Clair Castle*

Getting the Central RDM Message Across: A Case Study of Central versus Discipline-Specific Research Data Services (RDS) at the University of Cambridge

https://doi.org/10.1515/libri-2018-0064
Received May 28, 2018; revised August 31, 2018; accepted October 24, 2018

Abstract: RDS are usually cross-disciplinary, centralised services, which are increasingly provided at a university by the academic library and in collaboration with other RDM stakeholders, such as the Research Office. At research-intensive universities, research data is generated in a wide range of disciplines and sub-disciplines. This paper will discuss how providing discipline-specific RDM support is approached by such universities and academic libraries, and the advantages and disadvantages of these central and discipline-specific approaches.

A descriptive case study on the author's experiences of collaborating with a central RDS at the University of Cambridge, as a subject librarian embedded in an academic department, is a major component of this paper. The case study describes how centralised RDM services offered by the Office of Scholarly Communication (OSC) have been adapted to meet discipline-specific needs in the Department of Chemistry. It will introduce the department and the OSC, and describe the author's role in delivering RDM training, as well as the Data Champions programme, and their membership of the RDM Project Group. It will describe the outcomes of this collaboration for the Department of Chemistry, and for the centralised service.

Centralised and discipline-specific approaches to RDS provision have their own advantages and disadvantages. Supporting the discipline-specific RDM needs of researchers is proving particularly challenging for universities to address sustainably: it requires adequate financial resources and staff skilled (or re-skilled) in RDM. A mixed approach is the most desirable, cost-effective way of providing RDS, but this still has constraints.

***Corresponding author: Clair Castle,** Department of Chemistry Library, University of Cambridge Department of Chemistry, Cambridge, UK, E-mail: cmc32@cam.ac.uk
https://orcid.org/0000-0002-6897-582X

Keywords: research data management, research data services, chemistry data, discipline-specific research data services, subject librarians

Introduction

Centralised RDS are mainly designed to support the RDM needs of researchers at a generic level in order to consistently get central messages across to the highest number of people, for example in the area of RDM policy and procedures. Research data itself is generated, especially at research-intensive universities, across a wide range of disciplines and sub-disciplines. How far are these disciplinary differences taken into account when planning RDS? What are the advantages and disadvantages of discipline-specific versus centralised RDS? How sustainable is either approach? A descriptive case study on the author's experiences of collaborating with a central RDS at the University of Cambridge, as a subject librarian embedded in an academic department, is a major component of this paper.

Library Collaboration with Other RDM Stakeholders

Morais and Borrell-Damian (2018, 30) describe RDS in the Open Access 2016–17 EUA (European University Association) Survey Results as "A. Support services provided by library or other specialised staff, typically on a one-to-one basis. B. A dedicated office or service supporting researchers is available at some institutions. C. Training for researchers and support staff. D. Training for graduate students, including doctoral candidates. E. Information events focusing on research data management and/or open access to research data. F. Institutional website with information about research data management and open access to research data, blogs, newsletters. G. Specific financial support for researchers, including doctoral candidates, to attend

events on open access to research data and/or for open access publications." Some libraries also provide 'technical RDS' e. g. "providing technical support for data repositories, preparing data sets for a repository, deaccessioning or deselecting data sets from a repository, or creating metadata for data sets" (Tenopir, Birch, and Allard 2012, 7).

Providing RDS usually involves collaboration with other university RDM stakeholders. Cox et al.'s (2017, 2186) international survey of RDS in libraries in Australia, Canada, Germany, Ireland, the Netherlands, New Zealand, and the UK found that "RDM policy is a multistakeholder process with a wide range of participants (library, IT services, research office, legal office, and academic contributors)" and that "leadership and initiative in RDM policy development came most commonly from the library or research office or a close partnership between the two".

For example, at the University of Manchester the library has established a new "'Research Services Team' which acts as a single point of contact for the service but works in liaison with the other key stakeholders as part of a wider network" (Williamson and Parsons 2013, 12). There is now a "single point of contact for RDM enquiries", a "website on RDM", and "it has delivered training events, and created templates and examples of good data sharing statements aligned with what the funding bodies are demanding" (Williamson and Parsons 2013, 12).

Centralised and shared services can offer broad and consistent RDS across a university, taking advantage of the skills and knowledge available from the various stakeholders involved. For example, at the University of California, Berkeley, where research is "highly distributed", the UC Berkeley Library and Research Information Technologies (Research IT) joined forces to develop an RDM program "which will bring together the campus-wide systems and technical knowledge of Research IT with the research support and preservation expertise of the Library" (Wittenberg and Elings 2017, 90).

The main advantage of this collaborative approach is that of sustainability in RDS provision. Centralised, shared services are more cost-effective in terms of staffing and technical service infrastructure and, as such, are typically cross-disciplinary. "Most campus libraries today are structured by discipline to support academic departments" but "RDM support requires activities that cut across this departmentally aligned organizational structure" (Flores et al. 2015, 92).

Central versus Discipline-Specific Repositories

One example of 'technical RDS' is the provision of an institutional repository. Although some university libraries do operate their own institutional repository, which is generally multi-disciplinary and preserves the data underpinning its own authors' research, subject-specific repositories also exist. Researchers may prefer or be required by their funder to deposit their data in these: "Many subject areas are covered by well-developed data management facilities run by national or international data centres, reflecting disciplinary differences in the academic culture around the deposit and re-use of datasets" (Lewis 2010, 11). Some examples are the EBI (European Bioinformatics Institute), NERC (Natural Environment Research Council), UK Data Archive, the UK Data Service, and CCDC (Cambridge Crystallographic Data Centre) (Lewis 2010, 11). A useful resource for identifying subject-specific repositories is re3data.org (re3data.org 2018). Encouraging researchers to use either institutional or external subject-specific repositories, or supporting researchers' use of them, is one sustainable way of providing both central and discipline-specific RDS. At the University of Oxford for example, if

> there are national or international subject-specific repositories for the data in question, it will be far more sensible to use these and make use of the specialized curation skills of their repository staff than to use the more generic curation services associated with the institutional archive (Wilson and Jeffreys 2013, 238).

Discipline-Specific Aspects of Planning RDS

The examples mentioned so far are mainly of centralised RDS. However, it has been acknowledged that disciplinary differences should be considered when planning RDS. After all, the LIBER working group on e-sciences made ten recommendations for libraries to get started with research data management in 2012. No. 9. was "Subject-specific RDM support" (Christensen-Dalsgaard et al. 2012, 1).

At Emory University Libraries, different disciplines "vary widely in their research funding, technical infrastructures, collaboration networks, source materials, subject populations, methodologies, ethical considerations and types of research outputs. Therefore, to be most effective, data curation requires discipline-specific

approaches" (Akers and Doty 2013, 14). The importance of considering "both the similarities and dissimilarities among disciplines" is highlighted, as this "will help guide academic librarians in developing a range of data management-related services that can be tailored to the unique needs of different researchers, thereby resulting in more effective and comprehensive approaches to data curation" (Akers and Doty 2013, 17).

Pinfield, Cox, and Smith (2014, 25) have constructed "a tentative model of an RDM programme within an academic institution" which is "intended to address the 'who?', 'what?', 'why?', and 'how?' of RDM, particularly in relation to the library's involvement" (Pinfield, Cox, and Smith 2014, 22). Different disciplinary approaches are represented in a separate layer in the model, that "any institution-wide model needs to take into account" (Pinfield, Cox, and Smith 2014, 23). They admit that one "under-developed aspect of the model is the treatment of the way that any initiative needs to be adaptive to the diverse disciplinary cultures of research communities". Some disciplines already share data openly whereas some do not even use the term "data" and "such issues are incorporated in but are not central to" this model (Pinfield, Cox, and Smith 2014, 25).

The University of Nottingham has created "pragmatic models" for institutional RDS which "provides the outline of three levels of service (minimal, mediated, consultancy)" (Williamson and Parsons 2013, 15). Core RDS are categorised as Level 1 – Minimal, Level 2 – Mediated, Level 3 – Consultancy. Four activities mention subject-specific elements, as Level 3 services only: "Data Management Plans (Tailored approach, subject-specific advice & training)"; "Active data management and storage (Tailored approach & subject-specific advice to meet funders' requirements)"; "Data archiving and long-term preservation (Tailored approach & subject-specific advice to meet funders' requirements)"; "Website (Consultancy, subject-specific and embedding skills)" (Williamson and Parsons 2013, 15–16).

These examples indicate perhaps that, when planning RDS, libraries take discipline-specific services into account at a later stage of their development. Nevertheless, they do see discipline-specific RDS as a way of providing a more comprehensive service overall. Tenopir et al. (2017, 38) raise the need for more research on discipline-specific RDS provision and how [European] libraries will "customize services to meet the needs of different subject disciplines". It is not cost-effective for individual departments within a university to duplicate effort in developing its own RDS, but there is a demand for discipline-specific services – these can be central ones customised for a particular subject.

Centralised and Discipline-Specific RDM Training

One centralised area of RDS that can be successfully adapted to suit disciplinary RDM needs is training and advocacy. Cox and Pinfield (2013, 312) undertook a survey of UK universities in 2012 to understand the ways in which libraries are involved in RDM. In the area of advocacy, although "28 respondents (37 % of the 76 answering this question) stated that in their view RDM was 'best approached through institutional advocacy and support' and 12 (16 %) disagreed, a large number of respondents (36 or 47 %) were reluctant to choose between institutional or subject-community approaches." In addition, several respondents commented "subject-based support is likely to be variable across different disciplines with institutional approaches filling gaps and providing consistency for reasons such as regulatory compliance" (Cox and Pinfield 2013, 312). Consistency in central RDM messages, especially in the area of central RDM policy, should be maintained but should also be adapted to suit different disciplines.

At the University of Oxford, for example, "there is a distinction between training and services that can be most effectively provided at the institutional level, and those which need to be provided at the divisional or departmental level in order to be relevant and applicable to researchers" (Wilson and Jeffreys 2013, 235). The advantage of discipline-specific training and advocacy is that it is often more effective because it is more relevant to its recipients. The customised approach at the University of Oxford is highly appreciated:

> it is appropriate that research data management training be undertaken in a large part at disciplinary level, as experience suggests that researchers relate better to such training when they have examples which they can relate to, rather than when data management is dealt with at an abstract high level (Wilson and Jeffreys 2013, 244).

At the University of Bath "there is an agreement that generic training with discipline-specific examples is the most effective approach with resourcing constraints in mind" (Guy 2013a).

At the University of Bristol, whilst

> many of the areas covered in workshops were generic and relevant to all disciplines, both researchers and PGRs still seem much more motivated to attend a session if they believe it is faculty or discipline-specific, and are grateful for any discipline-specific examples which can be provided throughout workshops. (Hiom et al. 2015, 479).

However, due to the success of the initial, broad training, Bristol is facing increased demand for "next level" discipline-specific training, which it cannot sustain for every discipline it supports (Hiom et al. 2015, 489). It is considering handing over responsibility for "truly discipline-specific research data management training to the academic leads of taught Postgraduate courses" and "grouping research activity in a way which is of greater use when planning training", e. g. by identifying "research themes or methodologies and designing future training courses accordingly" to "allow the service to achieve the goal of offering both a broad and deep training provision" (Hiom et al. 2015, 490). "Themes" identified include GIS (Geographic Information Systems), sensitive data, and imaging data. (Hiom et al. 2015, 490). This, too, raises the issue of sustainability of discipline-specific RDM training for subject librarians, whether they are part of a larger 'campus' library or embedded within an academic department.

Bridging the RDM Skills Gap

Training and other RDS will often be provided by librarians with subject expertise. The literature refers to 'subject' and 'liaison' librarians. These terms can mean the same thing in theory but can be different in practice. Subject librarians can be physically based in a central library service where they are responsible for a particular subject or range of subjects. Where there is a devolved library service, as with the University of Cambridge, subject librarians are often 'embedded' in a particular department, such as Chemistry. Both roles can be described as part of the traditional liaison model because they involve liaison with faculty and students to support discipline-specific teaching and research needs. A Research Data Service (RDS) can struggle to support different disciplines if it does not incorporate or have access to staff with essential discipline-specific knowledge and skills.

The RDM skills gap amongst librarians has been widely discussed in the literature (Auckland 2012; Bresnahan and Johnson 2013; Corrall 2010; Cox et al. 2017; Cox and Pinfield 2013; Gabridge 2009; Guy 2013a, 2013b; Kennan 2016; Koltay 2016; Pinfield, Cox, and Rutter 2017; Pinfield, Cox, and Smith 2014; Swan and Brown 2008; Tenopir et al. 2014, 2017; Williamson 2013). Auckland (2012, 42) identified three skills or knowledge areas related to RDM that subject librarians may need to have, which are included in the "nine areas where over 50 % of the respondents with Subject Librarian responsibilities indicated that they have limited or no skills or knowledge" and which were also "deemed to be of increasing importance in the future". These are: "Knowledge to advise on data management and curation (48 % essential in 2–5 years, 16 % now)", "Knowledge to advise on potential data manipulation tools used in the discipline/subject (34 % essential in 2–5 years, 7 % now)", "Knowledge to advise on data mining (33 % essential in 2–5 years, 7 % now)" (Auckland 2012, 43). The re-skilling of librarians in RDM is therefore essential for them to be able to support researchers' data needs.

With regard to science and engineering librarians at MIT, Gabridge (2009) describes the "last mile" problem:

> "Studying the research needs of individual researchers and of the institution as a whole is a major component of the work of library subject liaisons. As such, these librarians are well positioned, and will be essential in building the "last mile" of research data cyberinfrastructure – the part of the network that will provide connections between the systems and the researchers, and ultimately, to new users of the data".

Subject librarians occupy a pivotal role in RDS provision across a university.

Tenopir et al.'s (2014, 89) study into RDS practices in U.S. and Canadian academic research libraries finds that as librarians increasingly get involved with research data, many "feel they have the subject knowledge necessary to help their constituents with research data services, but need the opportunity to take advantage of continuing education. Whether consultative or hands-on services, librarians need opportunities to learn more about these services either on their own campus or through attendance at workshops and professional conferences."Guy's (2013b) case study on the RDMRose Jisc-funded project "to produce taught and continuing professional development (CPD) learning materials in RDM tailored for information professionals" suggested

> Librarians do not know enough about what researchers do, and so researchers often do not feel that librarians can support them in their work. Much of the RDMRose training focuses on librarians understanding the diverse perspectives of researchers in different disciplines, getting out there and learning more about the research process and where the various roles that librarians could play are really needed. 'Getting out there' could take the form of shadowing, networking, and working on real data sets (Guy 2013b).

RDM skills can also be viewed as a career development opportunity in a new and exciting area.

New and existing librarians can develop their own skills in RDM by taking advantage of in-house programmes. One example is the Research Support Ambassadors Programme at the University of Cambridge, an initiative which skills staff moving into research support roles in areas such as RDM and OA (Sewell and Kingsley 2017). Other methods include attendance at external meetings, seminars, workshops, webinars, and conferences. Examples of sources of online continuing education in RDM for librarians include:

- NNLM RD3 (National Network of Libraries and Medicine) website (NNLM n.d.a.) lists courses and workshops in the form of online courses and tutorials, listservs. Includes guidance in subject areas, e. g. Physical Sciences, Chemistry.
- NNLM Library Roles (NNLM n.d.b.) website. Recommends introductory primers, books, general readings, case studies, competencies for data librarians.
- Data Curation Profiles (Data Curation Profiles n.d.a.). Website resource for those who want information about the specific data generated and used in research areas and sub-disciplines that may be published, shared, and preserved for re-use. Data Curation Profile Toolkit can be used to interview researchers about their RDM practices.
- ACRL (Association of College & Research Libraries) Scholarly Communication Toolkit (ACRL 2017). Can be used to support advocacy efforts designed to transform the scholarly communication landscape.
- DCC (Digital Curation Centre) Disciplinary RDM training (DCC 2018). Created a body of discipline-focused postgraduate training units which can be reused by others.
- FOSTER Open Science portal (FOSTER n.d.). An e-learning platform that brings together the best training resources addressed to those who need to know more about Open Science, or need to develop strategies and skills for implementing Open Science practices in their daily workflows.

Until RDM is included in library and information science courses, both for new and mid-career library staff, this could help fill the skills gap (Cox et al. 2017, 2194).

Cox et al. (2017, 2191) found the most common strategies libraries have used to "develop staff capacity and capability for RDS" were to "reassign existing staff (25%) or to recruit and reassign staff (25%). Other common responses were that they had already recruited new staff (12%) or planned to reassign existing staff (12%)". This is another way of addressing the RDM skills gap.

Discipline-specific RDM expertise can be achieved by a centralised service "through training and development of the existing workforce, and the recruitment of new staff with the necessary skills and knowledge" (Auckland 2012, 3). Subject librarians can acquire generic RDM skills through continuing education activities and find ways of applying them within their particular discipline.

Case Study of Centralised and Discipline-Specific RDM Support at the University of Cambridge

The University of Cambridge

The University of Cambridge (University of Cambridge 2018a) was founded in 1209. It was ranked among the top five institutions for research in the UK in the latest 2014 Research Excellence Framework (REF) exercise, which determines the amount of funding it receives from the government for research (Research Excellence Framework 2017). It is a member of the Russell Group of universities, an association of 20 major research-intensive universities of the UK (Russell Group n.d.).

The University of Cambridge is a confederation of six Schools and around 100 faculties, departments and other institutions, and 31 colleges. The colleges are governed by their own statutes and regulations, but form an integral part of the university, which currently has around 12,000 undergraduate and 7,200 postgraduate students (national and international). There are around 4,000 contract research staff (post-doctoral researchers) and around 10,000 administrative, technical, and manual staff who support the researchers and students (University of Cambridge 2018b).

The Department of Chemistry

The department (University of Cambridge 2018c) belongs to the School of Physical Sciences. It is currently placed second in the *QS World University Rankings by Subject 2018* (TopUniversities 2018) and is one of the largest departments in the university, with around 200 post-doctoral researchers, 300 postgraduate students, 60

academic staff, and 100 administrative and technical support staff. The department produces ~ 500 publications per year, and many more data sets.

The department's research structure is organised around five Research Interest Groups (RIGs):

- Biological RIG. Understanding biological systems at a molecular level, with relevance to health and disease, and drug discovery;
- Materials RIG. Developing new materials and surfaces e. g. in solar cells, batteries, and computer sensors;
- Physical RIG. Investigating the physical properties of molecules, in relation to climate change and pollution. Involves computer modelling of the atmosphere;
- Synthetic RIG. Making chemical compounds for use in industrial processes e. g. in pharmaceuticals, plastics;
- Theory RIG. Understanding the physics of chemical processes e. g. computer modelling of molecules to predict their behaviour when interacting with each other.

A Principal Investigator (PI, in charge of research Groups and responsible for applying for research grants) is generally a member of at least two RIGs, which is intended to encourage inter-disciplinary research within chemistry.

Individual researchers generate large volumes of a wide variety of data types, ranging from spectra (the raw data generated through performing experiments using Nuclear Magnetic Resonance, NMR machines), to computer code created for text and data mining of the scientific literature as part of the drug discovery process. File sizes vary but some, for example images from electron microscopes, can be terabytes each in size. Research data supporting papers published by members of the department can be viewed on the university repository, Apollo (University of Cambridge 2018d).

Libraries at the University of Cambridge

Cambridge University Library (University of Cambridge 2018e) is one of six in the UK and Ireland which receive a copy of everything published in the UK, by law (ALDL n.d.). The University Library alone is home to 8 million books, journals, and other items.

There are over 100 libraries in the university; in general, one at each department or faculty and college. Members of the university can use most of the libraries, though college libraries are typically restricted to members of that college. Faculty or department libraries are effectively subject libraries which are 'embedded' within a faculty or department.

The Department of Chemistry Library is therefore discipline-specific and supports all the teaching and research that takes place in the department (University of Cambridge 2018f). It does this in the usual way; by providing and managing space for users, providing relevant collections (printed and electronic), as well as providing services such as training, enquiries, document supply etc. 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, as well as RDM, and they have needed considerable support with this.

Centralised RDS at the University of Cambridge

The Office of Scholarly Communication was established in 2015 (University of Cambridge 2018g) and reports to Cambridge University Library. The OSC had initially organised "a series of information sessions, to which we invited researchers, research staff and students. The main message delivered at these sessions was that research data needs to be shared due to funders' requirements" (Teperek, Higman, and Kingsley 2017, 88). These presentations were perceived by the researcher community as "yet another 'checkbox' activity, dictated by funders and the central University administration". Additionally, "our initial approach was not accompanied by new resources or new services developed" to support the new open access (OA) requirements (Teperek, Higman, and Kingsley 2017, 88). In an effort to better understand the research community's RDM needs, the OSC had direct discussions with researchers, held structured interviews with and surveyed them, and also held open door meetings with funders. This bottom-up approach to RDM of engaging with researchers was combined with a top-down, policy-driven approach (Teperek, Higman, and Kingsley 2017, 91).

The feedback obtained allowed the OSC to "start developing services requested by the research community: "a central website with information on RDM" (University of Cambridge 2018h), "RDM training and support, and a data repository" (Teperek, Higman, and Kingsley 2017, 90). The OSC's Research Data Facility offers "individual advice on data management plans,

support to deposit data in the institutional repository, and consultancy on all aspects of RDM" to researchers across the university (Higman, Teperek, and Kingsley 2017, 97). The institutional repository, Apollo (University of Cambridge 2018i), is mainly intended to preserve university research outputs such as publications and the data that supports them. As well as events and workshops, communication methods with researchers include a Twitter account @CamOpenData, a newsletter, "e-mails, and traditional post" (Higman, Teperek, and Kingsley 2017, 2). There is therefore a considerable amount of outreach into departments and faculties where research takes place. However, the OSC's resources are limited: for example it is not possible for it to deliver RDM training in every department or faculty in the university. This is one disadvantage of centralised RDS.

The OSC broadly targets RDM training and advocacy towards supporting STEM (Science, Technology, Engineering, and Medicine) and HASS (Humanities, Arts, and Social Sciences) researchers. Developing more deeply discipline-specific RDS would require much more staffing capacity if it were to be carried out centrally. The structure of many departmental/faculty libraries is still highly devolved, so the OSC cannot control how RDM messages are communicated and what kind of training is carried out, or to what extent.

Implementing Centralised RDS at the Department of Chemistry

In 2016 I arranged for members of the OSC to deliver an information session to faculty. The reaction from researchers was that this was "yet another thing to remember to do on top of everything else". The EPSRC, one of the department's main funders, had just made their open access mandates stricter and were monitoring compliance closely, as many research-intensive universities have experienced as a major driver for RDM support (EPSRC 2014). Researchers were frustrated: there was a perceived lack of communication from funders about what their requirements were, and the reasoning behind them. Researchers also needed reassurance that the support infrastructure to help them meet the requirements was going to be put in place.

This highlighted my own RDM skills gap. I qualified as a librarian over 20 years ago, when research data management was not on the horizon. I learnt about RDM by attending external conferences, and anything in terms of meetings and workshops etc. that was being offered by the OSC. Having observed researchers' negative or apathetic attitudes to RDM activities I was determined to work with colleagues in the department to take steps to change them.

I designed a new library website which includes an area on open research that would act as a constantly developing source of information on discipline-specific open access publishing, open data, and RDM. Although this is a discipline-specific website supporting RDM, I ensure that I link back to central services offered by other RDM stakeholders in order to uphold consistency in central RDM policy.

The change in the OSC's approach to educating researchers about RDM had been successful, and I could see that it was extremely important to collaborate with its 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 more involved in the provision of discipline-specific RDM support for researchers in my department.

RDM Training for Graduate Students

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

- It would filter knowledge upwards to their PIs;
- 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, understanding that this is necessary if they want to be able to use others' research;
- It would get PhD students into the habit of thinking about how they manage their research data in general, and for specific projects, and to learn how to write a DMP (Data Management Plan), 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 2017b). 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 customised it with chemistry related content.

My proposal to our Head of Graduate Education to include RDM training in our Graduate Education Programme 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. As a result, the Head of Graduate Education made a successful pitch to the Graduate Education Committee that sessions should take place regularly and be made compulsory for all first-year graduate students to attend.

As a typical example, in the academic year 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, enabling me to learn more about the types of data students generate and need to manage throughout their PhD. Teaching RDM, along with attendance at RDM related events, reflects the 'learning on the job' approach that many librarians have needed to take in order to equip themselves with the skills and knowledge they need to support RDM.

Membership of the RDM Project Group

In 2016 the OSC established the RDM Project Group, a group of individuals who volunteered to take part in development and overseeing of RDM services at the University of Cambridge (Open Science Framework, 2018). Its terms of reference are:

- Provide a multi-stakeholder perspective on various aspects of research data management, including, but not limited to: outreach and support, standards and integration, preservation, solutions for personal/sensitive data;
- Recommend and develop solutions to address gaps in RDM provisions at the University of Cambridge;
- Provide reports and recommendations on relevant items to the university's Open Access Project Board (University of Cambridge 2018j).

I am one of the researchers, librarians, and research support staff who volunteered to join. Working groups were subsequently created: Standards, Preservation, and Systems Integration, Outreach and Support (of which I am currently Co-Chair), Sensitive Data, and Data Privacy.

We meet bimonthly, with the Working groups meeting more frequently and reporting back on the progress of their work at Project Group meetings, to ensure alignment and coordination of efforts of individual Working groups. The Outreach and Support Working Group is concerned with:

- Cultural change: incentivising and rewarding researchers, and encouraging reproducibility, throughout the research lifecycle;
- Reaching researchers: targeting disciplines, reaching out to where they are, communication. This includes allocating Data Champions to each department;
- Best practice and implementation: support and outreach activities. Resources should be tailored to researchers, and involve using case studies;
- Coordination and collaboration with external bodies.

Discipline-specific RDM support is not actually mentioned in the terms of reference of either the RDM Project Group or the Outreach and Support Working Group. Many of the members of these groups come from discipline-specific libraries across the university. Discipline-specific needs are partially represented, but by no means comprehensively, and there is no particular focus on this aspect of RDS provision.

Involvement in 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", and it is "also difficult for a central support service to develop the expertise needed to provide in-depth advice in every discipline, from Architecture to Zoology, given the range of data and research methods that these entail" (Higman, Teperek, and Kingsley 2017a, 97). In 2016 the Outreach and Support Working group organised a call for Data Champions to researchers, students, librarians, IT managers, data managers, other members of staff, and anyone else with a keen interest in RDM (Higman and Teperek 2017a dataset). The aim of the programme "is that other researchers engage with the Champions as peers in their discipline"

(Teperek, Higman, and Kingsley 2017, 98). The programme "not only solves the problem of making discipline-specific training on RDM sustainable, but also helps maintain the engagement within the research community by recognising and rewarding those championing research data management" (Teperek, Higman, and Kingsley 2017, 91).

Three members of the Department of Chemistry successfully applied and I have been mentoring them on various RDM activities. I carried out an initial training needs analysis with them to identify what RDM activities are needed and what could be achieved. This identified the following:

- How to make data generated through experimental instruments into open file formats for future sharing and preservation;
- Backups. This is addressed in the RDM training I do for all new graduate students;
- How to make papers and data open to satisfy funder and publisher mandates;
- What data should be shared when publishing;
- Case studies of how researchers in the department incorporate open research and RDM activities in their day-to-day workflows. Perhaps in the form of 5–10 minute presentations, in person or even online via video;
- A list of experts in the department on various aspects of RDM;
- Drop-in sessions where researchers can bring their own devices to quickly create ORCIDs, unique personal identifiers for researchers (ORCID n.d.), and learn about synching their ORCID with other platforms such as Symplectic Elements that researchers use to manage their research outputs at the university (Symplectic n.d.);
- Provide more information about using social media platforms to share data;
- More strongly promoting links to information already provided on compliance with open access and open data mandates in an attempt to improve the current compliance rate.

We then considered ways in which these issues could be addressed and have achieved the following so far:

- 'Chemistry Data FAQs' have been posted on the library's open data website (University of Cambridge 2018k);
- Slides on using cloud storage service Dropbox within a research group, and inspirational talks on open science, have been incorporated into RDM training sessions for postgraduate students;
- An 'Introduction to GitHub for chemists' session has been delivered via GitHub (GitHub 2018a);
- The first protocol for converting proprietary data generated through experiments into open data formats that can be shared easily has been posted on the library's open data website (University of Cambridge 2018l). A survey was distributed to all researchers in the department to find out what knowledge they have in this area in the hope that they can share best practice;
- An 'Introduction to ORCID' session was delivered as part of the department Careers Programme 2017–18 on GitHub (GitHub 2018b).

There was a new call for Data Champions early in 2018 and as a result 20 new Champions were recruited across the university, taking the total number to around 50. There are two new Data Champions in the Department of Chemistry, taking the total number to four. Their support is invaluable for me in terms of reaching a larger number of researchers in the department, and in identifying their unmet RDM needs.

Outcomes for the Department of Chemistry

Students consistently and overwhelmingly state in feedback that the RDM training should be compulsory for all new graduate students. The answers to the question "Please state one thing you have learnt at this workshop" in feedback forms demonstrate that graduate students have really engaged in all of the various aspects of RDM and have taken central RDM messages on board. I have seen that first-year graduate students are inspired by the RDM advocacy carried out by fellow students in the department as Data Champions; they respond best to talks by their peers. However, this level of training is just about sustainable for me alone to deliver within the department. There are many sub-disciplines within chemistry with varying RDM needs that could be supported. The Data Champions within the department may be able to support me but it might be desirable to ask post-docs, PIs, and academics from the department to contribute, as the University of Bristol is considering doing.

Having communicated RDM activities run centrally and carried out some 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 Project 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 researchers from my department 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 RDS

The RDM Project Group helps inform central policymaking decisions with the input of people interested and active in RDM from across the university, although it does not officially support discipline-specific RDM needs.

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: "Based on our experience so far, a Data Champions initiative seems to be an effective way to increase both advocacy for RDM and discipline-specific training available to researchers in larger universities" (Higman, Teperek, and Kingsley 2017, 104).

The bimonthly Data Champions Forum meetings are a good mechanism for the OSC to acquire feedback on existing RDM activities in the university, as well as to generate ideas for new ones.

I am currently undertaking a five-month secondment as Data Research Coordinator at the OSC Research Data Facility, where I will be responsible for the Data Champions programme, RDM outreach and support activities, and curating data sets uploaded to the university repository. Hopefully this will be of mutual benefit: I will bring my experience and knowledge of providing discipline-specific RDM support to the centralised service and will be able to bring back a much greater understanding of how different types of research data are shared in the university to my department. Work placements such as this are one of the methods that have already been identified for librarians to develop their skills in supporting RDM, enabling them to "get out there", as Guy (2013b) advocated.

Conclusion

Central and discipline-specific RDS each have their own advantages and disadvantages. Centralised RDS are effective at coordinating RDM services and communicating RDM policy to researchers across a university. Centralised RDS are able to support researchers' generic RDM needs across a university in a cost-effective way, but they struggle to cover discipline-specific needs in any depth. Discipline-specific RDS are effective at personalising RDM support for researchers in their own subject, and this encourages them to participate in RDM activities and engage them in RDM as a concept. But because there are often many sub-disciplines within one discipline this approach cannot be sustainable either. There needs to be more focus on providing discipline-specific RDS at the planning stage and this could be informed by more research into how this could be achieved.

Research-intensive universities have been concentrated on in the literature; they are naturally leading the way in RDS provision along with other RDM stakeholders, in order to support their highly productive researchers. Some of the older, more prestigious universities, such as the University of Cambridge, are quite different to others in the way their research, teaching, and libraries are structured. However, they and the more centralised 'campus' university libraries do face similar issues with regard to providing RDS sustainably and consistently, and across disciplines.

I consider a mix of centralised and discipline-specific RDM support within a university to be the optimum model. A completely centralised approach without any discipline-specific context can, as is evident from the OSC's initial information sessions, alienate a subject community. A completely discipline-specific approach may miss out on disseminating vital centralised messages. A mixed approach allows centralised messages to be tailored to researchers in their particular field, which in my experience increases their relevance and therefore the likelihood that they will be truly taken on board.

Acknowledgements: I thank Marta Teperek, formerly Data Facility Manager at the OSC, now Data Stewardship Coordinator at TU Delft, who has consistently encouraged and supported me in all my RDM activities and who gave me the confidence to write and talk about this.

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.

I thank my colleague at the OSC, Claire Sewell, for her wise words of advice.

References

ACRL. 2017. "Scholarly Communication Toolkit." Last updated April 17, 2018. http://acrl.libguides.com/scholcomm/toolkit/.

Akers, K. G., and J. Doty. 2013. "Disciplinary Differences in Faculty Research Data Management Practices and Perspectives." *International Journal of Digital Curation* 8 (2): 5–26. https://doi.org/10.2218/ijdc.v8i2.263.

ALDL. n.d. "Agency of the Legal Deposit Libraries." Accessed May 28, 2018. https://www.legaldeposit.org.uk/index.html.

Auckland, M. 2012. "Re-Skilling for Research: An Investigation into the Role and Skills of Subject and Liaison Librarians Required to Effectively Support the Evolving Needs of Researchers." *RLUK, Research Libraries UK, Report*. http://www.rluk.ac.uk/wp-content/uploads/2014/02/RLUK-Re-skilling.pdf.

Bresnahan, M. and A. Johnson. 2013. "Assessing scholarly communication and research data training needs." *University Libraries Faculty & Staff Contributions*. 7. https://scholar.colorado.edu/libr_facpapers/7

Christensen-Dalsgaard, B., M. van der Berg, R. Grim, W. Horstmann, D. Jansen, T. Pollard, and A. Roos. 2012. "Ten Recommendations for Libraries to Get Started with Research Data Management." *Final report of the LIBER working group on E-Science/Research Data Management (Ligue des Bibliothèques Européennes de Recherche – Association of European Research Libraries*. https://www.fosteropenscience.eu/content/ten-recommendations-libraries-get-started-research-data-management.

Corrall, S. 2010. "Educating the academic librarian as a blended professional: A review and case study." *Library Management* 31 (8): 567–593. http://dx.doi.org/10.1108/01435121011093360. https://www.emeraldinsight.com/doi/abs/10.1108/01435121011093360.

Cox, A. M., M. A. Kennan, L. Lyon, and S. Pinfield. 2017. "Developments in Research Data Management in Academic Libraries: Towards an Understanding of Research Data Service Maturity." *Journal of the Association for Information Science and Technology* 68 (9): 2182–200. https://doi.org/10.1002/asi.23781.

Cox, A. M., and S. Pinfield. 2013. "Research Data Management and Libraries: Current Activities and Future Priorities." *Journal of Librarianship and Information Science* 46 (4): 299–316. https://doi.org/10.1177/0961000613492542.

Data Curation Profiles. n.d. "Data Curation Profiles." Accessed May 28, 2018. http://datacurationprofiles.org/

DCC. 2018. "Disciplinary RDM Training." Accessed May 28, 2018. http://www.dcc.ac.uk/training/train-trainer/disciplinary-rdm-training/disciplinary-rdm-training.

EPSRC. 2014. "Expectations." Accessed May 28, 2018. https://epsrc.ukri.org/about/standards/researchdata/expectations/.

Flores, J. R., J. J. Brodeur, M. G. Daniels, N. Nicholls, and E. Turnator. 2015. "Libraries and the Research Data Management Landscape." In *The Process of Discovery: The CLIR Postdoctoral Fellowship Program and the Future of the Academy*, edited by J. C. Maclachan, E. A. Waraksa, and C. Williford, 82–102. https://www.clir.org/pubs/reports/pub167/ Washington, DC: Council on Library and Information Resources.

FOSTER. n.d. "FOSTER Open Science Portal." Accessed May 28, 2018. https://www.fosteropenscience.eu/.

Gabridge, T. 2009. "The Last Mile: Liaison Roles in Curating Science and Engineering Research Data." *Research Library Issues: A Bimonthly Report from ARL, CNI, and SPARC* 265 (August 2009): 15–21. https://doi.org/10.29242/rli.265.4.

GitHub. 2018a. "An Introduction to GitHub for Chemists." Accessed May 28, 2018. https://github.com/semacu/20171024_GitHub_Chemistry_Cambridge/blob/master/README.md.

GitHub. 2018b. "Introduction to ORCID." Accessed May 28, 2018. https://github.com/semacu/20180223_ORCID_Chemistry_Cambridge/blob/master/README.md.

Guy, M. 2013a. "Increasing Participation in Internal RDM Training Sessions." *DCC RDM Services case studies*. http://www.dcc.ac.uk/resources/developing-rdm-services/increasing-participation-training.

Guy, M. 2013b. "RDM Training for Librarians." *DCC RDM Services Case Studies*. http://www.dcc.ac.uk/resources/developing-rdm-services/rdm-training-librarians.

Higman, R., and M. Teperek. 2017a. dataset. "Creating a Community of Data Champions." Accessed January 31, 2019. https://doi.org/10.17863/CAM.7417.

Higman, R., and M. Teperek. 2017b. "Template Research Data Management Workshop for STEM Researchers." Accessed January 11, 2019. http://doi.org/10.5281/zenodo.239090.

Higman, R., M. Teperek, and D. Kingsley. 2017. "Creating a Community of Data Champions." *International Journal of Digital Curation* 12 (2): 96–106. http://dx.doi.org/10.2218/ijdc.v12i2.562.

Hiom, D., D. Fripp, S. Gray, and K. Snow. 2015. "Research Data Management at the University of Bristol: Charting a Course from Project to Service." *Program: Electronic Library and Information Systems* 49 (4): 475–93. https://www.emeraldinsight.com/doi/abs/10.1108/PROG-02-2015-0019.

Kennan, M.A. 2016. "Data Management: Knowledge and skills required in research, scientific and technical organisations." Paper presented at IFLA World Library and Information Congress, Columbus, OH, 2016. http://library.ifla.org/1466/1/221-kennan-en.pdf

Koltay, T. 2016. "Are You Ready? Tasks and Roles for Academic Libraries in Supporting Research." *New Library World* 117 (1/2): 94–104. https://doi.org/10.1108/NLW-09-2015–0062.

Lewis, M. J. 2010. "Libraries and the Management of Research Data." In *Envisioning Future Academic Library Services*, edited by S. McKnight, 145–68. London: Facet Publishing. http://eprints.whiterose.ac.uk/11171/

Morais, R., and L. Borrell-Damian. 2018. "Open Access 2016–2017 EUA Survey Results." European University Association (February): 1–37. https://eua.eu/resources/publications/324:open-access-in-european-universities-results-from-the-2016-2017-eua-institutional-survey.html.

NNLM. n.d.a "NNLM RD3: Resources for Data-Driven Discovery." Accessed May 28, 2018. https://nnlm.gov/data.

NNLM. n.d.b "Library Roles." Accessed May 28, 2018. https://nnlm.gov/data/library-roles.

Open Science Framework. 2018. "Research Data Management Project Group." Accessed May 28, 2018. https://osf.io/wbt2x/.

ORCID. n.d. "ORCID." Accessed May 28, 2018. https://orcid.org/.

Pinfield, S., A. M. Cox, and S. Rutter. 2017. "Mapping the Future of Academic Libraries: A Report for SCONUL." https://www.sconul.ac.uk/news/mapping-the-future-of-academic-libraries.

Pinfield, S., A. M. Cox, and J. Smith. 2014. "Research Data Management and Libraries: Relationships, Activities,

Drivers and Influences." *PLos One* 9 (12): 1–28. https://doi.org/10.1371/journal.pone.0114734.
re3data.org. 2018. "Registry of Research Data Repositories". Accessed August 28, 2018. https://doi.org/10.17616/R3D.
Research Excellence Framework. 2017. "REF 2021." Accessed May 28, 2018. http://www.ref.ac.uk/.
Russell Group. n.d. "Russell Group." Accessed May 28, 2018. https://russellgroup.ac.uk/.
Sewell, C., and D. Kingsley. 2017. "Developing the twenty-first Century Academic Librarian: The Research Support Ambassador Programme." *New Review of Academic Librarianship* 23 (2–3): 148–58. https://doi.org/10.1080/13614533.2017.1323766.
Swan, A., and S. Brown. *The Skills, Roles and Career Structure of Data Scientists and Curators: An Assessment of Current Practice and Future Needs*. Report to the Jisc (Truro: Key Perspectives Ltd., 2008). https://eprints.soton.ac.uk/266675/.
Symplectic. n.d. "Symplectic. Elements." Accessed May 28, 2018. http://symplectic.co.uk/products/elements/.
Tenopir, C., B. Birch, and S. Allard. *Academic Libraries and Research Data Services: Current Practices and Plans for the Future*. White Paper (American Library Association: June 2012). http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf.
Tenopir, C., R. J. Sandusky, S. Allard, and B. Birch. 2014. "Research Data Management Services in Academic Research Libraries and Perceptions of Librarians." *Library & Information Research* 36 (2): 84–90. https://doi.org/10.1016/j.lisr.2013.11.003.
Tenopir, C., S. Talja, W. Horstmann, E. Late, D. Hughes, D. Pollock, B. Schmidt, L. Baird, R. Sandusky, and S. Allard. 2017. "Research Data Services in European Academic Research Libraries." *LIBER Quarterly* 27 (1): 23–44. http://doi.org/10.18352/lq.10180.
Teperek, M., R. Higman, and D. Kingsley. 2017. "Is Democracy the Right System? Collaborative Approaches to Building an Engaged RDM Community." *International Journal of Digital Curation* 12 (2): 86–95. https://doi.org/10.2218/ijdc.v12i2.561.
TopUniversities. 2018. "Chemistry." Accessed May 28, 2018. https://www.topuniversities.com/university-rankings/university-subject-rankings/2018/chemistry.
University of Cambridge. 2018a. "About the University. Timeline." Accessed August 31, 2018. https://www.cam.ac.uk/about-the-university/history/timeline.
University of Cambridge. 2018b. "Planning and Resource Allocation Office. Facts & Figures." Accessed August 31, 2018. https://www.information-hub.admin.cam.ac.uk/files/facts_figures_2018_poster_for_web.pdf.
University of Cambridge. 2018c. "Department of Chemistry." Accessed May 31, 2018. https://www.ch.cam.ac.uk/.
University of Cambridge. 2018d. "Apollo – University of Cambridge Repository. Data – Chemistry." Accessed May 28, 2018. https://www.repository.cam.ac.uk/handle/1810/247062.
University of Cambridge. 2018e. "Cambridge University Library." Accessed May 31, 2018. http://www.lib.cam.ac.uk/.
University of Cambridge. 2018f. "Chemistry Library." Accessed May 28, 2018. http://www-library.ch.cam.ac.uk/.
University of Cambridge. 2018g. "Scholarly Communication." Accessed August 31, 2018. https://osc.cam.ac.uk/.
University of Cambridge. 2018h. "Research Data Management." Accessed August 31, 2018. https://www.data.cam.ac.uk/.
University of Cambridge. 2018i. "Apollo – University of Cambridge Repository." Accessed May 28, 2018. https://www.repository.cam.ac.uk/.
University of Cambridge. 2018j. "Scholarly Communication. Open Access Project Board." Accessed August 31, 2018. https://osc.cam.ac.uk/open-access/open-access-policies/cambridge-open-access-policy-framework/open-access-project-board.
University of Cambridge. 2018k. "Chemistry Library. Open Data FAQs for Chemists." Accessed May 28, 2018. http://www-library.ch.cam.ac.uk/open-data-faqs-chemists.
University of Cambridge. 2018l. "Chemistry Library. Convert Your Files into an Open Data Format." Accessed May 28, 2018. http://www-library.ch.cam.ac.uk/convert-your-files-open-data-format.
Williamson, L. 2013. *Roles, Responsibilities and Skills Matrix for Research Data Management (RDM) Support*. ADMIRe Working Document. https://admire.jiscinvolve.org/wp/files/2013/05/RDM-roles-and-skills.pdf.
Williamson, L., and T. Parsons. 2013. "ADMIRe Project. UK HEIs RDM Service Models and Skills to Support Research Data Management." (February): 3–17. https://admire.jiscinvolve.org/wp/files/2013/05/RDM_HumanInfrastructure.pdf.
Wilson, J. A., and P. W. Jeffreys. 2013. "Towards a Unified University Infrastructure: The Data Management Roll-Out at the University of Oxford." *International Journal of Digital Curation* 2 (8): 235–46. https://doi.org/10.2218/ijdc.v8i2.287.
Wittenberg, J., and M. Elings. 2017. "Building A Research Data Management Service at the University of California, Berkeley: A Tale of Collaboration." *IFLA Journal* 43 (1): 89–97. https://doi.org/10.1177/0340035216686982.