Know More About SciCrunch and RRIDs: An Interview With Dr. Anita Bandrowski (Part 1)

Author

Enago Academy

Post Url

https://www.enago.com/academy/know-more-about-scicrunch-and-rrids-an-interview-with-dr-anita-bandrowski-part-1/

Research studies involve the analysis of crucial data and drawing appropriate conclusions from that data. With the constant increase in the number of research studies, the amount of data generated is now enormous, thus making it difficult for researchers to find the right data for their research study. In this interview, Enago's Kuntan Dhanoya (Vice President, Business Development) had the opportunity to speak with Anita Bandrowski, Founder and CEO of SciCrunch. SciCrunch is a data sharing and display platform designed for dynamic data and allows communities of researchers to create their own portals in order to provide access to resources, databases, and other tools that are relevant to their research.

Anita is currently the Project Lead at the Centre for Research in Biological Systems at the University of California, San Diego. Prior to SciCrunch, she was also the Scientific Lead and worked on the Neuroscience Information Framework Project (NIF). NIF maintains the largest searchable collection of neuroscience data and biomedical resources.

In the first part of this interview series, Anita gives us a brief understanding about SciCrunch and RRID, which help researchers cite the key resources used to produce

the scientific findings reported in biomedical literature. She talks about how authors can use RRIDs in their research and discusses the rising problem of scientific reproducibility. She also sheds some light on the challenges encountered by authors and journals concerning the reproducibility crisis and shares some effective measures to mitigate this issue.

Kuntan: SciCrunch is a data sharing and display platform for researchers where they can cite key resources from their scientific studies. Could you tell us more about SciCrunch and how it benefits the research community?

Anita: SciCrunch emerged from a project called the Neuroscience Information Framework, which asked us to put together all the important information and data from different databases that anyone would want to know about neuroscience. We were at the lead institution at the University of California at San Diego (UCSD) out of a five-institution project. Some of the other famous institutions that were part of the project are Harvard, Yale, and Caltech. Together, we were able to start creating a system that would allow us to put a lot of data together. However, the National Institutes of Health asked us specifically to help researchers find antibodies. It turns out that many people were looking for antibodies, but wasting their time. Why wouldn't people want to be able to know which studies are using which antibody? So we were asked to sort this out as a way, essentially, to automate the process of finding antibodies.

While we were able to do many things, we completely failed at this task, because it was too hard. The information was not present in the papers for us to find, so there was nothing that we could do with our algorithms or any of our fancy math. We could not put into the papers something that was not there to begin with. This is what started us trying to think about ways that we could get journals to include additional data to help us with the text-mining task and to help others find antibodies. This initiative then grew into other research resources such as the mice, the flies, the cell lines, and, of course, antibodies and software tools—those were all the things that people were looking for; however, researchers could not find or readily identify them. We thought that if we can get people to put a little bit more of this information into their papers straight away, then our task would become much easier, thereby enabling easier access to information for readers.

Essentially, at the beginning, we failed to identify antibodies and wanted to address the problem. We did this by bringing together many journal editors who said that any possible solution would depend on having certain indexes. An index is a whole list of numbers similar to GenBank identifiers. Currently, in order to publish a genetic study, you need the GenBank identifier. If you want to ensure that you are talking about the same snip, you need an RS number for it.

Similarly, we need to have a way of identifying mice, flies, antibodies and cell lines. When we started this, we really needed to bring many databases under one place and have one uniform way of accessing all this different data with one uniform display. Then, with our experience in the Neuroscience Information Framework, we could very quickly

create a portal that would allow people to search across all this different data.

An Introduction to SciCrunch from Neuroscience Information Framework on YouTube

Kuntan: What steps do authors or researchers have to perform to generate RRIDs and cite them in their papers?

Anita: First, they have to go to *scicrunch.org/resources* and type in the catalog number of the antibody they are using in the search to get a list of results. For example, if they typed in MAP377, they would get several results. If the researcher used a particular antibody from Millipore then they would click the Cite This option, copy that particular piece of text, and paste it into their paper. By following this process, the antibody can now be easily identified.

Kuntan: In a 2016 *Nature* study, 80% of respondents attested to a scientific reproducibility crisis. What is your take on the report? Could you elaborate on how SciCrunch and RRIDs are going to resolve this problem around scientific reproducibility?

Anita: There have been many different reports in *Nature* and other journals describing some of the problems with reproducibility. In fact, the heads of the National Institutes of Health have authored some and have illuminated some of the problems and also actually addressed some of them. In 2016, they changed the reviewer guidelines for all grants. They analyzed several different places where we need to pay more attention. For instance, statistical challenges certainly need more attention, and certain efforts have been made to address these challenges.

Another part of the problem comes that the National Institutes of Health have identified are naming key biological resources such as antibodies, cell lines, other chemical reagents, and transgenic organisms. To fulfill this requirement, RRIDs are quite important; however, certain requirements need to be fulfilled for this purpose. First, we need to be able to identify particular organisms, cell lines, and antibodies. Second, we absolutely need to know if we are using the right reagent or if the cell line being used is not contaminated at the source. Many cell lines commonly used these days have drifted over the years and become something else. Committees of research scientists are looking at these issues. However, the problem is that these scientists are working all across the world and they do not always know about each other. An important thing about identifiers is that you can actually bring in additional information to a particular research object.

For example, I have a cell line from a particular biocenter that is known to be contaminated and this report exists in an excel file in Australia. However, with an identifier, such a report can be brought into that particular research object. Now, when a researcher tries to find a RRID, they might find that the cell line is contaminated and

might rethink their study's conclusions. In this manner, RRIDs should help authors bring in additional information into a study before it gets published.

Kuntan: Can RRIDs help identify mislabeled cells or questionable lab supplies thereby promoting reproducibility in science

Anita: Absolutely, on the one hand, RRIDs help identify and also answer the simple question about which resources were used. When enough people use these resources, they answer another question, i.e. "Who else used this particular resource?" However, when you get information about how RRIDs are actually functioning, and when there is enough data around them, then they actually provide other answers. The obtained information can be "attached" to individual cell lines or antibodies. When enough of these pieces of information are attached, or if the information is by a particular authority you know that you should probably trust it and act accordingly and re-examine whatever evidence you have with you in a proper way.

Kuntan: How are you currently increasing awareness amongst authors regarding RRIDs and using them in their manuscripts?

Anita: We work closely with several journals. We have recently brought online the journals published by Cell Press. If you are applying to publish in *Cell* or *Cell Systems*, or any of their other journals, on acceptance, the editors will ask you to add the RRIDs. Many authors already do this before they start the process; however, if you are working with a journal that directs authors to our website, you will have to do this. Even if the journal does not require it, you can still add RRIDs. We have had very few journals that feel this should not be done because they all recognize that this is a better way to identify research resources even if it is not in their Instructions to Authors. That way, you can track the reagent instead of trying to figure out where the company is located—which is useless these days since they are all on the Internet.

([To be Continued](#))

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

Cite this article

Enago Academy, Know More About SciCrunch and RRIDs: An Interview With Dr. Anita Bandrowski (Part 1). Enago Academy. 2017/09/04. https://www.enago.com/academy/know-more-

about-scicrunch-and-rrids-an-interview-with-dr-anita-bandrowski-part-1/