IFLA WLIC 2017 WROCŁAW

Submitted on: 25.05.2017

Adapting central RDM messages to discipline-specific needs at the Department of Chemistry, University of Cambridge

Clair Castle
Department of Chemistry, University of Cambridge, UK.
Email address: cmc32@cam.ac.uk

Abstract:
The University of Cambridge established its Office of Scholarly Communication (OSC) in 2015 (http://osc.cam.ac.uk/). Since then it has developed many services to support Research Data Management (RDM), including a central website, RDM training and support, and a data repository. It communicates with researchers and support staff including librarians and administrators across the University using a variety of methods. It has also had direct discussions with researchers, and carries out structured interviews and surveys, in a bottom-up approach to RDM which engages with researchers, combined with a top-down, policy-driven approach (Teperek, Higman and Kingsley, 2017). There is therefore a considerable amount of outreach into departments and faculties where research takes place. However, its resources are limited: it is not possible for it to deliver RDM training in every department or faculty in the University, for example.

Most departments and faculties have an embedded library service, which is discipline-specific. Librarians in the University are in a key position to be able to collaborate with the OSC and their own researchers in developing and implementing RDM services locally. This paper presents a case study of how centralised RDM services have been rolled out in the Department of Chemistry, thus adapting the central RDM messages to discipline-specific needs. I will begin by introducing the Department of Chemistry and the OSC and the centralised RDM services it offers to researchers. I will then cover my role as the Chemistry Department Librarian in delivering RDM training, and my involvement in the Data Champions programme (http://www.data.cam.ac.uk/intro-data-champions). I will describe the outcomes of this for the Department of Chemistry, and for the centralised service. Finally, I will outline future developments in the Department of Chemistry regarding RDM services offered in collaboration with the OSC.

Teperek, M., Higman, R., and Kingsley, D. (2017). Is Democracy the Right System? Collaborative Approaches to Building an Engaged RDM Community *[pre-print]*. https://doi.org/10.1101/103895

Keywords: Research Data Management, Chemistry RDM, Data Champions, training, chemistry data

Introduction

The Department of Chemistry at the University of Cambridge (http://www.ch.cam.ac.uk/) is currently placed third in the QS World University Rankings by Subject 2017 (https://www.topuniversities.com/university-rankings/university-subject-rankings/2017/chemistry). The department is one of the largest in the University, with around 200 post-doctoral researchers, 300 postgraduate students, 60 academic staff, and 100 administrative and technical support staff. The department's research structure is organised around five Research Interest Groups (RIGs): Biological, Materials, Physical, Synthetic, and Theory. The department has internal and external collaborations with other University departments, universities, and industry and covers an exceptionally wide range of chemical science from molecular biology to geophysics (http://www.ch.cam.ac.uk/about). Individual researchers generate large volumes of a wide variety of data types, from NMR (Nuclear Magnetic Resonance) spectra to computer code.

The Chemistry Library supports all the teaching and research that takes place in the department (http://www-library.ch.cam.ac.uk/). Since I joined the department as Librarian in October 2013 researchers have had to comply with increasingly strict funder mandates for open access and open data and they have needed considerable support with this.

Centralised RDM services at the University of Cambridge

The Office of Scholarly Communication (OSC) at the University of Cambridge was officially established in 2015, and reports to Cambridge University Library and the Research Office (http://osc.cam.ac.uk/). The OSC held direct discussions with researchers in an effort to better understand their Research Data Management (RDM) needs. Feedback from these discussions allowed the OSC to start developing the services researchers requested: a central website with information on RDM (http://www.data.cam.ac.uk/), RDM training and support, and a data repository (https://www.repository.cam.ac.uk/). The OSC's Research Data Facility offers individual advice on data management plans, support with depositing data in the institutional repository, and consultancy on all aspects of RDM to researchers across the University, as well as running events and workshops (Higman, Teperek and Kingsley, 2017).

Implementing centralised RDM services at the Department of Chemistry

The initial reaction from researchers at earlier OSC information sessions on RDM certainly was that this was "yet another thing to remember to do on top of everything else". There had been poor communication from funders about their requirements and the reasoning behind them.

I could see early on that it was extremely important to collaborate with the OSC's Research Data Facility to get its messages about RDM across to our researchers. Neither of us had the resources in terms of staffing or funding to do it alone.

Apart from simply advertising OSC events, and arranging briefings on the RDM services available, I decided to get much more involved and to provide discipline-specific support for researchers at my department.

RDM training for graduate students

The Department of Chemistry agreed to pilot RDM training with our PhD students. The rationale behind training this particular group was that:

- It would filter knowledge upwards to their PIs (Principal Investigators, who are in charge of research Groups and are responsible for applying for research grants);
- PhD students often act as corresponding authors on articles and their PIs ask them to make them OA. They therefore needed support with this;
- PhD students are more open to sharing data;
- It would get PhD students into the habit of thinking about how they manage their research data, as potential future post-doctoral researchers and PIs.

Research Data Facility staff designed a Powerpoint presentation that could be used as a template for customisation by departments across the University (Higman and Teperek, 2017). It comprises four sections: backup and exchange strategies, how to organise your data, data sharing, and writing data management plans. Practical activities are incorporated throughout, which makes it highly interactive, and easier to keep participants' attention.

I proposed including the RDM training in our Graduate Education Programme to our Head of Graduate Education. This was approved, and an initial trial session was delivered by myself and members of the Research Data Facility. It was well attended and there was positive feedback from participants. The Head of Graduate Education made a successful pitch to the Graduate Education Committee that sessions should take place regularly and should be made compulsory for all first year PhD students to attend.

As a typical example, in 2016-17 I spent 8 hours training 71 students in four two-hour long sessions. Smaller classes (~24 people) are most effective for this type of course so multiple sessions are required throughout the academic year in order to cover the whole annual cohort of around 75 students. I continue to customise the content for chemists and act upon the feedback received after each session.

The Data Champions programme

The Research Data Facility's training programme is heavily subscribed but currently lacks stable funding to employ enough people to meet demand. It is also difficult for a central support service to develop the expertise needed to provide in-depth advice in every discipline, given the range of data and research methods that these entail (Higman, Teperek and Kingsley, 2017). In 2016 myself and other stakeholders from across the University organised a call for Data Champions to researchers, students, librarians, IT managers, data managers, other members of staff and anyone with a keen interest in RDM (Higman and Teperek, dataset). Three members of the Department of Chemistry successfully applied, and, as a member of the RDM Working Group that oversees the Data Champions programme (https://osf.io/wbt2x/), I have been mentoring them on various RDM activities (http://www.data.cam.ac.uk/intro-data-champions).

I asked the Data Champions what issues researchers keep asking about and were not getting support with, and how these gaps could be covered. As a result, we have achieved the following so far:

- 'Chemistry Data FAQs' have been posted on the library's Open Data website (http://www-library.ch.cam.ac.uk/open-data-faqs-chemists). This addresses the types of data chemists should share when publishing;

- Slides on Dropbox and Github, and inspirational talks on open science, have been incorporated in the RDM training sessions;
- The first list of instructions for converting the data generated through experiments using techniques such as NMR spectroscopy, mass spectrometry, etc. into open data formats that can be shared easily has been posted on the library's Open Data website (http://www-library.ch.cam.ac.uk/convert-your-files-open-data-format). This shares best practice and avoids researchers having to reinvent the wheel each time they need to share data;
- The first of hopefully many training sessions on data management plans (including RDM best practice) for post-doctoral researchers has been delivered.

Outcomes for the Department of Chemistry
Students consistently and overwhelmingly state in feedback that the RDM training should be compulsory for all new PhD students.

Having communicated RDM activities run centrally and carried out some RDM activities internally, the library is now associated with all things 'open research' and is the main point of contact for our researchers. I can also refer them to the OSC for further support.

Membership of the RDM Working Group allows me to have a direct influence on RDM policy and activities within the University, and ensures that the RDM interests and needs of our researchers are represented.

It is helpful for researchers to be able to call upon the expertise of their peers in the form of Data Champions rather than a librarian who may know a lot about RDM principles but not necessarily about particular research techniques and the types of data involved.

Outcomes for the centralised RDM service
The RDM Working Group helps inform central policy making decisions with the input of people interested and active in RDM from across the University.

The Data Champions can help deliver RDM training and other RDM related activities across the University, in discipline specific areas that would be otherwise hard to reach.

Future RDM developments at the Department of Chemistry
There is still much that the Data Champions would like to do:

- Case studies of how researchers in the department incorporate RDM activities (including open access and open data) in their day-to-day workflows. This could perhaps be in the form of 5-10 minute presentations, in person or online through videos;
- Make and promote a list of experts in the department in various RDM aspects;
- Drop-in sessions where researchers can bring their own devices to quickly create ORCID accounts (https://orcid.org/), and learn about synching their ORCID with other accounts such as Symplectic Elements (http://symplectic.co.uk/products/elements/);

- More strongly promote links to information already provided on compliance with open access and open data mandates to improve the current compliance rate.

I will continue to work with them to find ways that they can help communicate RDM best practice.

Conclusion

The OSC has achieved such a lot in a relatively short time. It has been easy to collaborate closely with them and to share ideas. I know that chemists have attended OSC events and that this has benefitted them. We now need to build upon the Data Champion activities within departments.

There are of course always significant barriers to researchers taking on central and local RDM messages and implementing them, including researcher apathy and the lack of time available to researchers. However, if we all continue to work together to develop and implement RDM solutions that help make researchers' lives easier this will generate further interest in other RDM activities. I look forward to supporting the OSC with this in future.

Acknowledgments

I thank Marta Teperek, Data Facility Manager at the OSC, who has consistently encouraged and supported me in all my RDM activities and who gave me the confidence to write this paper and present it at the IFLA Satellite Conference.

I thank Deborah Longbottom, Head of Graduate Education at the Department of Chemistry, for her enthusiastic support for the RDM training that has been so successful.

References

Higman, R., Teperek, M., and Kingsley, D. (2017). Creating a community of data champions [pre-print]. https://doi.org/10.1101/104661

Higman, Rosie, & Teperek, Marta. (2017, January). Template Research Data Management workshop for STEM researchers. Zenodo. http://doi.org/10.5281/zenodo.239090

Teperek, M., Higman, R., and Kingsley, D. (2017). *Is Democracy the Right System? Collaborative Approaches to Building an Engaged RDM Community* [pre-print]. https://doi.org/10.1101/103895

Higman, R., & Teperek, M. Research data supporting "Creating a Community of Data Champions" [Dataset]. https://doi.org/10.17863/CAM.7417