

WHY THE WIZARD?

The Data Sharing Wizard is created to teach students on how to answer the questions:

- **Who owns a dataset?**
- **What are the legal rights the data creator has?**
- **Who gets to make decisions about a research dataset?**

The Problem

Confusion over intellectual property of research datasets emerged as a top issue for students in a series of interviews that we performed at Oregon State University in 2015:

- Intellectual property issues in research data are complex. They are case dependent and involve multiple partners.
- Lack of guidance and resources to answer their data related questions.
- Many faculty and researchers at the university also cannot help their students on the questions of dataset ownership and legal rights.

Our Answer

The Data Decision Making Wizard! The Wizard is a Scenario Based Learning (SBL) tool that offers an interactive approach to engage students in learning about authors rights in the context of decision making related to research datasets.

With the Wizard, students and researchers are able to identify:

- (1) the people involved in data decision making
- (2) the documents, laws, or policies that have a role on the data decision making for the dataset in question and
- (3) the resources on campus and beyond that they have access to that may help them answer their questions.

Version 1.0 is available in <http://dcwizard.library.oregonstate.edu/> (<http://dcwizard.library.oregonstate.edu>)

FLOW OF THE WIZARD

This video shows the basic workflow of the Data Sharing Wizard. It is 2:03 minutes long.

[VIDEO] https://www.youtube.com/embed/_Hwn_61QLRE?feature=oembed&fs=1&modestbranding=1&rel=0&showinfo=0

NAVIGATING BAR

This video shows how to revise previously made options with the Data Sharing Wizard. It is 2:06 minutes long.

[VIDEO] <https://www.youtube.com/embed/QmHkuKHqGgk?feature=oembed&fs=1&modestbranding=1&rel=0&showinfo=0>

SAVING LOGS

This video shows the Step Log feature of the Data Sharing Wizard. It is 2:12 minutes long.

[VIDEO] <https://www.youtube.com/embed/SvPNDf-XXxA?feature=oembed&start=2&fs=1&modestbranding=1&rel=0&showinfo=0>

BACKEND

The Data Sharing Wizard is a single-page web application built with the ReactJS JavaScript library.

Why use the single-page architecture

- The Wizard is fast loading and friendly to mobile devices.
- It is easier to implement as the Wizard requires minimal server-side setup and development.
- Latest technologies (Javascript, modern browser) provide better support for the development and debug of single-page application.

Wizard Implementation

- Server side: CentOS Linux with Node.js
- Client side: React Javascript library
- Use React Route to track current user selections and the history of selections user has made
- The step log saves all the user selections together with the answers and resources provided by the Wizard
- User can export the step log and save as a pdf file

Oregon State University Data Sharing Wizard

Questions?
Contact the OSU Research Data Services at
researchdataservices@oregonstate.edu

Current Step:
Formal documents

Previous Steps:
[Your questions](#)
[Decision maker](#)
[Home](#)

Topic: What can I do with my own data? [Show Step Log](#)

Question: Are there formal agreements or documents that limit what you can do with the data?

☐ Yes, and they clarify what I can do with the data. ☒ No. If they exist, they are not helpful.

Explanation:
You may be the person who gets to make the decision of making the data available, sharing it or publishing. Before doing that, though, you should make sure that there aren't any other formal agreements that make a difference over what you can do with your data. There may be formal agreements related with where you work: lab policies, university policies, etc. Formal agreements with your funders or other institutions or companies that you are working with: contracts, data sharing agreements, funder requirements, etc. Consider, also, if there are any discipline specific practices that you are expected to follow, even if they are not written anywhere. For example, in some disciplines sharing your data in specific repositories is expected, and you would not be considered a good professional if you did not do it.

Resources:

FEEDBACK

Usability study:

We designed a usability study in which we asked grad students and undergraduate students with experience in research to complete a questionnaire.

The questionnaire had three scenarios that addressed issues related to

- Data sharing
- Data ownership
- Reuse of publicly available data

We asked the participants to give us answers to the scenarios **before** and **after** using the wizard, to see if their answers showed a better understanding of these processes after using the wizard.

Results

We sent the questionnaire to 6 people and got 2 responses. There were 3 scenarios, so we have 6 examples of how two different people learned or did not learn by using the Wizard.

We observed that 1 out of 6 answers show more confusion on the topic (data ownership) after using the Wizard, but that 5 out of 6 answers show an improvement after using the Wizard.

These are the 3 ways that the answers show improvement:

1. Improvement is shown by giving answers that are more detailed, and that justify the answer.

Example 1: justified answers

After asking if the student in the scenario could use a dataset in his research...

Answer before: *"Yes, he should site the repository and authors accordingly in his paper."*

Answer after: *"Yes, since there is no clear license, he must go off copyright law. In this case, the data presented is not covered by copyright as it is composed of facts so Blake can use the data in his paper so long as he cites the data source."*

Both answers are right, but the second answer shows a better understanding about why in this particular case the reuse of data is appropriate.

2. Improvement is shown by answering a question right after having answered it wrong the first time.

Example 2: right after wrong

When asking who owns a dataset in a particular scenario to figure out who can make decisions about the dataset...

Answer before: *"NSF or the two researchers she worked with."*

Answer after: *"Since it is a grant funded project, OSU - the receiver of the grant - owns the data and will then be the entity who decides if Anne can take her research along with her"*

The first answer is partially incorrect. NSF and the researchers can have a role on decision making, but the owner of the data is the organization that got the grant, OSU. The answer after using the wizard identifies the right data owner.

3. Improvement is shown by expressing awareness of more options after using the Wizard.

Example 3: more options

When asking what actions should a student do to share a dataset while making sure that the other group won't publish the results before her...

Answer before: *"She could wait until she has published her thesis before letting Jie use her data. She also should talk to her research advisor."*

Answer after: *"She'll need to make a formal agreement with Jie."*

The answer before is a possibility, but the answer afterward is also a good option. The second option would give the student in the scenario more control and more options in her research.

Future of the wizard

The work of the wizard has been planned in 5 phases:

Phase 1: Development of the first version of the Wizard, v1.0, that we are presenting in this poster at ACRL in April 2019.

Phase 2: Improvement on the wizard based on feedback from colleagues and the preliminary usability study presented here. This phase will culminate in August 2019 with the release of version v2.0

Phase 3: Promotion of the Wizard, mostly within the university. Usability study in early fall 2019.

Phase 4: Improvement of the Wizard based on feedback from the usability study from phase 3. Writing of a case study with a description of the wizard and the results of the usability study. Archival of a snapshot of the code and content of the wizard in ScholarsArchive@OSU. Planned for late 2019 and early 2020.

Phase 5. Maintenance.

The code and content of the wizard will be shared via GitHub with an open license when version v2.0 is available.

AUTHOR INFORMATION

Clara Llebot Lorente is the Data Management Specialist at Oregon State University. She has a PhD in oceanography and worked as a researcher for several years in the field of biogeochemical modeling. Now Clara works at the OSU library to support researchers to manage their research data.

Hui Zhang is the Data Services Librarian at Oregon State University. He has a PhD in Information Science from Indiana University Bloomington with interest in information retrieval. His work at OSU libraries including support of discovery and digital repositories services.

ABSTRACT

Who owns a dataset and what are the legal rights in a dataset? Who gets to make decisions about a research dataset? These are common questions that researchers ask scholarly communication librarians or data specialists. Confusion over intellectual property of research datasets also emerged as a top issue for students in a series of interviews that librarians from a public research university conducted in 2015. Copyright issues in research data are complex because they are case dependent and involve multiple partners (Carroll, 2015). Students face barriers due to a lack of guidance and resources to answer their data related questions. The problem also affects faculty and researchers at the university, who are not always knowledgeable enough about the topic to help their students. This poster describes the Data Usage Wizard (the Wizard), a Scenario Based Learning (SBL) tool that offers an interactive approach to engage students in learning about copyright in the context of decision making related to research datasets.

Using the Wizard, students and researchers are able to identify (1) the people (2) the documents, laws, or policies that have a role on the data decision making for the dataset in question and (3) the resources on campus and beyond that they have access to that may help them answer their questions. With these learning goals in mind, we designed a series of questions and answers structured around 3 main questions:

- What can I do with my data?
- What can others do with my data?
- What can I do with other people's data?

Each question asked by the Wizard is accompanied by (a) an explanation (b) resources to learn more about that question, and (c) a collection of 2-5 predefined answers. Choosing an answer will prompt the user to the next question. The Wizard is a single-page web application built with the ReactJS JavaScript library, it loads and re-writes page contents dynamically from the server without interruption of user experience. The input of the Wizard is a JSON file that encodes the underlying structure of the questions. Users of the Wizard can export all questions and answers as a "handout" of the session.

Attendees to this poster session will learn the motivations of the authors of the Wizard to create the tool, the result of the project, and how the Wizard has been received in our campus as well as an initial user analysis. They will also learn about possibilities to reuse and adapt the content and the software for their own institutions, and our plans for the future, which have facilitating reuse of the tool by others as one of the main goals.