Description

In the second part of this interview series, Anita shares with us the importance of Research Resource Identifiers (RRIDs) and how they can improve the value of a research paper. She also discusses the importance of reproducibility in academic research and highlights how journals, publishers, and funding agencies are requiring strict adherence to reproducibility guidelines. In addition, Anita also talks about the impact SciCrunch has on researchers. In this interview, she mentions how SciCrunch unifies multiple databases on a single platform, making it easier for researchers looking for specific data.

Kuntan: Can RRID increase the impact factor of a paper as well?

Anita: That is a very good question. It seems like it should. If you do a good job at labeling and are doing your work very carefully, that should increase your paper's impact. I think that no matter what else is going on, careful work should be something we all strive for. We can all take better care in doing our science and reporting on it. That should be its own reward and I think this is something we need to measure. We need to look at journals in which the same kind of paper is published with and without a RRID, or set of RRIDs, and then try to figure out whether people can find reagents more easily when the RRIDs are known.

I think this will start to emerge over the next few years. This initiative is still relatively new. We now have a lot of uptake in some big journals. However, if one *Cell* paper is published without a RRID and another with a RRID even if they are published only a few months apart. In a few years, we will know the impact of papers with and without a RRID and possibly be able to examine the likely increase in the impact factor of journals because of RRIDs. At present, we do not have enough data. It is a little bit too new, but there is no reason to be sloppy. Being a little bit better in reporting these materials can only help you, your colleagues, and the people reading your article. If we can help them read more easily, I think they will cite you more frequently than they would if they had to track down the antibody you used.

Kuntan: Some publishers and institutions have their own data repositories like *Nature's* Protocol Exchange, where researchers can store their research data. Does SciCrunch unify these systems?

Anita: We have not really reached out to a lot of institutional repositories. We have aggregated a lot of

the big and small open databases like ModelDB, which is not really an institutional repository but a database with a particular kind of data. We find that the institutional repositories we have looked at typically have data at random. Dryad or Figshare are generic repositories, but you can do less with Figshare or Dryad data than with PDB data. In a repository like PDB, you are going to see, for instance, x-ray crystallography. However, in Dryad, it could be ice core samples, or it could be temperature measurements of a mouse. We find that if you put a lot of data into a "community repository," it becomes better, richer, and actually of value to a particular community.

I keep using ModelDB as an example; however, there are now thousands of neuronal models on different platforms like Neurogenesis, Matlab, and others. If one wanted to find the code for all the neuronal models that cover the hippocampus, one would actually be able to search that. Another repository, neuromorpho.org, simply has traces of neurons exportable into models. Again, it does not become interesting to search as a data set until there are a whole lot of things in there. However, neuromorpho and all of these other community data repositories have been growing very significantly over the last few years. People are depositing more and more data into public repositories, which is wonderful. I think in the future, we are going to have a bigger ecosystem with many community repositories.

Anita Bandrowski – Identifying research resources in biomedical literature should be easy (2014) from **INCF** on **YouTube**

Kuntan: When constructing a paper, what is the most critical part to ensure reproducibility?

Anita: Methods. Work on the methods. It is not great if you keep using the same protocol, like a daisy chain. Reusing protocols is absolutely a terrible practice, one of the worst things, because compared to the paper published almost 15 years ago, there were completely different methods used along with different sets of reagents.

One of the easiest things is to think of it as a recipe. Some journals now ask for such lists in a tabular format. However, even if they don't, it is a really easy thing to provide by adding the catalog number. Here is the RRID for a particular resource. Would this make your research more or less reproducible? You need to know exactly which resource was used in order to reproduce the research.

First, you need to check the list of ingredients. Second, you have to know exactly what you did, the exact protocol. The *Journal of Visualized Experiments* (JoVE) helps record the exact protocol that was followed in a video-based format. They follow you around with a camera and record your experiments along with its processes. In other words, the method section is really critical and should be written very carefully. You should really focus on writing a good, complete methods section with a list of reagents and protocols. Use lots of pictures, tables, and graphs to show people how to reproduce the study because this will really help bring together all the information that somebody else may need.

Kuntan: What are some of the most common issues encountered by authors or journals in terms of reproducible research?

Anita: A recent paper by Leonard Freedmen from GBSI examined the places where reproducibility is a problem. About half of the reproducibility problems have something to do with reagents. Therefore, if we solve this problem, we will be solving a good portion of the reproducibility problem. Another place is statistics. Having a statistical reviewer on your journal's editorial board is a wonderful thing. It is also important to ask your statistics colleagues to look over your statistics to make sure the paper is robust. Are you using enough males and females in the particular study? Most of the time, people who used mice didn't report whether they were using male or female mice. When they were using rats, they were using only male rats. When they were using humans, they mostly divided equally between both populations. However, imagine what this can do to drug trials. So now, you are basing a clinical trial for human males and females based on studies strictly using male rats. Many clinical trials fail due to this because maybe the animal subjects are not good enough, or we are not using enough females. Many people are not randomizing or blinding, and these are very basic methods of ensuring that the study is reproducible since these methods remove bias.

Kuntan: Will Open Science or Open Data resolve the reproducibility issue or will it add to it?

Anita: That is a very good question. Certainly, if the data is accessible, then certain things will come out faster, and perhaps because they come out faster, people will be able to catch problems earlier. However, by itself, releasing data is not going to make that data better. Openness is a very good policy, and I believe people will potentially pay much more attention to it if I am going to put my reputation on the line with my data. In such a situation, I will look at the data more frequently and more carefully. However, actively sharing data will probably make it a more important part of the paper and a more important part of the scholarly work.

(To be Continued)

(This interview is a part of our interview series of Connecting Scholarly Publishing Experts and Researchers.)

Category

1. Expert Interviews
2. Industry News

Date Created
2017/11/09
Author
daveishan