

Maturing research data services and the transformation of academic libraries

1432

Received 19 December 2018
Revised 4 May 2019
Accepted 6 May 2019

Andrew M. Cox

Department of Information, University of Sheffield, Sheffield, UK

Mary Anne Kennan

Department of Information Studies, Charles Sturt University, Albury, Australia

Liz Lyon

*Department of Information Culture and Data Stewardship,
University of Pittsburgh, Pittsburgh, Pennsylvania, USA, and*

Stephen Pinfield and Laura Sbaffi

Department of Information, University of Sheffield, Sheffield, UK

Abstract

Purpose – A major development in academic libraries in the last decade has been recognition of the need to support research data management (RDM). The purpose of this paper is to capture how library research data services (RDS) have developed and to assess the impact of this on the nature of academic libraries.

Design/methodology/approach – Questionnaire responses from libraries in Australia, Canada, Germany, Ireland, the Netherlands, New Zealand, the UK and USA from 2018 are compared to a previous data set from 2014.

Findings – The evidence supports a picture of the spread of RDS, especially advisory ones. However, future ambitions do not seem to have seen much evolution. There is limited evidence of organisational change and skills shortages remain. Most service development can be explained as the extension of traditional library services to research data. Yet there remains the potential for transformational impacts, when combined with the demands implied by other new services such as around text and data mining, bibliometrics and artificial intelligence. A revised maturity model is presented that summarises typical stages of development of services, structures and skills.

Research limitations/implications – The research models show how RDS are developing. It also reflects on the extent to which RDM represents a transformation of the role of academic libraries.

Practical implications – Practitioners working in the RDM arena can benchmark their current practices and future plans against wider patterns.

Originality/value – The study offers a clear picture of the evolution of research data services internationally and proposes a maturity model to capture typical stages of development. It contributes to the wider discussion of how the nature of academic libraries are changing.

Keywords Information services, Academic libraries, Scholarly communication, Research data management, Data curation, Research data services

Paper type Research paper

Introduction

One of the most significant changes to academic library services in the last decade has been the development of research data services (RDS) (Pryor *et al.*, 2014). Driven by the “data deluge”, funder policy and open scholarship, libraries, in collaboration with other professional services and researchers, have developed a range of advisory and



technical services to support research data management (RDM). A burgeoning literature on the topic reflects an intense professional debate about the meaning of RDM for librarianship.

There are a number of interlinked ways in which RDM might be considered to constitute a fundamental shift in the role of academic libraries. First, this could happen through a greater focus on research support and deeper engagement with the research process, including embedding library roles in research teams. Academic librarians' focus on research has typically paid most attention to historic collections and publications as formal outputs of the research process. RDM, especially the notion of data management planning (DMP) from project initiation, implies that there is a library role in supporting information management throughout the research process and through the whole data lifecycle. A second aspect that could be seen as transformational is the way curation of research data turns the collection "inside out". RDM is perhaps the classic instance of moving from purchasing externally-produced material for institutional communities, to stewarding the outputs of such communities for the wider world (Dempsey, 2016). It also implies a re-emphasis on curation, preservation and re-use. Of the two traditional roles of access and preservation academic libraries' focus have increasingly been on access, and there is indeed an access issue around data, but issues around integrity, reproducibility and transparency and the long-term preservation of data mark a potential shift of emphasis. A third aspect of transformation could be that RDM potentially brings with it a restructuring of professional relationships. RDM can be seen as initiating new forms of collaboration on campus between professional services and faculty. New forms of relationship in the wider "coopetitive" (a balance between collaboration and competition) environment are also potentially created with other libraries or with new forms of commercial vendor (Pinfield *et al.*, 2017). A fourth linked aspect of transformation could be the impact on competencies and professional identity. Some have seen the potential for librarians to be much more heavily involved in data analysis, visualisation and research integrity, implying a significant change to professional competencies. Commentators and practitioners have coined terms such as data librarian or "databrarian" to reflect very new types of role emerging in the RDM area (Rice and Southall, 2016; Kellam and Thompson, 2016).

These trends have the potential to transform the nature of the academic library. However, it is also possible to interpret academic library involvement in RDM as essentially an extension of traditional activities, such as:

- advisory and support services;
- information literacy training: hence the term data information literacy; and
- repository management, with associated issues around collection management, metadata management and resource discovery.

This less transformative response might be indicated by a focus on simply extending existing services to encompass RDM, translating existing skills and organisational structures to the new context (rather than developing significant new competencies or restructuring) and out-sourcing some services, such as long-term preservation (if such services are not seen as based on core competencies).

In this context, this study investigates the nature of library RDS and their impact on the nature of academic libraries, using data from international surveys conducted in 2014 and 2018. This is done by seeking to answer the following empirical research questions:

RQ1. What RDM policies and types of RDS are being developed by libraries?

RQ2. What internal and external collaborations have been involved in such developments?

RQ3. What changes in organisational structure and skill sets have there been to meet the needs of RDS?

RQ4. What are perceived to be the underlying drivers and challenges?

Literature review

In their review of the “RDM service bundle”, Bryant *et al.* (2017) offer a number of general comments about the services emerging in the RDM landscape. RDS can be divided into three types, they argue: education, expertise and curation. These are done rather differently in different institutions, shaped by local and external factors. RDS are not just what is provided locally, but involves external services, access to which is brokered. Typically, services are based on local and wider collaborations of different sorts. Yet there remains continuing “fluidity and uncertainty” around RDM (Bryant *et al.*, 2017, p. 29).

A number of studies have attempted to capture in more detail the character of the development of research data policy and services at a national and international level, usually focussing on library based RDS. Overviews based on national studies have been produced of the USA (Tenopir *et al.*, 2012, 2015) and the UK (Cox and Pinfield, 2014) and more recently France (Rebouillat, 2017). Other research at a national level based on website analysis includes the study Kouper *et al.* (2015) of American research libraries (ARL). Survey studies encompassing a number of countries by Corral *et al.* (2013), Cox *et al.* (2017), Tenopir *et al.* (2016) have sought to compare development in RDS internationally and to reflect on the relative strengths of drivers and barriers. A similar approach was taken for a recent study of countries in Southern Africa (Chiwere and Becker, 2018).

Although the terms used to describe RDS differ somewhat in the different studies, some common patterns emerge. One is that “consultative” services are more common than “hands on or technological” services (Tenopir *et al.*, 2017). RDS seem to be aligned to services libraries have traditionally offered. They are typically collaborative, often with libraries working with IT centres, research offices and academic departments; they may also be collaborative with other institutions. Another common finding is the continuing limits on staff skills to support RDM. There also appear to be different trajectories of development in different regions and at national level, although librarians express a strong commitment to the library role in RDM.

In trying to make sense of these patterns in a broader way, an obvious framework would be a maturity model. These are used in a wide range of fields to map evolutionary steps in growing formalisation, control and measurement of an area of activity (Qin *et al.*, 2017). They allow organisations to benchmark performance. They are not, however, without their critics, partly because they may fail to acknowledge context and alternative paths of development and partly because they do not explain how to move between levels (Pöppelbuß and Röglinger, 2011). Nevertheless, they are widely recognised ways of developing a framework for understanding and guiding development.

There are many such models that could have generic relevance to the area of RDM, e.g. maturity models for digital preservation (Kenney and McGovern, 2003) and for digital asset management or for information governance as a whole (Proença *et al.*, 2016). Maturity models developed specifically for RDM include the Stewardship Maturity Matrix (Peng *et al.*, 2015) and the Capability Maturity Model for RDM (Qin *et al.*, 2017; Crowston and Qin, 2011). These can be seen as evaluating maturity of projects or fields in terms of achieving the ends of better data management and reuse. With a focus more on the infrastructure to support these ends, in terms of benchmarking an infrastructure of policy, support services, competencies and culture that enable RDM at an institutional level, there is

the Collaborative Assessment of Research Data Infrastructure and Objectives (CARDIO) tool (www.dcc.ac.uk/resources/tools/cardio). CARDIO was designed to be used for internal discussion to develop the case for developing services. ANDS (2018) has a similar framework for assessing maturity in five areas: policies and procedures, support services, IT infrastructure, metadata and research data.

More empirically driven, and focussing specifically on library RDS, Cox and Pinfield (2014) identified three levels of service for a wide range of different RDM activities: “basic”, “well developed” and “extensive”. Kouper *et al.* (2015) also made a simple differentiation between three levels of service maturity in ARL libraries associated with particular types of service provision:

- (1) Basic – DMP support, advice and training.
- (2) Intermediate – repositories and associated metadata services.
- (3) Advanced – a number of services such as data curation, data citation and researcher IDs.

Based on our earlier empirical study, Cox *et al.* (2017) proposed a model of RDM services, moving from compliance through capacity-building, re-engineering to stewardship associated with basic, developing and extensive services (see Figure 10 reproduced below). The model sought to map stages in the development of RDS and associated governance structures and competencies. The current study returns to the topic of developing a maturity model for RDS, drawing on a new data set, four years after the original study and therefore carrying out a longitudinal analysis of the RDM space.

Methodology

In seeking to explore libraries’ perspectives on the development of policy and services, collaboration, structures and skills and drivers and challenges, a survey designed in the form of a questionnaire was selected as the most appropriate data collection strategy (Pickard, 2012). As well as allowing direct comparison to earlier work, a web-based survey enables the participation of a large number of academic libraries spread across the globe. Therefore, a revised version of the 2014 questionnaire (reported in Cox *et al.* (2017)) was prepared. Changes were made partly to reduce the overall number of questions and also to ensure the survey’s currency, e.g. to ask about principles or standards governing policy and practice. An open text question about drivers was added to complement an existing one on challenges. A redesigned survey of 24 questions, hosted as before on SurveyMonkey (n.d.), was piloted, and then distributed between February and April 2018[1]. The research approach was approved by the University of Sheffield research ethics approval procedure, based on gaining informed consent from all participants and a commitment from the research team to ensure anonymization of individuals and organisations in the reporting of the data.

As in the 2014 study, the countries surveyed were Australia, Canada, Germany, Ireland, the Netherlands, New Zealand and the UK. In addition, a request to participate was extended to the ARL group in the USA. In most cases, an invitation to respond was emailed directly to individual library directors, with some more broadcast distribution (Germany and USA). The survey received 209 valid responses. As is apparent from Table I, the response rate from Australia, Ireland and New Zealand was good. In addition, around 50 per cent of UK institutions responded, comparable to 2014. There were slightly more disappointing responses from other countries, though the number participating from Canada was increased. The variable response rate, combined with the low total number of institutions in some countries prevents some more sophisticated forms of statistical analysis of the data. The large number of responses from the UK potentially skews the results. There is also

likely to be some non-response bias, with institutions with a commitment to RDM more likely to respond. As a result, analyses need to be treated with caution, nevertheless, the study provides a substantial amount of data on current trends in RDS.

The data from the questionnaire were analysed through descriptive statistics and some factor analysis. In addition to closed questions, a substantial amount of data was collected in the form of text responses to open-ended questions, principally about drivers and challenges, amounting to around 15,000 words of data. This text was analysed through content analysis.

Findings

The results of the survey are presented here in sections broadly corresponding to the research questions outlined above, relating to policies and types of services, collaboration and competition, structures and skills and drivers and challenges.

User requirement gathering, institutional policy making and service development

Two questions in the survey examined the extent to which institutions had gathered data about needs in their institution, differentiating a “data audit, survey or evaluation” and a survey of staff attitudes. Whereas in relation to the first question, participants were given the option to say one was planned; it was a simple yes/no question for the second. In addition to the slightly different options, the results have to be understood through differences in wording: the phrase “audit, survey or evaluation” as being more inclusive of different options of undertaking a study than a survey (Figures 1 and 2).

Table I.
Response rate
by country

	Australia	Canada	Germany	Ireland	The Netherlands	NZ	UK	USA	Total
Invited institutions 2018	39	74	279	12	16	8	169	86	404
Number of responses 2018	34	24	23	11	6	8	80	23	209
% response rate 2018	87	32	10	92	38	100	47	27	
Invited institutions 2014	39	74	48	8	16	7	169	na	361
Number of responses 2014	34	17	8	7	12	7	85	na	170
% response rate 2014	87	23	17	88	74	100	50	na	

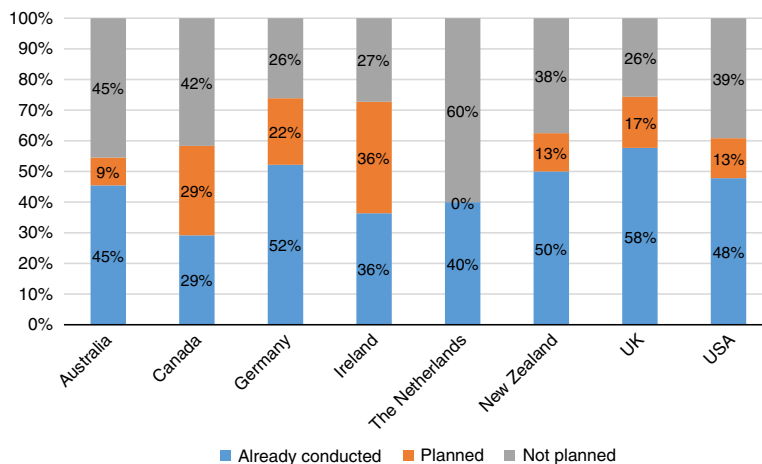
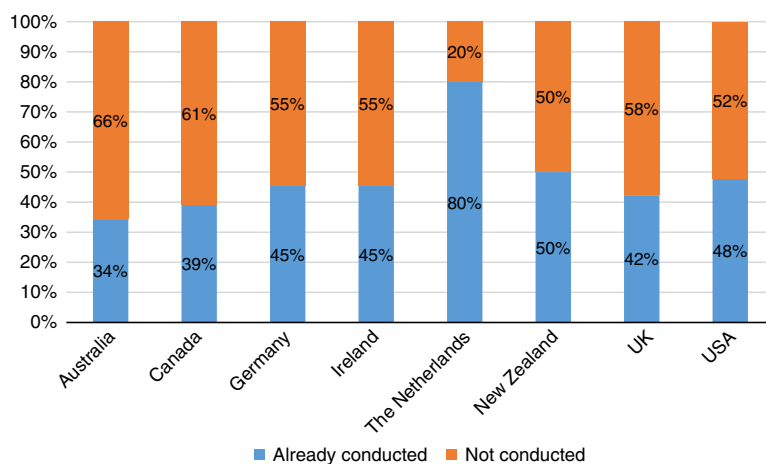


Figure 1.
Audit, survey or
evaluation of the
institution's data (see
also Table A1)

Note: n=205



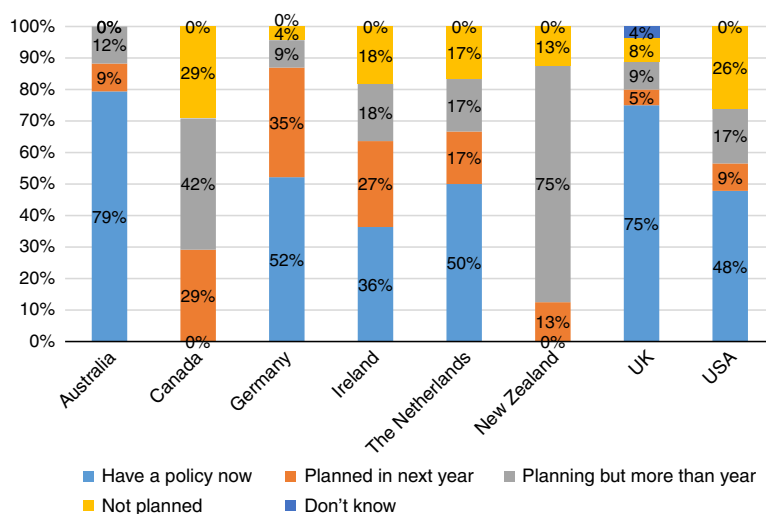
Note: n=202

Figure 2.
Frequency of having undertaken a survey of faculty/academic staff attitudes to RDM (see also Table AII)

In total, 62 per cent of institutions had undertaken at least one of these types of study and 29 per cent done both (sometimes in the form of a single survey covering both topics). Overall 49 per cent of responding institutions had undertaken a data study; and 43 per cent a user study. Of the former most had had library involvement. There was a surprising number of institutions still planning such studies, rather than having undertaken them already, especially in Canada and Ireland.

A second area of interest for the survey was policy development. As Figure 3 reveals, the pattern of policy development was quite varied.

Over half, 117 (56 per cent) of responding institutions had a formal RDM policy; 29 (14 per cent) were planning to have one in the next year; and 36 (17 per cent) in a



Note: n=209

Figure 3.
Formal RDM policy in place or planned (see also Table AIII)

longer timescale. Only 11 per cent had no plan for a policy. Three quarters of Australian and UK institutions did already have a policy at the time of the survey. In contrast, none of the responding institutions in Canada or New Zealand had one currently. In 2014, 64/167 (38 per cent) of institutions had a policy in place; rather more 76 (46 per cent) had plans for one. This indicates “progress” across the sector, but it seems reasonable to infer that not all the institutions who said they were planning a policy in 2014 did actually implement one. “Progress” is also discernible in relation to the broad range of services offered by libraries. Based on treating respondents indicating they have no service = 0, basic service = 1 and well developed or extensive service = 2, Figure 4 offers a clear visualisation of the progress in service delivery made in developing RDS, but without a major shift in emphasis away from advisory services. Note some items were not available as options in 2014; simplified wording for the options is used.

Advisory services, such as DMP, web guides, data discovery and support on copyright remained the main types of services being offered by libraries. Of the top ten services provided by libraries in 2018, nine were “advisory”. Running a repository was the only common technical service. This is a similar pattern to 2014. Advisory services on data analysis, data mining and services for cleaning data, analysing and visualising data and rescuing legacy data remained relatively rare. Rewording of the option from “Directly participate with researchers on a research project (as a team member)” to “Embed librarians in the laboratory or research project” has resulted in a fall in agreement.

Figure 5 suggests relatively little change has taken place in what libraries consider priorities between the two studies.

Comparing current services with priorities, it does seem that the repository was a higher future priority (3rd highest priority) than current service (ranked 10th). Thus, it appears that the ambition is to create more technical services, although the differences in ranking were not found to be statistically significant.

Collaboration and competition

In addition to asking respondents about user requirements studies, policy and RDS development, the questionnaire sought to discover the nature of the collaborations with research administrators, IT services, academics and legal teams that lay behind these developments.

The data revealed that while IT services and academics were commonly involved in both policy and RDS development, they were rarely perceived to have led such activity. In the development of policy, it was the library and research office who were typically seen

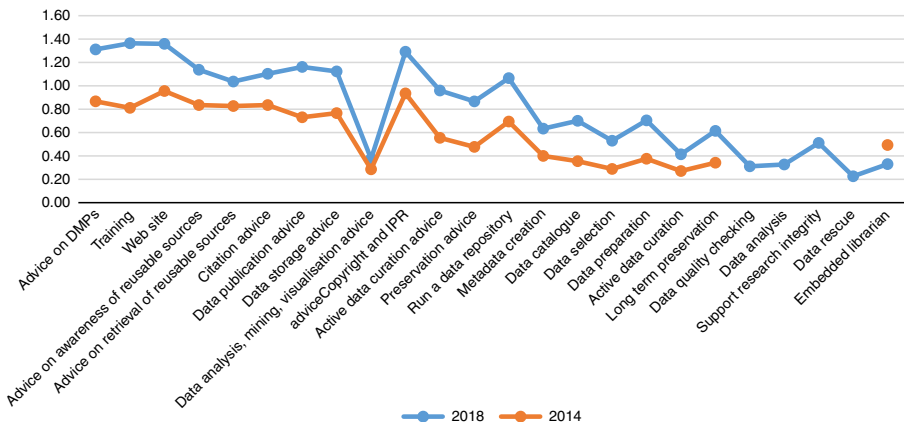


Figure 4.
A comparison of RDS between 2014 and 2018 (see also Table AIV)

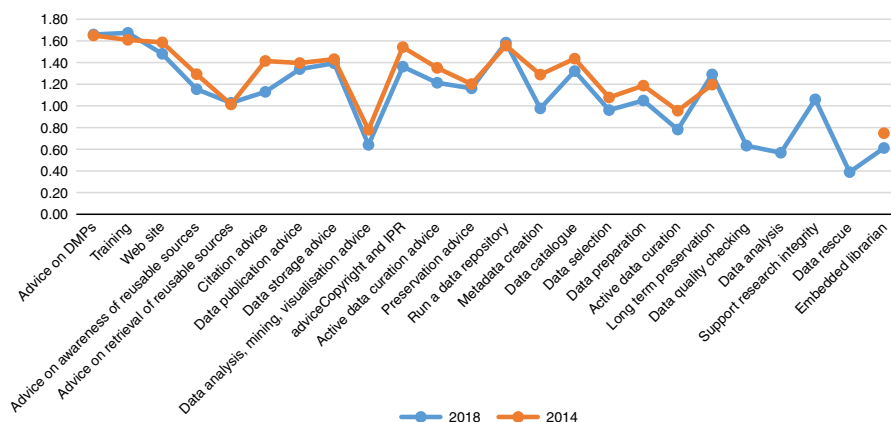
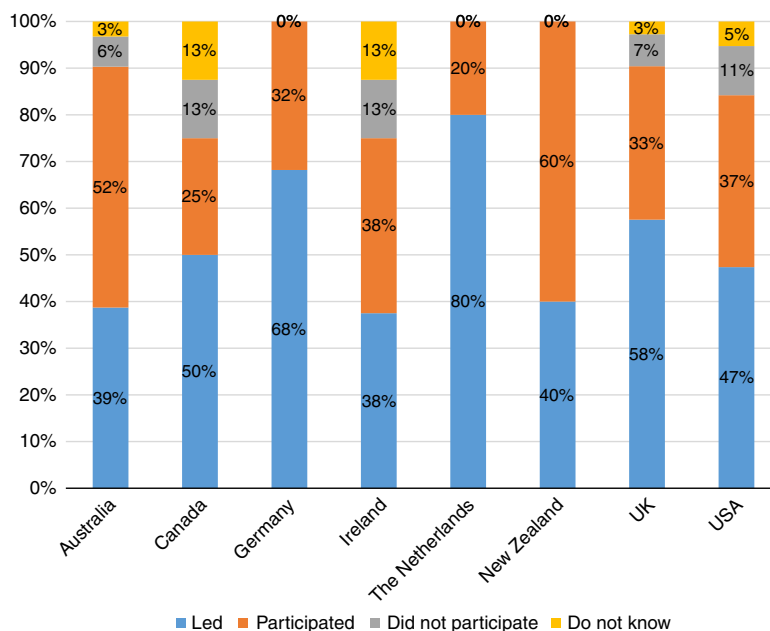


Figure 5.
A comparison of
RDS priorities 2014
and 2018 (see also
Table AIV)

as leads; while in development of services it was usually the library that had led development (Figures 6–9 below). The exception was Australia where research office seemed to be much more likely to leading in policy development and equally likely to be involved in service development.

The survey also asked about whether the library used or partnered with “any external organisations, for example commercial providers such as Figshare and Arkivum or other universities or university libraries, to provide research data services (Table II)”.

Nearly 50 per cent of respondents said that their library was cooperating with an external organisation; another 20 per cent planned to do so. This was somewhat similar with the overall figures in 2014. However, there did seem to have been some changes, for while reported



Note: n = 179

Figure 6.
Library involvement
in research data policy
development (see also
Table AV)

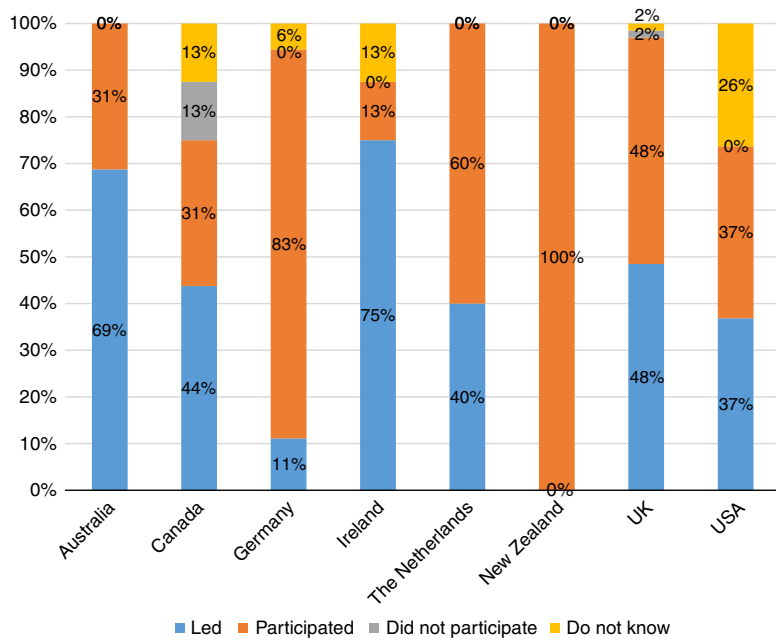


Figure 7.
Research office involvement in research data policy development (see also Table AVI)

Note: n = 168

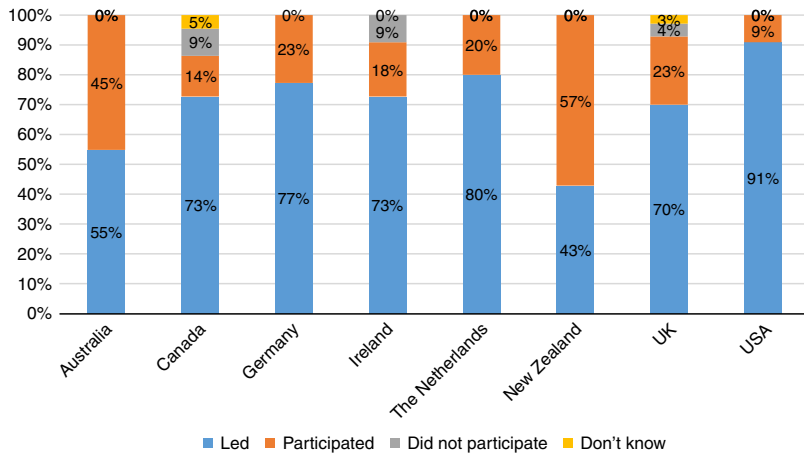
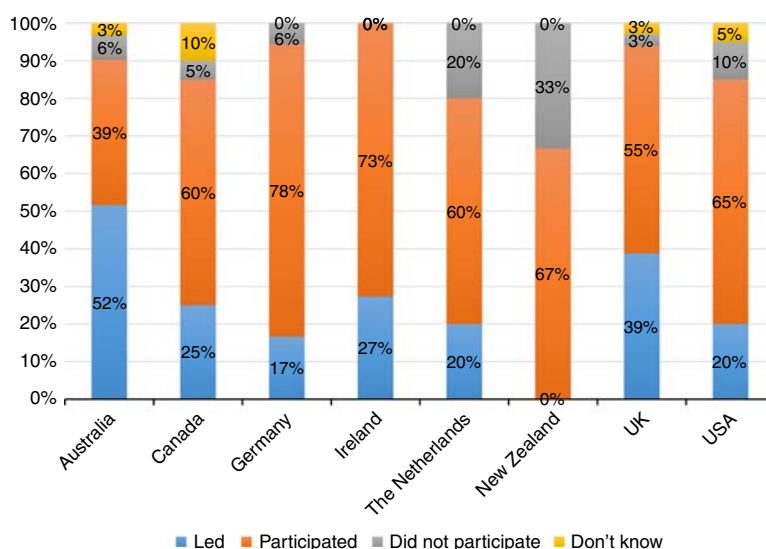


Figure 8.
Library role in research data services (see also Table AVII)

Note: n = 190

collaboration by Australian institutions had declined; UK collaborations had increased. A free text box requested details of collaborations. Analysis of this further information suggests a need for caution in interpretation because some saw simply using services such as DMPonline (which is a service run by DCC but can be customised for local users) as a collaboration, whereas others appeared to be referring to much more substantial forms of collaboration such as shared



Note: n = 178

Figure 9.
Research office role in
research data services
(see also Table AVIII)

	2018 (%)			2014 (%)		
	Yes	No (planned)	No (not planned)	Yes	No (planned)	No (not planned)
Australia	50	16	34	76	9	15
Canada	50	13	38	67	0	33
Germany	43	35	22	60	20	20
Ireland	18	27	55	14	14	71
The Netherlands	60	20	20	82	18	0
New Zealand	38	25	38	14	14	71
UK	47	23	30	22	29	49
USA	52	9	39	nc	nc	nc
Overall mean value	46	21	33	42	20	38

Table II.
A comparison of
current and planned
external collaborations
(see also Table AIX)

services. However, some patterns do seem clear. Most mentions were of commercial organisations that are providing a service such as Arkivum (19 mentions) or Figshare (15 mentions). A very wide range of commercial services were mentioned by Australian respondents, such as cloudstore and redbox; UK respondents mentioned a narrower range. National level initiatives such as Portage (CAN) and Jisc shared service (UK) were also mentioned frequently. Across all the data, collaboration with other universities was given far less emphasis than commercial collaborations. So part of the pattern seems to be to move towards out-sourcing and commercial relationships, but there was not much evidence of institutional collaborations among libraries. One participant commented:

We all seem to be solving research data management problems at a local or regional level and there is a place, in my opinion, for a national or international consortia approach. (AUS)

Structures and skills

To explore organisational structures in place to support RDM one question asked respondents to identify “who in the library has primary leadership responsibility for plans

and programs of research data services". Some people chose more than one option; and it may be hard to articulate complex organisational structures to fit the question, nevertheless the result does give some suggestion of differing experiences (Table III).

Responses diverged, reflecting a diversity of organisational approaches. About a quarter of institutions have a single person responsible, presumably reflecting a relatively low commitment to, or capacity for, RDM. About a fifth had a team with a remit for research support as a whole. 15 per cent spread responsibility quite widely. A quarter had a team with a focus on research data, a structure that was suggestive of a strong commitment to RDM. There seemed to be some strong national differences with Australian universities more likely to have a team dedicated to research support in general; whereas in the Netherlands the dominant model was to have a team for research data as such. In Ireland the commonest pattern was a single individual.

To analyse this further, Table IV seeks to correlate structure and service type. This presents the total number for each type of organisational structure for the 26 services the number of times no service, basic service or more than basic service was selected. In other words, 549 out of 1,323 total replies from organisations that said a single individual was responsible for RDM indicated that they provided no service under one of the 26 services. The cases where respondents chose more than one option were excluded for this analysis.

Table IV suggests that institutions with a specific team dedicated to RDM typically had stronger services. Thus looking at the average across all services 37 per cent of institutions with a RDM team had well developed services, compared to around only 20 per cent for each of the other forms of organisation. The other options for organisational structure do not seem to be very strongly correlated with service level. Looking across services, while well developed services were rare, they were particularly concentrated in institutions with an RDM team, e.g. of the six institutions with well developed services around "Clean data and carry out data quality checks" all were in institutions with dedicated RDM teams.

Table III.
Organisational
structures for RDS
(see also Table X)

	A single individual is responsible (%)	A team with a specific focus on research data is responsible (%)	A team with a general remit for research support is responsible (%)	It is spread across multiple teams (%)
Australia	18	18	31	26
Canada	11	29	11	14
Germany	15	42	15	12
Ireland	40	13	20	13
The Netherlands	20	80	0	0
New Zealand	20	10	20	10
UK	25	17	22	10
USA	31	26	11	26
Overall mean value	23	23	19	15

Table IV.
Service level
correlated with
organisational
structure

	A single individual is responsible	A team with a specific focus on research data is responsible	A team with a general remit for research support is responsible	It is spread across multiple teams
No service	549 (41%)	334 (25%)	277 (40%)	207 (41%)
Basic service	502 (38%)	490 (37%)	273 (40%)	202 (40%)
Well developed service	272 (21%)	487 (37%)	139 (20%)	94 (19%)
Total	1,323	1,311	689	503

Another question sought to identify potential areas of skills gap. It was clear that most libraries did perceive there to be a skills gap, in nearly every area identified, but particularly around data curation skills, knowledge of research methods and data description (Table V). It is quite interesting that an area such as data description is a gap, given that this is an area where librarians could be expected to have existing expertise. This might reflect the way that libraries have moved to outsource technical processes including metadata creation.

A comparison between the two surveys (Table VI) does not suggest a major shift in the skills gap, except around the research lifecycle, which now seems to be considered less important.

Drivers and challenges

Of 209 respondents to the survey, 167 wrote something in the open text box, “Briefly comment on what you think are the major drivers, now and looking into the future, for libraries working in research data management” amounting to a total of about 5,000 words of data. Most responses mentioned a number of drivers, so 350 items were coded (Table VII).

Compliance with funder policy was the most commonly mentioned driver, followed by the library ability/need to find a role and researcher need. In total, 57 per cent of those responding mentioned funder policy and another 8 per cent mentioned the UK REF or equivalent:

Funders’ requirements as that drives the need for DMPs, on which can be hung RDM services. (IRE)

The REF is a major driver, and a lever for getting additional resources in this area, so that the University can demonstrate it is providing adequate support for RDM as part of a healthy research

	AUS	CAN	GER	IRE	NETH	NZ	UK	USA	Total	Rank
Data curation skills	21	16	16	10	3	8	57	21	152	1
Technical and ICT skills (e.g. data storage, infrastructure, architecture, etc.)	21	13	14	6	1	7	45	16	123	5
Subject and or disciplinary knowledge	11	6	14	6	1	4	27	14	83	8
Knowledge of a variety of research methods (e.g. data analysis, data visualisation)	22	15	12	10	2	5	51	20	137	2
Knowledge of the research lifecycle	11	7	12	7	0	5	37	8	87	7
Data description and documentation	17	15	14	10	2	7	50	18	133	3
Legal, policy and advisory skills (e.g. intellectual property, ethics, licencing, etc.)	19	14	16	9	3	7	44	16	128	4
Understanding of research integrity, reproducibility and transparency principles	15	12	11	7	1	8	45	16	115	6
Total response rate for the survey as a whole	34	24	23	11	6	8	80	23	209	

Table V.
Skills needing
development

	Total 2014		Total 2018	
	(%)	Rank	(%)	Rank
Data curation skills	74	1	73	1
Knowledge of a variety of research methods (e.g. data analysis, etc.)	65	4	66	2
Data description and documentation	72	2	64	3
Legal, policy and advisory skills (e.g. intellectual property, etc.)	71	3	61	4
Technical and ICT skills (e.g. data storage, infrastructure, etc.)	60	5	59	5
Understanding of research integrity, reproducibility and transparency principles	na	na	55	6
Knowledge of the research lifecycle	58	6	42	7
Subject and or disciplinary knowledge	36	7	40	8

Table VI.
Skills requiring
development
comparison

Table VII.
Major drivers of
RDM services (from
167 responses)

	AUS	CAN	GER	IRE	NETH	NZ	UK	USA	Total	%
Funder compliance	17	14	4	8	1	4	39	8	95	57
Library role – having the skills/need to stay relevant	7	6	8	5	5	3	25	9	68	41
Needs of researchers	7	5	7	3	1	2	9	7	41	25
Integrity	7	1	1	1	1	0	12	4	27	16
Open science	6	1	0	1	1	0	12	1	22	13
Publishers	3	4	1	0	0	1	6	4	19	11
Impact of research	3	2	1	1	1	1	6	4	19	11
Institutional policy	7	1	0	1	0	0	4	2	15	9
REF (or equivalent)	0	0	0	0	0	1	13	0	14	8
FAIR	4	0	1	0	0	0	1	0	6	4
Competition from commercial suppliers	1	0	0	0	0	0	0	1	2	1
Consortia	1	1	0	0	0	0	0	0	2	1

environment[.] Compliance – if the REF doesn't work in terms of getting additional resource, then we are able to use the stick of compliance with Funder Policies! (UK)

Thus, it seemed that compliance to a top-down policy, rather than bottom-up demand was the key factor. Publishers' policies were quite often cited as a driver, usually in tandem with funder compliance. Institutional policy was less often mentioned.

In total, 41 per cent focused on the role of the library, either in terms of the contribution the library could make, or the need of the library to re-invent its role:

It is a natural extension of the 'traditional' library tasks. (NETH)

Academic libraries have the expertise to lead and support RDM on campuses and we are already involved in the data discussions at the early stages of research. We also have a need to access data resulting from research so there is a wonderful "full circle" reuse/recycle/upcycle system in place when we are involved. (CAN)

In some ways it is the organisation, management and preservation of the data that appeals to the curatorial instincts of the library. There may also be a sense that the re-usability of data will be critical for the future as budgets continue to shrink – so ensuring its discoverability and provenance is important. I think also it is a "territory" that librarians feel they should claim because there is what might be called a vacuum with IT and research units stepping back from the tasks. (UK)

The need to expand the services we offer to keep the library "current". (UK)

Only 25 per cent mentioned researcher requirements, and it seemed to be a discourse of researcher need rather than direct demand:

[...] learning about the needs of researchers. (GER)

Other factors that did get some mention were research integrity and open science:

Research integrity is a major driver from the Dean of Research now both to ensure good scholarship and minimise the university's exposure to risk from poor practice. Open scholarship links closely with the university's mission and is gaining prominence as a good in itself. (UK)

Slightly more participants (170) wrote a response to a similar question about "major challenges". In total, 6,000 words were coded under 400 items (Table VIII). Headings developed in 2014 were used; only two new codes were added (competition from publishers and lack of mandate/rewards).

Table VIII.
Major challenges
(coded as per 2014
analysis; new codes
marked *) from 170
responses

	AUS	CAN	GER	IRE	NETH	NZ	UK	USA	Total	%
Skills	10	9	9	6	0	4	23	6	67	39
Resourcing – financial	9	6	4	4	1	2	23	6	55	32
Engagement of academic staff	9	5	3	2	1	5	19	4	48	28
Collaboration with other support services	5	5	3	4	0	6	10	1	34	20
Resourcing – staffing	3	4	4	0	0	0	12	3	26	15
Infrastructure	8	3	2	4	0	0	7	2	26	15
Senior support – institutional	1	0	1	1	0	2	8	3	16	9
Understanding disciplinary differences	2	1	5	1	2	0	3	2	16	9
Acceptance in the institution	2	0	2	2	0	0	6	2	14	8
Lack mandate/rewards*	0	2	1	1	1	2	6	1	14	8
Reprioritising in the library	3	0	0	0	0	1	4	2	10	6
Acceptance of RDM role in the library	3	1	0	0	0	0	2	3	8	5
Acceptance of data sharing	1	2	2	0	0	0	3	0	8	5
Legal issues	1	1	1	1	0	0	4	0	8	5
Data - scale, variety	2	1	0	0	0	0	2	1	6	4
Rapid technology change	2	0	2	0	0	0	2	0	6	4
Competition from publishers*	2	2	0	1	1	0	0	1	7	4
Acceptance of the need for RDM	2	1	1	0	0	0	1	0	5	3
Collaboration between institutions	1	0	1	2	0	0	1	0	5	3
Preservation	0	0	0	0	0	0	4	0	4	2
Senior support – library	0	1	0	0	0	0	0	1	2	1
Compliance with funder requirements	0	0	0	0	0	0	2	0	2	1

The most frequently mentioned challenges were resources (financial – 32 per cent staffing – 15 per cent) and library skills (39 per cent). The perceived shortfall in library skills combined with resource shortages both in financial and staffing terms seemed to be the main barrier:

[...] resources and time, there are many areas we are being pulled into but we do not have the staffing or relevant expertise. (UK)

A major challenge is doing this as well as everything else. Also, RDM is much more complex than most other things we do. (UK)

[...] staff skills and willingness to take on new tasks that are not viewed as traditional “library” tasks. (AUS)

Presumably, this in turn links to less than whole-hearted funder commitment. Working with other central services and immaturity of infrastructure were other issues. Lack of engagement from faculty was also a major factor, mentioned in 28 per cent of responses:

Faculty apathy for sharing; culture of research antithetical to sharing in some disciplines. (AUS)

The chicken and egg scenario of RDM remains. You need to have a service in place to promote effective RDM practices, but it is hard to fund and develop a service without evidence of demand for that service, or to decide how to scope it. We are still in advance of academic demand for RDM. (UK)

Other factors mentioned were on collaboration between support services and infrastructure:

Technologies – both the increasing range of technology used in the creation and curation of data and the infrastructure solutions that libraries and associated partners will need to develop. (AUS)

It is interesting that competition from publishers emerged as a challenge, though it was mentioned by only a handful of respondents:

Commercial competition. Large international publishers foresee that there is hardly any growth in licensing, so they will focus on research data too. Within five years Elsevier will offer a comprehensive data management tool. (NETH)

Discussion

Thus our findings broadly confirm previous studies that have shown a similar strength and type of commitment to RDS by libraries (Corrall *et al.*, 2013; Cox *et al.*, 2017; Tenopir *et al.*, 2016, 2017). Consultative services are usually more strongly developed by libraries than technical services (Tenopir *et al.*, 2017). Service development is collaborative with other entities in the institution. Librarians still perceive a skill gap in providing all the RDS they would wish. The drivers and challenges, particularly around the engagement of researchers with RDM, confirm those uncovered by previous qualitative studies, such as Faniel and Connaway (2018) (Figure 10).

Reviewing the maturity model

The model proposed in our earlier paper is based on a logical development of activities, supported by data from our 2014 survey and suggests a number of levels, indicative of RDM maturity. The dimensions of maturity are associated with particular activities:

- (1) audit current status to develop an understanding of needs (at Level 0 before any service exists);
- (2) develop basic web guides (during Level 1);
- (3) develop policy/governance for RDM (higher up in Level 1);
- (4) develop advisory services (Level 2); and
- (5) develop a data repository and associated technical services (Level 3).

In terms of the common types of service, this general pattern still seems to fit the 2018 data. It remains the case that advisory services are more common than technical services. However, it seems that basic web guides no longer stand out as the commonest type of service, and therefore this emphasis could be removed from the model.

However, the model does not appear to reflect the 2018 data in two main ways. First, the responses to questions about surveying data/users show that fewer institutions have ever conducted a formal audit than have developed services. One interpretation of this is that if the driver is compliance the incentive to discover user need is reduced. It is widely acknowledged that actual demand is low. In this context, a service may be required to achieve compliance, regardless of need. An alternative reason could be that institutions relied on what others had

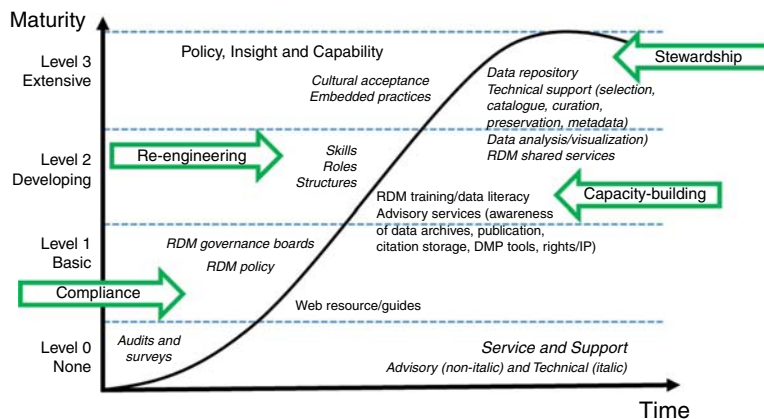


Figure 10.
The RDM
maturity model

Source: Authors

discovered about need. There could be seen to be a consensus at the level of the library community about services required; in these circumstances a study of need might seem unnecessary. It could also be either that such information gathering did happen but people who completed the survey were not aware, or that such data gathering was done informally.

The second area of difference from the original model is suggested by the fact that the number of institutions with policies is consistently lower than the number with some basic services. Indeed, in two countries (CAN and NZ) no institution had a formal policy, but there was development of services. This is slightly harder to explain, since the institutions that took part in the survey would tend to have a formal policy in areas which were regarded as important from an institutional point of view. It could be this reflects that the library sector has understood the need to respond to the RDM agenda, but the institution as a whole has not recognised it. Thus, it was evident that developing a policy was always a collaborative enterprise, often led outside the library. On the other hand, service development involved others but was almost always library led. If there is a lack of commitment to RDS outside the library, it could be this that would prevent policy formation; but it would not prevent some basic services being provided. Having services but no policy is a weak posture, but could be interpreted to reflect libraries dynamically following the sector standards, pushing institutions further than they as a whole acknowledged a need.

Both differences between the earlier model and our new findings suggest that while the model reflects a clear logic of development, in practice services can develop in a not wholly “rational” pattern. Stage 4 can come before Stages 1 and 3. Though in many cases development could happen in the order suggested, the data suggest that it is not necessarily the case. This could be less to do with differing patterns of development, than that if a consensus emerges within the library community some elements of the model are leapfrogged.

Determinants of service provision

Examining Table IX below “Ranking of services” it seems possible to argue that the frequency with which RDS are developed is ostensibly shaped by two (interlinked) factors, namely: how similar they are to libraries’ standard services and also to the level of resources required to offer them.

Thus many of the top 10 most frequently supplied services such as training, offering advice, etc. are close to services that libraries commonly already offer. The (simple) logic is that if the potential service is like what libraries already do they will extend the service to cover research data. The reasoning is practical: this approach requires them to make the least effort, in a context of resource scarcity. It also takes them least outside their traditional area of “jurisdiction” meaning that it fits in with their knowledge base and has recognition as an appropriate activity by other stakeholders. This argument would explain most of the pattern of the top 9 services.

The tenth commonest service was a data repository. This reflects the same logic. Libraries commonly already have an outputs repository, so delivering this kind of service is something they typically do. However, it is a specialist role and building a repository requires greater effort than simply creating a training programme or other advisory type services, and so is likely to be slightly further down the list of commonly supplied services. It may also be where libraries are acting more to broker use of existing services, rather than creating their own. In addition, several of the “technical services” listed in the survey are dependent on having a repository before they are required, so it is not surprising they are a little less common, specifically:

- offer a service creating or transforming metadata for data or data sets;
- provide a data catalogue including your institution’s research data;

	Any library service (%)	<i>n</i> (total responses)	Similar current library activities
Promote awareness of reusable data sources, such as data archives	83	168 (202)	Information discovery role
Offer advice on copyright and/or intellectual and/or licensing property rights relating to data and data management	81	164 (203)	Copyright advice role
Data management training and/or data literacy instruction (e.g. to research students, early career researchers, etc.)	81	164 (202)	Information literacy training
Maintaining a web resource/guide of local advice and useful resources for RDM	79	161 (203)	Advice/guides on information resources
Data management planning (DMP) advisory service	76	155 (203)	Not aligned
Offer data citation advisory services	76	155 (203)	Citation advice role
Offer data publication advisory services	75	152 (203)	Scholarly publishing advisory
Provide support for search and retrieval of external data sources	73	145 (200)	Information discovery role
Offer data storage advisory services	68	138 (203)	Not aligned
Run a data repository/archive/store	67	132 (198)	Open access repository role
Provide advisory services on the curation of active data	63	125 (199)	Not aligned
Provide advisory services on the technical aspects of long-term data preservation	61	122 (202)	Preservation role
Prepare data/data sets for deposit in a repository	52	102 (197)	Collection management
Offer a service creating or transforming metadata for data or data sets	46	92 (198)	Collection management
Provide a data catalogue including your institution's research data	44	88 (198)	Collection management
Carry out long-term preservation of research data	43	86 (198)	Preservation
Select, accession and/or deselect and deaccession data/data sets for deposit in a repository	39	77 (197)	Collection management
Support reproducibility, transparency in workflows and research integrity	34	67 (198)	Not aligned
Carry out the curation of active data	29	56 (196)	Not aligned
Offer an advisory service on data visualisation	28	56 (201)	Not aligned
Embed librarians in the laboratory or research project	27	53 (198)	Not aligned
Clean data and carry out data quality checks	26	50 (196)	Not aligned
Offer an advisory service on data analysis	24	48 (202)	Not aligned
Offer an advisory service on data mining	23	46 (201)	Not aligned
Analyse and visualise data sets using Python scripts, SPSS, R and MS Excel software	21	41 (196)	Not aligned
Rescue legacy data or perform data triage or forensic data recovery	16	31 (196)	Not aligned

Table IX.
Ranking of RDS

- select, accession and/or deselect and deaccession data/data sets for deposit in a repository; and
- prepare data/data sets for deposit in a repository.

Activities that seem to be relatively uncommon include advice on data analysis and visualisation, forensic activities and embedded roles. These are least close to what libraries already do and much more like the activities of researchers. Libraries have traditionally not been involved in analysis of content, more in facilitating discovery of and access to it. As well as requiring effort in terms of developing skills, these would not be so easily accepted as appropriate library activities and could potentially be seen as

competitive with other campus services, such as the IT service or even researchers themselves. Embedded roles are a very different form of organisational structure. Together these might be usefully differentiated as a separate level of activity or maturity in an adapted model. They could be legitimately interpreted as “transformative” because they go well beyond what are familiar areas of library activity and so signal a more fundamental shift of the jurisdiction of the profession.

One major exception to the general pattern is DMP. It is not closely aligned to what libraries already do, both in terms of the knowledge required or the timing of requests for support. It is often a pre-project-start activity, whereas libraries have tended to be involved primarily with output stage activity. Yet it is quite high up the ranking of services offered, at fifth, with three quarters of institutions offering a service. We know anecdotally that the main area of actual demand for services is around DMPs, where funders require them as a part of a project proposal. DMP advice is also resource intensive because it requires individualised feedback, often in a constrained time period. It seems to be the one area where currently libraries have extended their work outside usual territorial boundaries in order to address an area of explicit demand. However, work of this sort does seem to be an extension of activity associated with developing an RDM policy and guidelines, and, therefore, might be seen as a “natural” extension of that RDM-related role already undertaken by the library in many institutions.

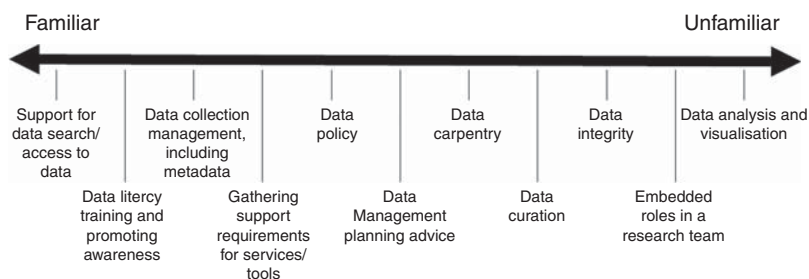
Thus in summary, we could argue that the determinants of service provision are:

- (1) closeness to existing skills/service patterns;
- (2) resources required; and
- (3) evidence of immediate demand.

This relates to what Cox (2018) has referred to as the data role spectrum the range of data related roles from those similar to library work to those more distantly connected (Figure 11).

Thus, the data suggest a potential revision to the stage model:

- (1) audit current status to develop an understanding of needs (at Level 0 before any service exists) – but optional;
- (2) develop policy/governance for RDM (Level 1) – but optional;
- (3) develop advisory services (Level 2);
- (4) develop a data repository and associated technical services (Level 3); and
- (5) develop services around data analysis, visualization, etc.; research integrity; embedded roles (Level 4).



Source: Adapted from Cox (2018)

Figure 11.
The data role
spectrum

Another way of interpreting this model is to see it less as a model of maturity, but more as defining levels of services appropriate to different types of institution. In a research-intensive institution, the whole gamut of services is justified, including support around data analysis. In a teaching orientated institution, where research is less crucial to service, there is still a need for advisory services, but there may not be a need for a technical infrastructure. Many institutions sit somewhere on the range in between.

One way to test this hypothesis would be to look at differences between the research-intensive institutions and others. This was hard to do across the whole data set, so we focussed on the UK. Simplifying significantly we can take the Russell Group of 24 institutions to represent the research-intensive institutions (Russell Group, 2019).

Figure 12 illustrates the difference in services between Russell Group and non-Russell Group. It shows that the Russell Group institutions have more well developed services in areas where most institutions have at least some service. There is less difference around relatively rare services such as data analysis or even having a repository. Thus, the difference was statistically significant for training, web guides, support for the retrieval of external sources, data citation advice and data publication advice and advice on copyright. The Russell Group institutions have stronger services in areas where services are common. Significantly, however, the differences of ambition represented in Figure 13 are small. This suggests that all institutions aspire to a similar service, but the Russell Group institutions are more likely to have the existing skills and resources to deliver them. Relating this to the model it suggests that they have a stronger commitment to compliance rather than one to stewardship, and certainly not transformation.

This data also point to a lack of evidence for different aspirations around RDM in different institutions, contrary to the arguments of Bryant *et al.* (2017). Although the lack of difference may reflect insufficient resources or indeed failure to understand their long-term needs, it does not point to strong variations in what are perceived to be appropriate RDSs. It is noticeable that Bryant *et al.*'s claim does rest on looking at the experience of a few pioneer institutions. It could be that while these do indeed approach RDM differently, because they are inventing the idea of RDS for themselves, outside this group of early adopters, institutions tend to have similar aspirations. This is consistent with Akers *et al.*'s (2014) findings that the rather diverging paths of development of service in eight US research-intensive institution were brought into closer alignment by funder policy requirements.

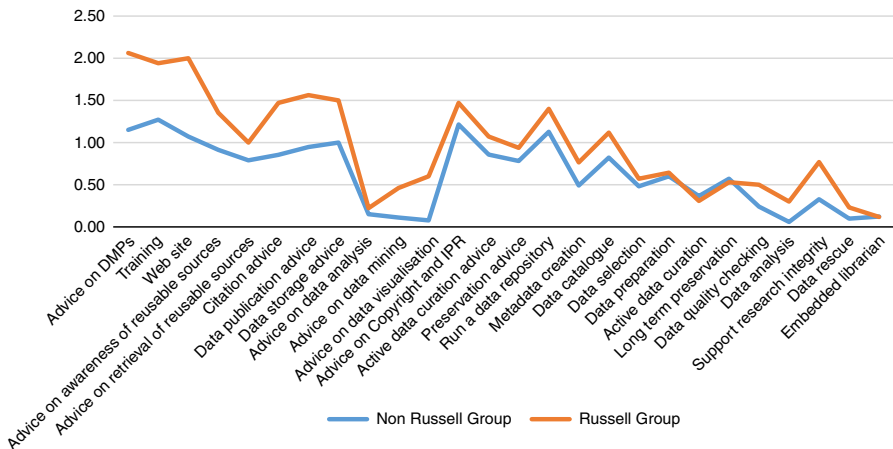


Figure 12. Comparison of UK russell group and non-russell group universities' RDS (see also Table AXI)

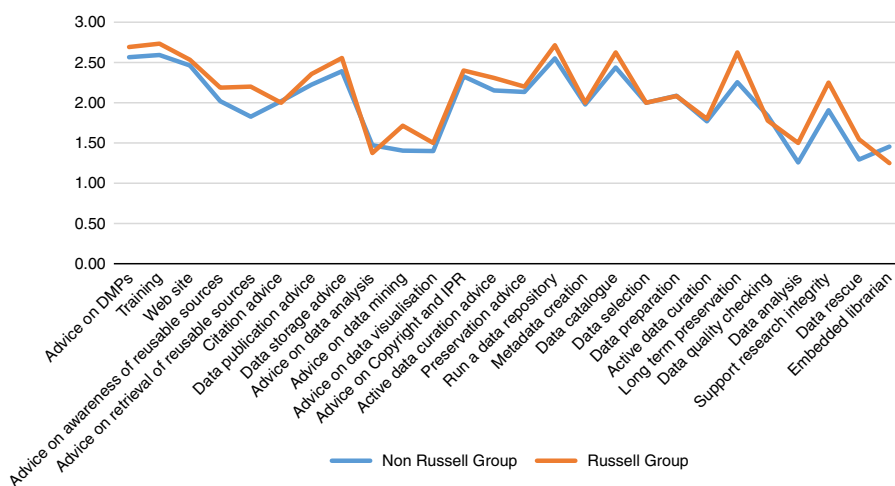


Figure 13. Comparison of russell group and non-russell group universities' plans for RDS (see also Table AXI)

Collaboration and competition

Our original model only mentions collaboration on or beyond campus as an aspect of maturity in listing “shared services” as an aspect of Level 3. However, within the logic discussed above, collaboration and competition could be seen as linked to the determining issues of resource scarcity and jurisdiction. Where the demands of RDS significantly go beyond existing practice, there is an implication of a potential need for collaboration on campus to find resources and to agree on who does what. The data support the impression of the need for collaboration both in policy and RDS development. Most responses implied new services were a collaborative effort, albeit generally led by the library. In fact, policy development seemed to require deeper collaboration than service development. This relates to the suggestions above of why services have often developed before policy.

In a similar way, external collaborations can be seen as reflecting either a lack of resource or knowledge. For example, in the UK many institutions seem to have chosen to outsource the role of long-term preservation of research data, presumably because it lies outside their resource capabilities.

These logics imply that the more novel and demanding the services required (the further up the maturity model an institution has moved) the more collaboration is needed. On these grounds, one could argue that maturity implies greater collaboration; deepening collaboration could appear on the X-axis of the maturity model. Competition, particularly from commercial suppliers, presumably arises when local demand or need are not met. The data presented in this study do not suggest an urgent sense of competition; though that might arguably be considered complacent. Competition is also occurring alongside cooperation within institutions, particularly as libraries and other stakeholders move outside the traditional boundaries of their traditional jurisdictions. Balancing the competition with cooperation (“coopetition”) is a major current challenge for libraries and other actors.

Structures and skills

In the existing model “skills, roles and structures” are placed at Level 2 (Figure 10). But because these always exist even where no service is provided, it might be better to differentiate three logical levels for skills and structures:

- Level 1 – translation of skills (existing skills are reapplied to the new context of RDM).

- Level 2 – upskilling (existing staff are significantly retrained).
- Level 3 – new skills (wholly new skills need to be acquired, e.g. through recruitment).

Unfortunately, our current data set does not provide evidence in these areas, and so further research would be needed to gather evidence to test this suggestion.

As regards organisational structures we could argue that certain structures are indicative both of the level of change that has happened in the organisation and resourcing:

- Level 1 – responsibility of individual/distributed among multiple individuals in many teams – both options imply a relatively low commitment of resources and a lack of the development of concentrated expertise.
- Level 2 – a general research support/scholarly communications team has the main responsibility for research data.
- Level 3 – a new team, potentially with members from outside the library, specialises purely on research data.

The data partially support this analysis as Table IV above shows that a team dedicated to research data typically had stronger services, although there did not seem to be a major difference between the other structures and the development of services. This may reflect the difficulty of summarising organisational structures within the simple terminology of a survey, rather than that the difference does not exist.

The change process

Our previous model depicted change as a smooth “S” curve. This is, of course, schematic. The reality of significant change in many organisations is that it does not happen on a continuous basis. Rather change is often experienced as part of “punctuated equilibrium” – bursts of transformational change followed longer periods of stability, followed in turn by further rapid change (Gersick, 1991; Hayes, 2014). The reason for this is that most organisations often have inertia built into them in key “domains of organizational activity”, such as “organisational culture, strategy, structure, power distribution, and control systems”, which constrain change (Romanelli and Tushman, 1994) and so change occurs when these different factors are aligned to facilitate change. Universities in general and libraries in particular are no exception to this.

We see some evidence of this experience of change in our research. A significant burst of change in the UK, for example, occurred around 2012–2013 in response to major changes in the environment: the requirement of major funders, notably the Engineering and Physical Sciences Research Council, to develop an institutional roadmap for RDM. This resulted in significant institutional activity around governance, policy development and initial work in the development of advisory services, much of it led by libraries. Such developments appear to relate to Romanelli and Tushman’s (1994) domains of strategy, structure and control systems. Arguably, changes since then have apparently been incremental rather than transformative. Similar bursts of development seem to have occurred in other countries to a greater or lesser extent, and at an institutional level, changes requiring strategic review or organisational restructuring, are always likely to take time to effect.

The revised model

Arising from this discussion a revised model can be proposed (Figure 14).

This revised version of the model retains the concept of four levels of development. The lowest level of commitment and services is Level 0. Audits and surveys may be undertaken at this level. Level 1 is compliance and is generally characterised by a formal policy coupled with

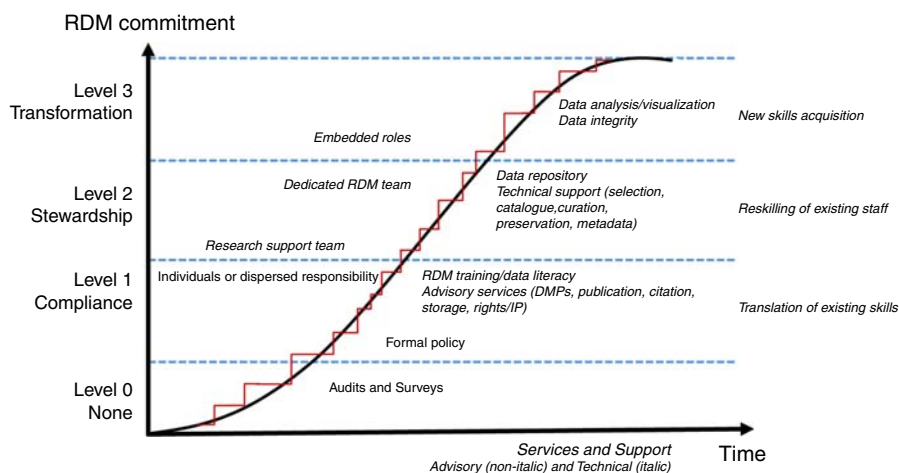


Figure 14.
Revised library RDS
maturity model

advisory services. RDM support is the responsibility of a single individual or widely dispersed; skills are primarily developed at the level of translation of existing skills. The stewardship level (Level 2) is associated with the creation of a repository and associated services. A dedicated RDM team is likely to be created. Significant reskilling may occur. At the top level (Level 4), library services are being transformed by a deep commitment to supporting high-level analytic activities. This implies new skills acquisition and novel organisational forms, including embedded roles. All of these developments are in reality likely to occur in bursts of activity (rather than as continuous change), and this is represented as a set of step changes, with the “S” curve being the general trend line (Figure 14).

Conclusion

This paper contributes to the literature on the development of RDS in academic libraries, drawing on international surveys of institutions conducted in 2014 and 2018. A revised maturity model based on the data collected has been presented, which maps levels of change in academic libraries linked to the development of RDS. For library practitioners the model can be used to reflect on how things are changing and to inform local strategy.

The data do not appear to support the notion that RDM in itself, as it is currently being pursued, is a transformative force in academic library work. Having a policy and various RDS has become common. The services provided are mostly advisory; and there is no evidence in future priorities that this will change. Correspondingly, the evidence on organisational structure does not suggest that academic libraries have significantly been reorganised to respond to RDM. This is not very surprising because libraries so far have tended to adjust by building the types of service which fit in with the existing activities and skillset of the profession, such as around advice, training and open access. The drivers fall short of an absolute and are not tied to adequate resourcing. As a result, RDM has been largely interpreted through “traditional” roles. There are some new forms of internal or external collaboration, but collaboration has become part of most academic library practice. The skills gap does not seem to have narrowed significantly. The implication for LIS education would seem to be that for now an evolutionary approach to extending the skills that students are taught remains a valid approach.

It is possible that exogenous factors could lead to a major shift in the near future, with consequences for library services. This might be a further shift in government policies in

the direction of demanding compliance. The linking of open research data into a wider agenda about open scholarship does not suggest this, however. It could be that external parties, including publishers, building new business models or introducing new technologies could prompt further change, that would impact libraries. Again, there is not yet strong evidence that this is likely.

However, when combined with other new areas of work such as around open access and open scholarship, text and data mining (TDM), bibliometrics and advice on research-related social media communication, it seems that academic libraries are facing several waves of change. Academic libraries' focus has turned to a greater emphasis on research and with a concomitant greater need to engage with issues around data such as data integrity and data analysis. Pinfield *et al.* (2017) identify a nexus of interrelated changes they dub "datafied scholarship" that together could bring about transformational change in libraries. They also note a parallel transformation through "connected learning", itself potentially tied to data via learning analytics. The current focus on (big) data both in research but much more widely in many areas of the economy and society point to significant underlying shifts that are likely to transform the context of academic libraries. Therefore, in the long run it seems highly plausible to imagine that there will be transformational effects: RDM is just one wave of this rising tide. Future research should explore this wider picture of accumulating change, by gathering data about changes in policy and service, cooperative relationships, organisational structures and staff competencies arising from data driven changes such as around TDM and bibliometrics, and emerging new impacts such as those linked to Artificial Intelligence. It will help us understand the new data oriented skills, knowledge and attitudes professionals will increasingly need and how LIS curricula can be transformed to provide them.

Note

1. The questionnaire and anonymised data are available via University of Sheffield's data repository, ORDA, at doi: 10.15131/shef.data.9204509.

References

- Akers, K.G., Sferdean, F.C., Nicholls, N.H. and Green, J.A. (2014), "Building support for research data management: biographies of eight research universities", *International Journal of Digital Curation*, Vol. 9 No. 2, pp. 171-191, available at: <http://doi.org/10.2218/ijdc.v9i2.327>
- ANDS (2018), *Creating a Data Management Framework*, Australian National Data Service, Melbourne, available at: <http://ands.org.au/guides/creating-a-data-management-framework>
- Bryant, R., Lavoie, B. and Malpas, C. (2017), *The Realities of Research Data Management: Part Two: Scoping the University RDM Service Bundle*, OCLC, Dublin, OH, available at: <http://doi.org/10.25333/C3Z039>
- Chiwere, E.R. and Becker, D.A. (2018), "Research data management services in Southern Africa: a readiness survey of academic and research libraries", *African Journal of Library and Archives and Information Science*, Vol. 28 No. 1, pp. 1-16, available at: <http://doi.org/10.1155/2012/490647>
- Corrall, S., Kennan, M. and Afzal, W. (2013), "Bibliometrics and research data management services: emerging trends in library support for research", *Library Trends*, Vol. 61 No. 3, pp. 636-674, available at: <http://doi.org/10.1353/lib.2013.0005>
- Cox, A.M. (2018), "Academic librarianship as a data profession: the familiar and unfamiliar in the data role spectrum", *eLucidate*, Vol. 15 Nos 1-2, pp. 7-10.
- Cox, A.M., Kennan, M-A, Lyon, L. and Pinfield, S. (2017), "Developments in research data management in academic libraries: towards an understanding of research data services maturity", *Journal of the Association for Information Science and Technology*, Vol. 68 No. 9, pp. 2182-2200, available at: <https://doi.org/10.1002/asi.23781>

- Cox, A.M. and Pinfield, S. (2014), "Research data management and libraries: current activities and future priorities", *Journal of Librarianship and Information Science*, Vol. 46 No. 4, pp. 299-316, available at: <http://doi.org/10.1177/0961000613492542>.Abstract
- Crowston, K. and Qin, J. (2011), "A capability maturity model for scientific data management: evidence from the literature", *Proceedings of the American Society for Information Science and Technology*, Vol. 48 No. 1, pp. 1-9.
- Dempsey, L. (2016), "Library collections in the life of the user: two directions", *LIBER Quarterly*, Vol. 26 No. 4, pp. 338-359, available at: <http://dx.doi.org/10.18352/lq.10170> (accessed 1 November 2018).
- Faniel, I.M. and Connaway, L.S. (2018), "Librarians perspectives on the factors influencing research data management programs", *College and Research Libraries*, Vol. 79 No. 1, pp. 100-119, available at: <https://doi.org/10.5860/crl.79.1.100>
- Gersick, C.J. (1991), "Revolutionary change theories: a multilevel exploration of the punctuated equilibrium paradigm", *Academy of Management Review*, Vol. 16 No. 1, pp. 10-36, available at: <http://doi.org/10.2307/258605>
- Hayes, J. (2014), *The Theory and Practice of Change Management*, 4th ed., Palgrave, London.
- Kellam, L.M. and Thompson, K. (Eds) (2016), *Databrarianship: The Academic Data Librarian in Theory and Practice*, American Library Association, Chicago, IL.
- Kenney, A.R. and McGovern, N.Y. (2003), "The five organizational stages of digital preservation", in Hodges, P., Sandler, M., Bonn, M. and Wilkin, J.P. (Eds), *Digital Libraries: A Vision for the 21st Century*, University of Michigan Scholarly Publishing Office, Ann Arbor, MI.
- Kouper, I., Fear, K., Ishida, M., Kollen, C. and Williams, S.C. (2015), "Research data services maturity in academic libraries", in Johnston, L. (Ed.), *Curating Research Data*, Association of College and Research Libraries, Chicago, IL, pp. 153-170.
- Peng, G., Privette, J.L., Kearns, E.J., Ritchey, N.A. and Ansari, S. (2015), "A unified framework for measuring stewardship practices applied to digital environmental datasets", *Data Science Journal*, Vol. 13, February, pp. 231-253, available at: <https://doi.org/10.2481/dsj.14-049>
- Pickard, A. (2012), *Research Methods in Information*, Facet, London.
- Pinfield, S., Cox, A. and Rutter, S. (2017), "Mapping the future of academic libraries: a report for SCONUL", available at: <https://sconul.ac.uk/publication/mapping-the-future-of-academic-libraries> (accessed 1 August 2019).
- Pöppelbuß, J. and Röglinger, M. (2011), "What makes a useful maturity model? A framework of general design principles for maturity models and its demonstration in business process management", *19th European Conference on Information Systems, Helsinki, June*, available at: <http://fim-rc.de/Paperbibliothek/Veroeffentlich/327/wi-327.pdf>
- Proença, D., Vieira, R. and Borbinha, J. (2016), "A maturity model for information governance", in Fuhr, N., Kovacs, L., Risse, T. and Nedjl, W. (Eds), *International Conference on Theory and Practice of Digital Libraries*, Springer, Cham, pp. 27-3815-26.
- Pryor, G., Jones, S. and Whyte, A. (Eds) (2014), *Delivering Research Data Management Services: Fundamentals of Good Practice*, Facet, London.
- Qin, J., Crowston, K. and Kirkland, A. (2017), "Pursuing best performance in research data management by using the capability maturity model and rubrics", *Journal of eScience Librarianship*, Vol. 6 No. 2, available at: <https://doi.org/10.7191/jeslib.2017.1113>
- Rebouillat, V. (2017), "Inventory of research data management services in France2", in Chan, L. and Loizides, F. (Eds), *Expanding Perspectives on Open Science: Communities, Cultures and Diversity in Concepts and Practices, Proceedings of the 21st International Conference on Electronic Publishing*, IOS Press, Amsterdam, pp. 174-181.
- Rice, R. and Southall, J. (2016), *The Data Librarian's Handbook*, Facet, London.
- Romanelli, E. and Tushman, M.L. (1994), "Organizational transformation as punctuated equilibrium: an empirical test", *Academy of Management Journal*, Vol. 37 No. 5, pp. 1141-1166.

- Russell Group (2019), "Russell group", available at: <https://russellgroup.ac.uk/>
- SurveyMonkey (n.d.), Available at: www.surveymonkey.com/
- Tenopir, C., Birch, B. and Allard, S. (2012), *Academic Libraries and Research Data Services: Current Practices and Plans for the Future: An ACRL White Paper*, Association of College and Research Libraries, Chicago, IL, available at: www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf
- Tenopir, C., Hughes, D., Allard, S., Frame, M., Birch, B., Baird, L., Sandusky, R. and Lundeen, A. (2015), "Research data services in academic libraries: data intensive roles for the future?", *Journal of eScience Librarianship*, Vol. 4 No. 2, available at: <http://dx.doi.org/10.7191/jeslib.2015.1085>
- Tenopir, C., Pollock, D., Allard, S. and Hughes, D. (2016), "Research data services in European and North American libraries: current offerings and plans for the future", *Proceedings of the Association for Information Science and Technology*, Vol. 53 No. 1, pp. 1-6, available at: <https://doi.org/10.1002/pra2.2016.14505301129>
- Tenopir, C., Talja, S., Horstmann, W., Late, E., Hughes, D., Pollock, D., Schmidt, B., Baird, L., Sandusky, R.J. and Allard, S. (2017), "Research data services in European academic research libraries", *LIBER Quarterly*, Vol. 27 No. 1, pp. 23-44, available at: <http://doi.org/10.18352/lq.10180>

Further reading

- Borgman, C.L. (2015), *Big Data, Little Data, no Data: Scholarship in the Networked World*, MIT Press, Cambridge, MA.
- Kenney, A.R. and McGovern, N.Y. (2005), "The three-legged stool: institutional response to digital preservation", 2nd Convocatoria del Coloquio de Marzo, available at: www.library.cornell.edu/iris/dpo/docs/Cuba-ark-nym_final.ppt (accessed 1 August 2019).
- Si, L., Xing, W. and Zhuang, X. (2015), "Investigation and analysis of research data services in university libraries", *The Electronic Library*, Vol. 33 No. 3, pp. 417-449, available at: <http://doi.org/10.1108/EL-07-2013-0130>

Appendix 1

Table AI.
Audit, survey or evaluation of the institution's data (data supporting Figure 1)

	Already conducted		Planned		Not planned	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Australia (<i>n</i> = 33)	15	45	3	9	15	45
Canada (<i>n</i> = 24)	7	29	7	29	10	42
Germany (<i>n</i> = 23)	12	52	5	22	6	26
Ireland (<i>n</i> = 11)	4	36	4	36	3	27
The Netherlands (<i>n</i> = 5)	2	40	0	0	3	60
New Zealand (<i>n</i> = 8)	4	50	1	13	3	38
UK (<i>n</i> = 78)	45	58	13	17	20	26
USA (<i>n</i> = 23)	11	48	3	13	9	39
Total responses (<i>n</i> = 205)	100	49	36	18	69	34

Appendix 2

	Already conducted		Not planned	
	<i>n</i>	%	<i>n</i>	%
Australia (<i>n</i> = 32)	11	34	21	66
Canada (<i>n</i> = 23)	9	39	14	61
Germany (<i>n</i> = 22)	10	45	12	55
Ireland (<i>n</i> = 11)	5	45	6	55
The Netherlands (<i>n</i> = 5)	4	80	1	20
New Zealand (<i>n</i> = 8)	4	50	4	50
UK (<i>n</i> = 78)	33	42	45	58
USA (<i>n</i> = 23)	11	48	12	52
Total responses (<i>n</i> = 202)	87	43	115	57

Table AII.
Frequency of having
undertaken a survey
of faculty/academic
staff attitudes to RDM
(data supporting
Figure 2)

Appendix 3

	Have a policy now		Planned in next year		Planning but more than year		Not planned		Do not know	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Australia (<i>n</i> = 34)	27	79	3	9	4	12	0	0	0	0
Canada (<i>n</i> = 24)	0	0	7	29	10	42	7	29	0	0
Germany (<i>n</i> = 23)	12	52	8	35	2	9	1	4	0	0
Ireland (<i>n</i> = 11)	4	36	3	27	2	18	2	18	0	0
The Netherlands (<i>n</i> = 6)	3	50	1	17	1	17	1	17	0	0
New Zealand (<i>n</i> = 8)	0	0	1	13	6	75	1	13	0	0
UK (<i>n</i> = 80)	60	75	4	5	7	9	6	8	3	4
USA (<i>n</i> = 23)	11	48	2	9	4	17	6	26	0	0
Total responses (<i>n</i> = 209)	117	56	29	14	36	17	24	11	3	1

Table AIII.
Formal RDM policy in
place or planned (data
supporting Figure 3)

Appendix 4

1458

Table AIV.
A comparison of RDS between 2014 and 2018 and A comparison of RDS priorities 2014 and 2018 (data supporting Figures 4 and 5)

	Current services		Future priorities	
	2018	2014	2018	2014
Advice on DMPs	1.31	0.87	1.66	1.65
Training	1.36	0.81	1.67	1.61
Website	1.36	0.96	1.48	1.59
Advice on awareness of reusable sources	1.14	0.84	1.15	1.29
Advice on retrieval of reusable sources	1.04	0.83	1.03	1.01
Citation advice	1.10	0.84	1.13	1.41
Data publication advice	1.16	0.73	1.34	1.40
Data storage advice	1.12	0.77	1.39	1.43
Data analysis, mining and visualisation advice	0.38	0.29	0.64	0.78
Advice on Copyright and IPR	1.29	0.94	1.36	1.54
Active data curation advice	0.96	0.55	1.21	1.35
Preservation advice	0.87	0.48	1.16	1.20
Run a data repository	1.07	0.69	1.58	1.56
Metadata creation	0.63	0.40	0.98	1.29
Data catalogue	0.70	0.35	1.32	1.44
Data selection	0.53	0.29	0.96	1.08
Data preparation	0.70	0.38	1.05	1.19
Active data curation	0.41	0.27	0.78	0.96
Long-term preservation	0.61	0.34	1.29	1.20
Data quality checking	0.31		0.63	
Data analysis	0.33		0.57	
Support research integrity	0.51		1.06	
Data rescue	0.23		0.39	
Embedded librarian	0.33	0.49	0.61	0.75

Appendix 5

Table AV.
Library involvement in research data policy development (data supporting Figure 6)

	Led		Participated		Did not participate		Do not know	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Australia (<i>n</i> = 31)	12	39	16	52	2	6	1	3
Canada (<i>n</i> = 16)	8	50	4	25	2	13	2	13
Germany (<i>n</i> = 22)	15	68	7	32	0	0	0	0
Ireland (<i>n</i> = 8)	3	38	3	38	1	13	1	13
The Netherlands (<i>n</i> = 5)	4	80	1	20	0	0	0	0
New Zealand (<i>n</i> = 5)	2	40	3	60	0	0	0	0
UK (<i>n</i> = 73)	42	58	24	33	5	7	2	3
USA (<i>n</i> = 19)	9	47	7	37	2	11	1	5
Total responses (<i>n</i> = 179)	95	53	65	36	12	7	7	4

Appendix 6

	Led		Participated		Did not participate		Do not know	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Australia (<i>n</i> = 32)	22	69	10	31	0	0	0	0
Canada (<i>n</i> = 16)	7	44	5	31	2	13	2	13
Germany (<i>n</i> = 18)	2	11	15	83	0	0	1	6
Ireland (<i>n</i> = 8)	6	75	1	13	0	0	1	13
The Netherlands (<i>n</i> = 5)	2	40	3	60	0	0	0	0
New Zealand (<i>n</i> = 4)	0	0	4	100	0	0	0	0
UK (<i>n</i> = 66)	32	48	32	48	1	2	1	2
USA (<i>n</i> = 19)	7	37	7	37	0	0	5	26
Total responses (<i>n</i> = 168)	78	46	77	46	3	2	10	6

Table AVI.
Research office
involvement in
research data policy
development (data
supporting Figure 7)

Appendix 7

	Led		Participated		Did not participate		Do not know	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Australia (<i>n</i> = 31)	17	55	14	45	0	0	0	0
Canada (<i>n</i> = 22)	16	73	3	14	2	9	1	5
Germany (<i>n</i> = 22)	17	77	5	23	0	0	0	0
Ireland (<i>n</i> = 11)	8	73	2	18	1	9	0	0
The Netherlands (<i>n</i> = 5)	4	80	1	20	0	0	0	0
New Zealand (<i>n</i> = 7)	3	43	4	57	0	0	0	0
UK (<i>n</i> = 70)	49	70	16	23	3	4	2	3
USA (<i>n</i> = 22)	20	91	2	9	0	0	0	0
Total responses (<i>n</i> = 190)	134	71	47	25	6	3	3	2

Table AVII.
Library role in
research data services
(data supporting
Figure 8)

Appendix 8

	Led		Participated		Did not participate		Do not know	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Australia (<i>n</i> = 31)	16	52	12	39	2	6	1	3
Canada (<i>n</i> = 20)	5	25	12	60	1	5	2	10
Germany (<i>n</i> = 18)	3	17	14	78	1	6	0	0
Ireland (<i>n</i> = 11)	3	27	8	73	0	0	0	0
The Netherlands (<i>n</i> = 5)	1	20	3	60	1	20	0	0
New Zealand (<i>n</i> = 6)	0	0	4	67	2	33	0	0
UK (<i>n</i> = 67)	26	39	37	55	2	3	2	3
USA (<i>n</i> = 20)	4	20	13	65	2	10	1	5
Total responses (<i>n</i> = 178)	58	33	103	58	11	6	6	3

Table AVIII.
Research office role in
research data services
(data supporting
Figure 9)

Appendix 9

1460

Table AIX.
A comparison of
current and planned
external collaborations
(data supporting
Table II)

	Yes		No (planned)		No (not planned)		Total
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	
<i>2018</i>							
Australia	16	50	5	16	11	34	32
Canada	12	50	3	13	9	38	24
Germany	10	43	8	35	5	22	23
Ireland	2	18	3	27	6	55	11
The Netherlands	3	60	1	20	1	20	5
The New Zealand	3	38	2	25	3	38	8
UK	36	47	18	23	23	30	77
USA	12	52	2	9	9	39	23
Total	94	46	42	21	67	33	203
<i>2014</i>							
Australia	26	76	3	9	5	15	34
Canada	10	67	0	0	5	33	15
Germany	3	60	1	20	1	20	5
Ireland	1	14	1	14	5	71	7
The Netherlands	9	82	2	18	0	0	11
New Zealand	1	14	1	14	5	71	7
UK	18	22	24	29	41	49	83
Total	68	42	32	20	62	38	162

	A single individual is responsible		A team with a specific focus on research data is responsible		A team with a general remit for research support is responsible		It is spread across multiple teams		Other answer	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Australia (<i>n</i> = 39)	7	18	7	18	12	31	10	26	3	8
Canada (<i>n</i> = 28)	3	11	8	29	3	11	4	14	10	36
Germany (<i>n</i> = 26)	4	15	11	42	4	15	3	12	4	15
Ireland (<i>n</i> = 15)	6	40	2	13	3	20	2	13	2	13
The Netherlands (<i>n</i> = 5)	1	20	4	80	0	0	0	0	0	0
New Zealand (<i>n</i> = 10)	2	20	1	10	2	20	1	10	4	20
UK (<i>n</i> = 93)	23	25	16	17	20	22	9	10	25	14
USA (<i>n</i> = 35)	11	31	9	26	4	11	9	26	2	6
Total responses (<i>n</i> = 251)	57	23	58	23	48	19	38	15	50	20

Note: Some people ticked more than one answer

Table AX.
Organisational
structures for RDS
(data supporting
Table III)

1462

Table AXI.
Comparison of UK
russell group and non-
russell group
universities' RDS and
comparison of russell
group and non-russell
group universities'
plans for RDS (data
supporting Figures 12
and 13)

	Current services		Future plans	
	Non-Russell Group	Russell Group	Non-Russell Group	Russell Group
Advice on DMPs	1.15	2.06	2.57	2.69
Training	1.27	1.94	2.59	2.73
Website	1.07	2.00	2.46	2.53
Advice on awareness of reusable sources	0.91	1.35	2.02	2.19
Advice on retrieval of reusable sources	0.79	1.00	1.83	2.20
Citation advice	0.85	1.47	2.02	2.00
Data publication advice	0.95	1.56	2.22	2.36
Data storage advice	1.00	1.50	2.39	2.56
Advice on data analysis	0.15	0.22	1.47	1.38
Advice on data mining	0.11	0.46	1.40	1.71
Advice on data visualisation	0.08	0.60	1.40	1.50
Advice on Copyright and IPR	1.21	1.47	2.33	2.40
Active data curation advice	0.86	1.07	2.15	2.31
Preservation advice	0.78	0.94	2.13	2.20
Run a data repository	1.13	1.40	2.55	2.71
Metadata creation	0.49	0.76	1.98	2.00
Data catalogue	0.82	1.12	2.44	2.63
Data selection	0.48	0.57	2.00	2.00
Data preparation	0.60	0.64	2.09	2.08
Active data curation	0.37	0.31	1.77	1.80
Long-term preservation	0.57	0.53	2.26	2.63
Data quality checking	0.24	0.50	1.84	1.78
Data analysis	0.06	0.30	1.26	1.50
Support research integrity	0.33	0.77	1.91	2.25
Data rescue	0.10	0.23	1.29	1.55
Embedded librarian	0.12	0.12	1.45	1.25

Corresponding author

Andrew M. Cox can be contacted at: a.m.cox@sheffield.ac.uk

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgroupublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com