Research Data Management at the University of Adelaide from the perspective of a researcher and research support staff

This is a two-part case study from the University of Adelaide.

The first part is a case study on the research of Nathan Watson-Haigh, a Research Fellow in Bioinformatics at the School of Agriculture, Food and Wine.

The second part is an insight into the university's research data engagement across campus, as told by Mary O'Connor, eResearch Support Project Librarian in Learning and Research Services.

The University of Adelaide's Figshare portal can be found at **adelaide.figshare.com**

Using Figshare to store large datasets

Nathan Watson-Haigh

Research Fellow in Bioinformatics School of Agriculture, Food and Wine

Nathan has a biology background, but has transitioned into bioinformatics and is currently studying wheat genomics. The group of researchers he works with study the phenotypes of wheat, particularly with regard to salt, drought, and heat tolerance in order to selectively breed new varieties of wheat that are able to withstand drier, hotter, and more saline conditions.

The wheat genome is about five times larger than the human genome; therefore, working with this genome creates a lot of data. Bioplatforms Australia, in consultation with the Australian research community, has sequenced the genomes of 16 wheat accessions - the plant materials collected from a particular area - with the purpose of developing a resource for the community to use.

Before the creation of the reference genome, this resource could only be used by groups with enough bioinformatics and computing resources to handle the large datasets. Ahead of the release of the wheat reference genome, Nathan's team took the raw datasets, aligned then to the reference genome and performed variant calling, taking the burden off researchers needing to access the data. This data has been made available in Figshare.

In total, Nathan uploaded about 2.2 TB of data; there are 17 filesets within the collection of data, each fileset broken down per chromosome. Because there was hundreds of data files, Nathan didn't want to use the web interface to upload the data as it would be too time consuming, so he opted for the Application Programming Interface (API).

"One of the reasons why I chose Figshare was because of the ability to host files of this type and size," said Nathan. "For example, there's the Short Read Archive and similar repositories for depositing raw sequence data, but there doesn't seem to be a place to deposit processed data. That's where Figshare came into its own."

"The upside of dealing with the University of Adelaide library was that they were very responsive to the future storage requirements of my work as well as that of my colleagues. We now have a 5 TB storage limit, which is great because I'll be able to continue to use my account to upload future datasets."

Read more about Nathan and his research

The University of Adelaide's engagement strategy

Mary O'Connor
eResearch Support Project Librarian in Learning and Research Services

The University of Adelaide implemented Figshare in 2017 and has since taken a structured, targeted approach to engagement and advocacy, highlighting the benefits of uploading the data to Figshare on a case-by-case basis. As such, the Research Data (ReDa) Project Team and the Liaison Librarians have been targeting specific departments of researchers; for example, they've approached the computer science department about using Figshare to store their code. The department hasn't been quite as responsive as the team hoped, but the engineering department have been using Figshare for their code, having heard about it through word-of-mouth. They also ran a HDR Conference Poster upload competition, with quite a few posters uploaded in order to get a figshare KeepCup - a popular prize!

They have several possibilities lined up for engagement, as well, from architecture students uploading 3D models of their designs to fourth year engineering students uploading their final-year posters into an aggregated portal.

They've also assessed research center websites for content that would benefit from being on Figshare. They've discovered, in particular, videos that are hosted on YouTube which receive low numbers of views and have advertising attached to them, which they're hoping to upload to Figshare. The university also has a Research Tuesdays lecture series where researchers prepare lectures for the general public. These videos may soon be uploaded to Figshare, after consultation with the marketing department.

The team feel this targeted approach has been useful and is the way forward for engagement, but it's time-consuming. There are also opportunities to discuss research data management with researchers when they approach the Project Team or the Liaison Librarians about a separate research issue.

Once a researcher shows an interest in using Figshare, a member of the team will provide him or her with a demo of the system. They use a combination of in-person meetings, phone calls, emails, and department meetings to discuss research data management. Some research centers have had demos to the entire center rather than a one-to-one. Generic Figshare training hasn't worked as well as they had expected because most attendees were PhD students who were attending as part of compulsory research student skills training, rather than because they had a need for a research data management system.

They also promote use cases on their Figshare support page on the university's website. This allows showcasing to other researchers interested in using Figshare but unsure of the benefits or how to use it.

The next steps for the rollout include providing IT support for those wishing to use the API to upload data. They will also start to offer a holistic approach to outreach with other systems, like Symplectic Elements and Altmetric Explorer, offered by the university rather than focusing solely on Figshare. They will also consider focusing on a researcher-lead approach which will allow researchers to come and talk about the systems they're interested in and have administrative support providing training for those systems. Liaison librarians are also going to do more outreach with researchers aided by a research support brochure which outlines the different research support systems and what they do.

Read more about Mary