Data Repository?
Domain Knowledge-4	Do you have experience using data from or depositing research data to any other data repositories?

Part 2. Upload Study

Task	Operator Script
Instructions	In this part of the study, I will walk you through our interface for submitting research datasets to the Duke Research Data Repository. It would be great if you could think about a recent dataset you have worked on or submitted as we go through all the steps required to complete a submission. We will discuss each section in detail as we go through it. Your feedback and patience are greatly appreciated.
Background on submissions	<p>Before we begin, would you mind sharing a bit about your background and experience with the dataset submission process in general?</p> <p><i>[Or, if already covered above] You mentioned earlier that you [have/have not] deposited datasets to the Duke Research Data Repository in the past. Is that correct?</i></p>
Setup	<p>Please return to this page in your browser.</p> <p>https://research.repository.duke.edu/ [paste into chat]</p> <p>This time, let's click on the "Deposit your data" option. You will be asked to log in with your netid.</p> <p>Click "Next" on this page to start the submission form.</p>
Restricted Information	<p>Please take a moment to read through the text on this page.</p> <ul style="list-style-type: none"> • • What is your experience working with restrictive data? • Are there terms on this page that you don't have a good understanding of? Need help to define?
Data Deposit Size	<p>Select "No" and then click "Next" to advance in the form. Please take a moment to read through the text on this page.</p> <ul style="list-style-type: none"> • How large are the datasets you've worked with in the past roughly? • Why do you think there are different buttons here? What is the purpose of these three different groupings? • Does the size of your dataset affect whether you're likely to deposit it in a repository?
Deposit Agreement	Click "Next" to advance the form. Scroll to the bottom of the agreement, select "I agree," and click "Next" again, unless you have any comments you'd like to share.
Submission Information	<p>Please read the description for each field, and then I'll ask you a few questions. For each field:</p> <ul style="list-style-type: none"> • Does the field description make sense to you? • Do you recall if this field was difficult to complete in the past? Why or why not? <p>For "List a name, ORCID identifier...":</p>

	<ul style="list-style-type: none"> Do you tend to submit an ORCID as a contact? Why or why not?
DOI Information	<p>Type the word “text” in each of the required fields, and then click “Next” at the bottom.</p> <p>Take a moment to read the text on this page.</p> <ul style="list-style-type: none"> How do you interpret this page? Have you ever had a situation where you did have a DOI for your dataset in advance?
CC Information	<p>Select “No” and click “Next” to advance the form.</p> <p>Take a moment to read the text on this page.</p> <ul style="list-style-type: none"> How do you interpret this page? How do you make a decision about the kind of license that you want to apply to your dataset?

Part 3. Post Upload Study Discussion

Question Type	Question
README	Have you been asked for a README during the submission process? How do you interpret this requirement?
Favorite Aspect	What was your favorite aspect of the submission interface? Why?
Least Favorite Aspect	What was your least favorite aspect of the submission interface? Why?
Free Text	Any additional comments?

Part 4. Download Study

Task	Operator Script
Discussion	<p>What experience do you have downloading datasets? From the Duke Research Data Repository specifically?</p> <p><i>[Or, if already covered above] You mentioned earlier that you [have/have not] downloaded datasets from the Duke Research Data Repository in the past. Is that correct?</i></p>
Discussion	<p>Now we’ll go through a series of tasks using the RDR. Please open this page in your browser and share your screen.</p> <p>https://research.repository.duke.edu/ [paste into chat]</p>

	Let's click on the "Browse datasets" option that you can see on the home page.
Search for a dataset	<p>Suppose you're a researcher looking to find data on the impacts of climate change on trees. How would you go about finding a dataset that might be relevant for you?</p> <p>[If their search includes the tree fecundity dataset somewhere in results] Please click on "Data from: Continent-wide tree fecundity..."</p> <p>[If their search does not include the dataset] Thank you. Next, please click on this link to a specific dataset: https://research.repository.duke.edu/concern/datasets/sn009z63q [past into chat]</p>
Dataset Homepage Summarize	What are your first impressions of this page?
Download single file with a long author list on homepage	Download the README.txt file contained in this dataset
Discussion	Did finding and downloading that file make sense to you? How was the experience of getting to the file browser? How would you expect to be able to download/find files in the dataset?
Discussion	Scroll to the bottom of the page, where there is a table with linked items. What is your impression of the rows without download buttons? What do you think those are?
Click a folder data type on the dataset homepage	Click on the name of one of the objects without a download button
Discussion	Do you understand what it is now? What is your impression of this approach to presenting the contents of the dataset?
Impression of title link	Scroll down to the table of items on this page. Notice that all of the titles are links. What do you think will happen if you click one of those links?
Click into a file set page	Click the title of the top file.
Discussion	What do you make of this page? Is any of this information useful? What is most confusing about it?
Click the usage stats button	Navigate back in your browser twice to get back to the dataset homepage and scroll up to the top of the page. Click the "usage stats" button on the right.
Discussion	(If applicable) Have you explored this page on your own datasets? What do you like about this page? What do you not like?
Go to dataset with only Globus option	Now enter the search terms "Nested circuits" and click into the entry that is found.

Discussion	Do you notice any differences between this page and the other dataset you just explored? What do you need to do to download a file here? Are you familiar with Globus?
Click Globus option	Scroll down to the table of items. Click one of the “Get file from Globus” links.
Discussion	What happened? What is your impression of this?
Click a failure file	Return to the dataset homepage in the other browser tab. On the dataset homepage, click the title link of the first file in the file browser.
Discussion	What do you notice about this page that is different from the other file details page we explored? What do you think the error means? Would this affect your decision to download the file?

Part 5. Post Download Study Discussion

Question Type	Question
Favorite Aspect	What was your favorite aspect of the submission interface? Why?
Least Favorite Aspect	What was your least favorite aspect of the submission interface? Why?
Free Text	Any additional comments?

Novice Users

Part 1. Pre-Study Discussion

Question Type	Question
Ice Breaker	Thanks so much for participating in our study! How are you today?
Introductions	We'll start by introducing ourselves. My name is Yasha Saxena, and I'm a 2 nd year PhD student in Biomedical Engineering. [Pass to Shelley] My name is Shelley Dong, and I am a junior majoring in Psychology and minoring in Computer Science. [Pass to Yasha] Shelly and I work for the Assessment & User Experience Strategy department in the Libraries. Today we'll be talking in depth about your experiences with research data repositories, both as a depositor and as a user.
Consent to record	Before we get started, we would like to ask to record this session to review it and ensure our notes are as accurate as possible. This will only be used by AUXS staff and project team members to make improvements to the tool. If that sounds okay to you, I'll start the recording, and then I'll ask you again to confirm that you're okay being recorded. [Begin recording]

	Just to reconfirm, do you consent to being recorded through this session? (Yes or no)
Context	Thanks so much. To get started, we'd like to hear a bit more about your background.
Background	What is your current title?
Background	What is your educational background?
Domain Knowledge	How often do you read research papers?
Domain Knowledge	Are you a part of a research lab or group?
Background	Describe your experiences with using and/or publishing research data.

Part 2. Download Study

Task	Operator Script
Discussion	<p>What experience do you have downloading datasets? From the Duke Research Data Repository specifically?</p> <p><i>[Or, if already covered above] You mentioned earlier that you [have/have not] downloaded datasets from the Duke Research Data Repository in the past. Is that correct?</i></p>
Discussion	<p>Now we'll go through a series of tasks using the RDR. Please open this page in your browser and share your screen.</p> <p>https://research.repository.duke.edu/ [paste into chat]</p> <p>Let's click on the "Browse datasets" option that you can see on the home page.</p>
Search for a dataset	<p>Suppose you're a researcher looking to find data on the impacts of climate change on trees. How would you go about finding a dataset that might be relevant for you?</p> <p>[If their search includes the tree fecundity dataset somewhere in results] Please click on "Data from: Continent-wide tree fecundity..."</p> <p>[If their search does not include the dataset] Thank you. Next, please click on this link to a specific dataset:</p> <p>https://research.repository.duke.edu/concern/datasets/sn009z63q [past into chat]</p>
Dataset Homepage	What are your first impressions of this page?
Summarize	What is your impression of what this dataset contains? Can you summarize what it may be about based on skimming the page?
Download single file with a long author list on homepage	Download the README.txt file contained in this dataset

Discussion	Did finding and downloading that file make sense to you? How was the experience of getting to the file browser? How would you expect to be able to download/find files in the dataset?
Discussion	Scroll to the bottom of the page, where there is a table with linked items. What is your impression of the rows without download buttons? What do you think those are?
Click a folder data type on the dataset homepage	Click on the name of one of the objects without a download button
Discussion	Do you understand what it is now? What is your impression of this approach to presenting the contents of the dataset?
Impression of title link	Scroll down to the table of items on this page. Notice that all of the titles are links. What do you think will happen if you click one of those links?
Click into a file set page	Click the title of the top file.
Discussion	What do you make of this page? Is any of this information useful? What is most confusing about it?
Go to dataset with only Globus option	Now enter the search terms “Nested circuits” and click into the entry that is found.
Discussion	Do you notice any differences between this page and the other dataset you just explored? What do you need to do to download a file here? Are you familiar with Globus?
Click Globus option	Scroll down to the table of items. Click one of the “Get file from Globus” links.
Discussion	What happened? What is your impression of this?
Click a failure file	Return to the dataset homepage in the other browser tab. On the dataset homepage, click the title link of the first file in the file browser.
Discussion	What do you notice about this page that is different from the other file details page we explored? What do you think the error means? Would this affect your decision to download the file?

Part 3. Post Download Study Discussion

Question Type	Question
Favorite Aspect	What was your favorite aspect of the submission interface? Why?
Least Favorite Aspect	What was your least favorite aspect of the submission interface? Why?
Free Text	Any additional comments?

Part 4. Upload Study

Task	Operator Script
Instructions	<p>In this part of the study, I will have you fill out information on a mock data set in our “Data Submission Form.” Please feel free to stop and ask me about anything you are confused about as you complete the assignment. You will imagine that the dataset you are submitting is the tree fecundity dataset from the previous part of the study. We have prepared a list of the information you will need from this dataset to complete the submission.</p> <p>Let’s begin by navigating to the “Data Submission Form.”</p> <p>Please return to this page in your browser.</p> <p>https://research.repository.duke.edu/ [paste into chat]</p> <p>This time, let’s click on the “Deposit your data” option. You will be asked to log in with your netid.</p> <p>Click “Next” on this page to start the submission form.</p>
Policies and Guidelines	<p>From this point, please go ahead and fill out the submission form as if you are trying to deposit the tree fecundity dataset. Please try to speak your thoughts out loud as you are completing the form, including your initial impressions of each question on the form and how you decide what to select or type.</p> <p>See if they explore the links or just move on to the next page.</p>
Restricted Information	<p>Are they confused about whether the dataset has restricted information? Do they click the additional info link? Do they ask questions? Do they say yes? Ask them why.</p>
Data Submission Form	<p>This dataset has over 300GB. Let them play around with the page, observe what they do/ask, and advise them to click one of the lesser options to move on with the study.</p>
Deposit Agreement	<p>Take any notes, but not very critical for the study.</p>
Submission Information	<p>Let them fill out the fields based on what is in the dataset homepage. Encourage them to speak out loud if they are confused about anything, need clarification on anything.</p> <p>If pressed for time, tell them they only need to fill out the required fields, but we would still like to hear their thoughts about the other fields.</p>
DOI Information	<p>Do they recognize the DOI in the dataset homepage and say yes? Do they understand what a DOI is?</p>
Creative Commons Information	<p>Take any notes if they click on any of the links or say anything out loud. These users will likely not know what this is, but they should see that the dataset being used for the study opted for CC0.</p>
Review and Submit	<p>Stop the study before they submit</p>

Part 5. Post Upload Study Discussion

Question Type	Question
README	Have you been asked for a README during the submission process? How do you interpret this requirement?
Favorite Aspect	What was your favorite aspect of the submission interface? Why?
Least Favorite Aspect	What was your least favorite aspect of the submission interface? Why?
Free Text	Any additional comments?

Dummy Submission Information for Novice User Upload Study

These metadata come from the dataset entitled [Data from: Continent-wide tree fecundity driven by indirect climate effects](#).

Task	Operator Script
Enter the full title for the study/dataset	Data from: Continent-wide tree fecundity driven by indirect climate effects
Enter the creators/authors (principal investigator, co-investigator(s)) of this study/dataset as follows: Lastname, Firstname (MI optional); Lastname, Firstname (MI optional); etc.	Zhu, K.; Zlotin, R.; Veblen, T.T.; Bergeron, Y.; Whitham, T.G.; Myers, J.A.; Daley, R.; Moore, C.M.; Kilner, Christopher L.; Rodman, K.C.; Whipple, A.V.; Clark, James S.; Gilbert, G.S.; Schlesinger, William; Aubry-Kientz, M.; Hille Ris Lambers, J.; Parmenter, R.; Franklin, J.F.; Kunstler, G.; Legg, K.L.; Greenberg, C.H.; Reid, Chantal; Ibanez, I.; Das, A.J.; Pearson, S.; Fer, I.; Koenig, W.D.; Poulton-Kamakura, Renata; Courbaud, B.; Lutz, J.A.; Johnstone, J.; Knops, J.; Stephenson, N.L.; Guo, Q.; Andrus, R.; Redmond, M.D.; Bragg, D.C.; Cohen, S.; Nunez, C.; Brockway, D.; Moran, E.; Shanahan, E.; Wion, A.P.; Sutton, Samantha; Bogdziewicz, M.; Scher, C. Lane; Luongo, Jordan; Macias, D.; Ready, Ethan; LaMontagne, J.M.; Gehring, C.A.; Sharma, Shubi; Swift, Maggie; Dietze, M.; Steele, M.; Cleavitt, N.L.; Schwantes, Amanada; Messaoud, Y.; Myers, O.B.; Swenson, Jennifer J.; McIntire, E.J.; Fahey, T.J.
Enter additional contributors (non-authors) to this study/dataset as follows: Lastname, Firstname (MI optional); Lastname, Firstname (MI optional); etc. Please note contributors will not appear as creators/authors in the citation created for this deposit.	Leave blank

<p>Enter department, school, research center or other organizational affiliation. Separate each with a semicolon.</p>	<p>Leave blank</p>
<p>List a name, ORCID identifier, phone number and/or email address for a designated contact for this study/dataset. This contact is responsible for answering questions about the data, documentation and/or code for this project.</p>	<p>James Clark: jimclark@duke.edu, ORCID: 0000-0002-5677-9733</p>
<p>Enter a description for your study/dataset. An abstract from a grant or publication is acceptable.</p>	<p>Indirect climate risks for tree fecundity that come through variation in size and growth (climate-conditions interactions; CCI) are not currently part of models used to anticipate forest regeneration.</p> <p>Yet CCI may be among the most-important mechanisms needed to accurately predict forest responses to climate change. Meta-analyses of trends and species distribution models to predict them will be misleading if responses depend on the condition of individual trees.</p> <p>A synthesis of tree species in North America shows that CCI dominate through two pathways, i) effects of growth on fecundity that depend on climate, and ii) effects of climate that depend on tree size. Because fecundity declines in large trees, climate changes that stimulate growth move small trees into more fecund and large trees into less fecund size classes. There is a biogeographic divide to this CCI, reducing fecundity in the West and increasing it in the East. Continental-scale responses to climate change are thus dominated by interactions at the individual scale, highlighting the importance of ecosystem management that considers multiple demographic rates.</p>
<p>Enter one or more key words to describe your study/dataset (e.g. Mechanical engineering, Sociology, Medical imaging, etc.). Please separate each key word with a semicolon.</p>	<p>Trend attribution; Indirect climate effects; Seeds-- Production (Biology); Climatic changes; Forest transformation</p>
<p>All remaining fields</p>	<p>Leave blank</p>

Appendix B: Recruitment Process and Text

Novice Recruitment via Qualtrics

For novice recruitment, we used the 2023-2024 STEP Qualtrics List (combination of people who agreed to be contacted after 2023 biennial student survey and prior STEP members). We excluded people who had already been contacted for the Fall 2023 data party. From the remaining students, we filtered to graduate students (either DUL or professional school libraries) and created a sample of 56. On top of that we directly invited 4 students who wanted to attend the data party but either had a last minute conflict (1) or were on the waitlist (3). We closed the survey after 15 people filled it out, and we selected 6 for the study.

Message Text

Subject:

\$15 for helping the Libraries: User study on research data

Body:

Dear \${m://FirstName},

The Duke Libraries are looking for graduate students to participate in a virtual user study focused on our Research Data Repository. The study will last less than an hour, and participants will receive a **\$15 Amazon gift card** as compensation for their time.

If you're available and interested, follow this link to **[\\${i://SurveyLink?d=fill out the volunteer form}](#)**.

The form will remain open until all spots are filled. Thank you for your willingness to help the Libraries improve to better meet your needs!

Best,

Angela Zoss

--

Angela M. Zoss, Ph.D. (she/her/hers)

Interim Head, Assessment & User Experience Strategy department

Duke University Libraries

Phone: 919-684-8186

Email: angela.zoss@duke.edu

I pledge to oppose bigotry, xenophobia, and racism in all their forms, to practice self-awareness, and to make equitable choices daily.

Follow the link to opt out of future emails:
\${!://OptOutLink?d=Click here to unsubscribe}

Expert Recruitment via RDR Product Council Member

A staff member who was a member of the RDR Product Council volunteered to contact RDR depositors directly to ask them to volunteer for the study. A total of 13 RDR users were invited to participate.

Message Text

Dear [First Name],

The Duke Libraries are looking for individuals with prior experience submitting datasets to the Research Data Repository to participate in a virtual user study. The study will last between 30 and 45 minutes, and participants will receive a **\$20 Amazon gift card** as compensation for their time.

If you're available and interested, follow this link to **fill out the volunteer form** [add link to form].

The form will remain open until all spots are filled. Thank you for your willingness to help the Libraries improve to better meet your needs!

Best,
[Name]