

# Data Sharing Seminar Series for Societies

## Data Sharing and Citation: How Societies Can Make a Difference February 5, 2021

10:08:29 From Laura Lyon, AGU to Everyone : FABBS Data Standards for Behavioral and Brain Sciences Webinar Series: <https://bit.ly/39VQVUI>

10:23:01 From Shelley Stall, AGU to Everyone : DataCite Commons: <https://commons.datacite.org>

10:24:48 From Diana Schlamadinger to Everyone : How do you decide what a citable dataset is? How big? What constitutes a citable dataset?

10:27:15 From Shelley Stall, AGU to Everyone : Connected Research: The Potential of the PID Graph, <https://doi.org/10.1016/j.patter.2020.100180>

10:28:48 From Jennie Larkin to Everyone : Connecting funding sources into the PID graph would also be valuable.

10:28:50 From B&E Wilson to Everyone : Is there a distinction between “this paper describes this dataset” vs “this paper uses this dataset”

10:28:57 From Michael N. Evans to Everyone : Q for afterward: Geological Society of America uses a repository system (e.g. for the journal Geology) in which you must search their website for a pdf of data, then download it, then redigitize it... archival, perhaps, but not very 'FAIR'. Any plans to work with them to bring them into the 21st century, for their existing repository?

10:29:43 From Elisha Wood-Charlson (she/her) to Everyone : Is there any interest or effort to have citations metrics to show up on someone's ORCID?

10:30:19 From Martin Halbert to Everyone : Helena, are you working directly with bibliographic, subject, and citation databases?

10:30:38 From Mark Servilla to Everyone : I see this as elevating research data to the same level as a science manuscript publication (this is, a 1st order object). Do you have thoughts on the peer review process for research data?

10:31:20 From Elisha Wood-Charlson (she/her) to Everyone : Is there a structured schema for meta-analysis? (Recombination of data sets)

10:31:57 From B&E Wilson to Everyone : @Diana — there's also some interesting debate going the other direction, creating DOI for a collection of data. In one specific case, it's understanding the best way to cite a collection of airborne data collection campaigns. Each campaign has its own DOI. But there's also value to the community of describing the full set.

10:32:43 From Graham Smith to Everyone : @Diana there is a Research Data Alliance group on data granularity addressing this very relevant question: <https://www.rd-alliance.org/groups/data-granularity-wg>

## Data Sharing Seminar Series for Societies

10:33:28 From susan borda to Everyone : Helena - Can a data repository or IR submit to DataCite a citation indicating a relationship between a dataset and a paper even if the dataset citation is not in the "References" section of the paper but mentioned elsewhere in the paper?

10:34:53 From Kyle Copas to Everyone : @Susan, we at GBIF (have to) do this all the time

10:36:08 From susan borda to Everyone : @Kyle Good to know!

10:37:18 From susan borda to Everyone : Our researchers/depositors have not been great at actually referencing their own datasets yet in our repository they cite a related paper.

10:37:25 From B&E Wilson to Everyone : @Helena @Kyle The ORNL DAAC has this sort of data as well, but haven't submitted it to anyone. We use the language of a paper refers to a dataset vs a paper cites a dataset. It might be interesting to see if that could be submitted. We have that cite/refer data for a decade or so. But, cobbler's kids don't have shoes, that particular data isn't FAIR. But I'd be willing to share it.

10:40:12 From Dennis Brown to Everyone : For better or for worse, our entire reward system is built upon the prestige of the journal in which our work is published. We need a major reeducation effort to make institutions change their assessment process, as well as helping researchers themselves sort out the meaningful data from the garbage that appears in the repositories.

10:41:55 From Jonathan Petters to Everyone : To pile onto the challenges already addressed, there's a lot of human subjects or otherwise sensitive data to be linked/made accessible (or not)...researchers need guidance in what to do with this data, and when

10:42:05 From Parnian Kiani to Everyone : I'm wondering if there is a Framework for sharing sensitive data like clinical Trials?

10:42:32 From Jake Yeston to Everyone : There's a good discussion on the CrossRef site here about how to specify relationships (eg, between datasets and papers and whether the dataset has just been created in association with that paper, or is just being referenced) <https://www.crossref.org/education/content-registration/structural-metadata/relationships/>

10:43:00 From Eric DeWeaver to Everyone : Dear Shelley: the biggest challenge to geoscience data sharing in the US is the government's desire to buy satellite data rather than flying in-house satellite missions. Purchased data is always cheaper with license restrictions and NOAA and NASA have already purchased data under strong restrictions. Is AGU looking at this issue? One reference here is the WMO conference convened to look at the issue, in particular this presentation: 06\_Rick Anthes\_WMO Presentation Final.pdf

10:44:24 From Sylvia Spengler NSF to Everyone : Love idea of data as world heritage

10:44:49 From susan borda to Everyone : Article mentioned in Helena's talk: <https://www.nature.com/articles/sdata2018259>

10:45:22 From susan borda to Everyone : "A data citation roadmap for scientific publishers"

## Data Sharing Seminar Series for Societies

10:46:02 From Hans-Georg Weinig GDCh to Everyone : Germany started the huge project "National Research Data Infrastructure" (<https://www.nfdi.de/en-gb>). GDCh, the German Chemical Society is partner in the sub consortium "NFDI4Chem" (<https://nfdi4chem.de/>

10:46:18 From Hans-Georg Weinig GDCh to Everyone : We are working towards worldwide standards

10:46:29 From Anne Cambon-Thomsen to Everyone : Re question from Wilson about difference in citation between "data cited" and "data used" in a paper, we published in 2015 a guideline that offers a system (format of citation) that makes the difference; you can have a look at this BMC medicine paper: <https://bmcmmedicine.biomedcentral.com/articles/10.1186/s12916-015-0266-y>:

10:46:49 From susan borda to Everyone : Also from Helena's talk: "A data citation roadmap for scholarly data repositories" <https://www.nature.com/articles/s41597-019-0031-8>

10:47:41 From Darrell Kaufman to Everyone : I am very happy to hear that journal editors are being engaged to help lead the charge. In my view, they have more influence than societies in promoting FAIR data practices.

10:48:21 From Nicholas Jarboe to Everyone : And reviewers.

10:48:33 From Sylvia Spengler NSF to Everyone : you see the empty spaces in the graph of the ARDC figure.. Are those real?

10:49:00 From Shelley Stall, AGU to Everyone : NISO Plus 2021 Meeting: <https://nisoplus21.sched.com>

10:49:14 From Jake Yeston to Everyone : Q: I'd love to hear more about ROR—are those ID's going into metadata on papers or datasets on a regular basis?

10:50:00 From Melissa Cragin to Everyone : +1 to Jake's question.

10:50:40 From Kevin Lindstrom to Everyone : I know that the WHOI has a ROR. I'm wondering how they are using their id pid graphs for WHOI publications?

10:51:09 From Richard Nakamura to Everyone : Please contact me about a NAS meeting on this topic which will occur virtually on April 28-29.: [Nak1963@gmail.com](mailto:Nak1963@gmail.com)

10:51:38 From Kristian Garza to Everyone : Here it's the current recommendation for the relationTypes for citations [https://support.datacite.org/docs/relationtype\\_for\\_citation](https://support.datacite.org/docs/relationtype_for_citation)

10:52:08 From Kevin Lindstrom to Everyone : I know our university has a ROR but if you asked our researchers what it is you might see some puzzled faces

10:53:25 From Rachael Lammey to Everyone : @Jake we'll start to accept RORs as part of Crossref metadata in our next schema release (coming as soon as we can, we're working on it now). But DataCite is already collecting RORs in their schema which is great. And you can start to search using ROR IDs in DataCite Commons etc. Which is really cool.

## Data Sharing Seminar Series for Societies

10:53:36 From Nicholas Jarboe to Everyone : Add "Accurate" to FAIR?

10:53:44 From Jake Yeston to Everyone : Thanks—that's great to know!

10:53:52 From Kevin Lindstrom to Everyone : I've applied for a ROR for one of our departments. This is a discussion for another day but it is not always clear what is understood as an institute

10:54:24 From Kyle Copas to Everyone : Data papers are intended to function essentially as peer-reviewed metadata.

10:54:53 From Nicholas Jarboe to Everyone : Data should be ready at paper submission and the reviewers can review them.

10:55:06 From Howard Ratner to Everyone : Hi we at CHORUS have been working to report on datasets of funded research for societies, funders, institutions and publishers. How can societies help promote the mandatory inclusion of dataset creators and institutions?

10:56:37 From Mark Hahnel to Everyone : We find that 7 out of 8 don't go in the reference lists

10:57:20 From Laura Lyon, AGU to Everyone : Hi everyone, if you have any additional thoughts, comments, or questions after the webinar, you can add it to our Google Doc and we will include it with the chat: <https://bit.ly/3tuTDaY>

10:57:27 From Michael N. Evans to Everyone : What is your position on systems like figshare, which are sometimes used for data repositories? Does it matter as long as it's FAIR?

10:58:41 From Kristian Garza to Everyone : here is more guidance on how data repositories can provide relationships <https://support.datacite.org/docs/contributing-data-citations>

10:59:23 From Howard Ratner to Everyone : @Jake - Once ROR is part of the Crossref Schema (this Spring) publishers will be paying more attention

11:00:02 From James Liu to Everyone : what is the difference between ROR and Ringgold?

11:00:03 From Rachael Lammey to Everyone : +1 Howard - we'll be doing a lot of outreach when we have a firmer date for that. But there is a publisher adoption WG as well.

11:00:15 From Erin Robinson to Everyone : +1 to PID webinar and societies

11:00:23 From susan borda to Everyone : Thanks Kristian! I just wasn't clear on whether we, as a data repo, could indicate this even if the paper didn't officially cite the dataset in the references section. Good to get this clarification from Helena.

11:00:32 From Dan Lear (MBA) to Everyone : That maturity level for PID's would be great

11:00:33 From Nici Pfeiffer to Everyone : can you share the link to that paper?