# ACADEMIC LIBRARIES AND RESEARCH DATA MANAGEMENT: A SYSTEMATIC REVIEW

# UPRAVLJANJE AKADEMSKIM KNJIŽNICAMA I ISTRAŽIVAČKIM PODACIMA: PREGLED

Juan-José Boté-Vericad Facultat d'Informació i Mitjans Audiovisuals & Centre de Recerca en Informació, Comunicació i Cultura, Universitat de Barcelona juanjo.botev@ub.edu

Sharon Healy Maynooth University, Ireland schealy.ire@gmail.com

> UDK / UDC: 001.103:027.7: 004.451.5 Overview paper / Pregledni rad https://doi.org/10.30754/vbh.65.3.1016 Received / Primljeno: 15. 8. 2022. Accepted / Prihvaćeno: 12. 10. 2022.

#### Abstract

**Purpose.** Open Science entails research reproducibility, with an emphasis on data sharing and reuse. Hence, research data management (RDM) is an essential asset in research institutions for supporting open science. This study offers a systematic review of the landscape of research data management in academic libraries. It further examines the influence academic libraries can have if they are involved in the research lifecycle process, and how this benefits research institutions that have started implementing research data management, especially in the data-intensive disciplines.

**Methodology.** In this study, the authors analysed Web of Science and Scopus databases, searching for papers connecting research data management and academic libraries. The authors found a total of N=387 articles. After removing duplicates and applying the exclusion and inclusion criteria process, the authors finally analysed N=32 articles, n=20 case studies, and n=12 research papers at both national and international levels.

**Limitations.** This study has some limitations. Although the authors retrieved as many papers as possible for the analysis, it should not be considered as an exhaustive analysis, as varying studies may also be missing from the sample. The authors observed that there are more case studies focused on one institution rather than research papers involving different institutions at the international or national level. Therefore, more research studies would enrich the literature and show best practices in RDM.

**Results.** The results show that research data management has some services implemented in different countries at the local or the international level. The authors argue that research data management generates new opportunities for academic libraries and librarians to acquire new skills as a part of the research data lifecycle.

**Originality/value.** This study reports the current state of research data management at the international level in academic libraries and the influence libraries can have if they are involved in the research lifecycle process.

**Keywords:** Academic libraries; librarian skills; open science; research data management; systematic review

#### Sažetak

Cilj. Otvorena znanost bavi se ponovljivošću istraživanja, s naglaskom na dijeljenje i ponovnu upotrebu podataka. Stoga je upravljanje istraživačkim podacima ključno sredstvo u istraživačkim institucijama koje podržavaju otvorenu znanost. Ova studija nastoji pregledno izložiti kako se upravlja istraživačkim podacima u visokoškolskim knjižnicama. Nadalje ispituje utjecaj koji akademske knjižnice mogu imati ako su uključene u proces životnog ciklusa istraživanja i kako to koristi onim istraživačkim institucijama koje su počele implementirati upravljanje istraživačkim podacima, posebno u onim disciplinama koje se na njih značajno oslanjaju.

**Metodologija**. U ovom radu autori su analizirali baze podataka Web of Science i Scopus, tražeći radove koji povezuju upravljanje istraživačkim podacima (RDM) i akademske knjižnice. Autori su pronašli ukupno 387 članaka. Nakon uklanjanja duplikata i primjene postupka isključenja i uključivanja odabranih kriterija, autori su konačno analizirali 32 članka, 20 studija slučaja i 12 istraživačkih radova na nacionalnoj i međunarodnoj razini.

**Ograničenja**. Ova studija ima određenih ograničenja. Iako su autori dohvatili što je moguće više radova za analizu, ne treba je smatrati iscrpnom analizom jer različite studije mogu nedostajati u analizi. Autori su uočili da postoji više studija slučaja usmjerenih na jednu instituciju nego istraživačkih radova koji uključuju različite institucije na međunarodnoj ili nacionalnoj razini. Stoga bi više istraživačkih studija obogatilo literaturu i pokazalo najbolje prakse u RDM-u.

**Rezultati**. Rezultati pokazuju da upravljanje pojedinim istraživačkim podacima dovodi do određenih usluga implementiranih u različitim zemljama na lokalnoj ili međunarodnoj razini. Autori tvrde da upravljanje istraživačkim podacima stvara nove prilike za akademske knjižnice i knjižničare za stjecanje novih vještina kao dio životnog ciklusa istraživačkih podataka.

**Izvornost/vrijednost**. Ova studija izvještava o trenutnom stanju upravljanja istraživačkim podacima na međunarodnoj razini u akademskim knjižnicama i utjecaju koji knjižnice mogu imati ako su uključene u proces životnog ciklusa istraživanja.

Ključne riječi: knjižničarske kompetencije; otvorena znanost; sistemski pregled; sveučilišne knjižnice; upravljanje istraživačkim podacima

# 1. Introduction

Academic libraries may gain more relevance and influence in the following years because of Open Science. Adopted by the General Conference of UNES-CO in 2021, the UNESCO (2021) report "UNESCO Recommendation on Open Science", recommends that State members foster Open Science and collaborative initiatives. The report embraces Open Science for establishing:

"a new paradigm that integrates into the scientific enterprise practices of reproducibility, transparency, exchange and collaboration resulting from the greater openness of the contents, the scientific tools and processes." (cf. ibid.: 7)

Thus, Open Science supports an ethos of free access to scientific publications, open research data, open educational resources, open-source software, and open-source computer equipment. The UNESCO (2021) recommendation was also agreed upon by the International Federation of Libraries Association, reporting implications for libraries such as the role of librarians in achieving Open Science and the relevance of the institutional repositories where the library discipline has long expertise (IFLA, 2021).

In the case of Europe, the European Commission reported that Open Science is a priority policy and a "system change allowing for better science through open and collaborative ways of sharing knowledge and data" (European Commission, 2019). Open Science implies new ways to gather, analyze, manage, or store research data. Abadal (2021) emphasises that Open Science allows transparency, effectivity, and reproducibility because the access to open data and research has a more social impact on citizens. This question has a diversity of implications. For instance, academic libraries need to implement policies and strategies to have data available for the researchers' institutions and to provide access to researchers from other institutions. Then, sharing research data requires procedures to create, disseminate, reuse, store and preserve the research data. Therefore, it is necessary to have strategies for research data management (RDM) in research institutions. RDM has acquired relevance across multiple disciplines in the past few years as a part of Open Science.

#### 1.1 Research data management

RDM is a set of activities involving the research data lifecycle. These activities include topics such as planning, creation, use, organisation, preservation, or reuse (Ashiq et al., 2020). Research libraries are supposed to provide RDM services as a part of the regular services (Cox et al., 2019). Nevertheless, there are several implications behind RDM. For instance, having an adequate technology structure or librarians having enough skills and knowledge to support researchers.

This paper is a systematic literature review of the landscape of RDM in line with academic libraries to identify the roles of academic libraries in supporting researchers with RDM.

The following research questions guided this research:

RQ1: What is the role of academic libraries in research data management?

RQ2: How are academic libraries promoting RDM across their institutions?

## 2. Literature Review

# 2.1. Research Data Management and Open Science

The relationship between Open Science and RDM is the storage and preservation of research data, so that it may be used to reproduce results to ensure research integrity, but also to offer the potential for the preserved data to be reused. More increasingly, grants and funders require data management plans (DMP) that will make research data open when possible. This has resulted in the development of data repositories and RDM strategies by universities and research institutions to ensure this requirement. For instance, Japan has started developing computing structures to keep and discover research data (Komiyama and Yamaji, 2017). In Scotland, librarians have started training their researchers in data management plans, evaluating RDM costs, or ensuring that sensitive data is kept safely where the storage costs increase because of the security for protecting data (Rice, 2019). In regard to training, the International Federation of Libraries Association (IFLA) has published a report "*IFLA Guidelines for Professional Library and Information Science (LIS) Education Programmes*" (Chu et al., 2022) emphasizing in 8 foundational knowledge areas (FKA). Three of these areas are especially involved in RDM: FKA4-research and innovation, FKA7-information needs and user services, and FKA8-literacies and learning.

RDM practices must guarantee the recoverability of the resources of any project (Lefebvre and Spruit, 2021). This entails the concept of Findable, Accessible, Interoperable and Reusable (FAIR) principles being applied to research data. The FAIR principles imply "that all research objects should be Findable, Accessible, Interoperable and Reusable (FAIR) both for machines and for people" (Wilkinson et al., 2016: 3). However, FAIR principles present technical and ethical challenges in different countries when it comes to access and the reusability of research data (Boté and Térmens, 2019), which is required to improve the reliability of the research, increasing the impact of the research (Shelly and Jackson, 2018). Legal issues are also presented as a challenge for reusability. For instance, in Sweden researchers may deposit data in digital archives and they cannot reuse them because of the Personal Data Act (Slavnic, 2017).

For proper data management and compliance with FAIR principles it is necessary to have advanced knowledge of information management. This includes creating a repository, developing good practices that allow data management, and promoting and facilitating open publishing. Training in open science and data management is necessary for faculty members and research librarians when faculty members are obliged to deal with DMP and data sharing or data reuse. A DPM allows for reflections on the realistic requirement of how a researcher will collect, share and preserve research data (Mushi et al., 2020) and in the case of librarians providing support in the proper management of research data. For example, the reusability of research data derived from web archives is often problematic due to copyright, data protection and privacy laws, which vary from country to country. This further affects the sharing and reuse of research data across borders (Truter, 2021, Healy et al. 2022; WARCnet, n.d.).

# 2.2. Projects related to research data management

There are different projects in scientific literature that provide informal training in RDM based on toolkits and the research lifecycle. One example is the Learn Project founded by Horizon 2020 with a consortium of the following institutions: University College London, Universitat de Barcelona, Stichting Liber, Universität Wien and the United Nations Economic Commission for Latin America and the Caribbean (Ayris, 2017; Traub et al., 2016). This toolkit is based on 23 case studies and 20 RDM best practices. Another toolkit is the Research Data Management Teaching toolkit (Read and Surkis, 2018), a one-hour introductory data management class for biomedical researchers. There are other toolkits, such as the RDM toolkit from the Joint Information Systems Committee (JISC), an interactive toolkit to support researchers during the research lifecycle. It covers topics such as collecting and capturing or reusing data (JISC, 2021). Finally, the policies and procedures openly published by the UK Data Archive, is a useful resource (UK Data Archive, 2022).

# 3. Methodology

The authors used the bibliographic databases Scopus and Web of Science to perform this systematic review. The reason to use these databases is that they concentrate on the major corpus in scientific literature worldwide. Studies were limited to articles published up to the end of 2021. The exclusion criteria were book chapters, editorials, and guest editorials. The inclusion criteria encompassed research articles and case studies. The authors restricted the languages of studies to what they could read and understand. Therefore, this study includes studies in the following languages: French, English, German, Portuguese, and Spanish.

After several tests, the authors decided to use these search key phrases.

**Scopus:** TITLE-ABS-KEY (library\* "research data management") AND (EXCLUDE (PUBYEAR, 2022)) AND (LIMIT-TO (DOCTYPE, "ar")) AND (LIMIT-TO (PUBSTAGE, "final")) AND (LIMIT-TO (SRCTYPE, "j"))

**Web of Science:** librar\* "research data management" (Topic) and 2022 (Exclude – Publication Years) and Article (Document Types)

The authors applied the PRISMA model to select the final sample of articles (Page et al., 2021), as shown in Fig. 1. The authors performed the search equations and documented (N=193) records from Scopus and (N=194) records from Web of Science. This resulted in a total of (N=387) records from both databases. After joining both databases, (N=108) duplicated articles were removed, obtaining a sample of (N=279) articles. Thereafter, the authors read all titles and abstracts, and in case of doubts, the authors read the complete papers. This resulted in the exclusion of (N=209) articles. The reasons to exclude papers were varied, such as only covering the librarian's role or minor aspects related to research data. In the end, a total of (N=37) articles were removed, leaving a final sample of (N=32) articles. The final sample comprises (n=20) case study articles and (n=12) research papers.

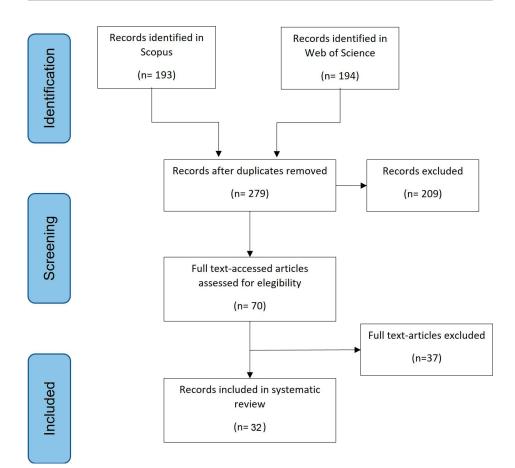


Fig. 1 PRISMA model for the systematic review.

# 4. Results

The results show different approaches to how academic libraries introduce RDM in their institutions. Table 1 offers the description of the sample and shows all included studies with the main topics covered in case studies and the emerging topics in the case of research studies. In this classification, case studies are centred in one institution or partnering projects among institutions in one country, and research papers focus on studies in one country with several institutions involved or international studies with participating institutions from different countries. In the case studies, the main topics covered are the promotion of RDM performed by the academic libraries, the review of the computing structure to manage research data, training addressed to librarians and researchers and finally, the relation with

the stakeholders of the academic library or with stakeholders of their institutions in case of consortia or partnerships. In the case of research papers, there are (n=5) performing quantitative methodologies, such as questionnaires, and (n=7) qualitative methodologies such as semi-structured interviews or focus groups. From all selected final papers, (n=8, 24.24%) articles reported a definition of RDM, arguing that it is a set of services along the research data lifecycle. Among these services are data management planning, digital curation, digital preservation, and data sharing.

Type of article	Works	Main covered topics
Case Study	(Ball, 2013; Bardyn et al., 2018; Blumesberger, 2020; Bradley-Ridout, 2018; Chiware & Mathe, 2015; Clem- ents, 2013; Hansen et al., 2021; Hen- derson & Knott, 2015; Kim and Syn, 2021; Kruse and Thestrup, 2014; Lind- städt and Schmitz, 2019; Mushi et al., 2020; Nie et al., 2021; Read et al., 2019; Rehwald and Stegemann, 2021; Searle, 2015; Shipman and Tang, 2019; Steiner, 2015; Strauch, 2019, 2020)	<ul> <li>Technology structure</li> <li>Promotion of RDM by academic libraries</li> <li>Relationship with stakeholders</li> <li>Training</li> </ul>
Research papers	(Bishop et al., 2021; Corrall et al., 2013; Cox et al., 2016, 2017; Cox and Pinfield, 2014; Faniel and Connaway, 2018; Hu- ang et al., 2021; Pinfield et al., 2014; Shelly and Jackson, 2018; Tenopir et al., 2014; Wu and Chen, 2016; Yoon and Sc- hultz, 2017)	<ul> <li>Data curation</li> <li>Data management plans</li> <li>Data storage</li> <li>RDM costs</li> <li>RDM development</li> <li>Training</li> </ul>

Table 1. Description of the sample papers

Figure 2 offers an overview of the methodological features of the sample, and how this is distributed by country. The implementation review process is the methodology most found to be performed in the sample of papers, and deals with how an institution has performed an RDM strategy or delivered RDM services to its communities. In addition to the implementation review process, semi-structured interviews are followed by a quantitative approach using questionnaires. Also, Mixed-Methods has been used to explore academic libraries in different studies such as cross-sectional studies (Bardyn et al. 2018; Huang et al., 2021). Other methodologies were less used such as website content analy-

sis (Shelly and Jackson, 2018; Yoon and Schultz, 2017) or document analysis (Bradley-Ridout, 2018).

In the following sections, the analysis of the sample is organised around five main topics as follows: (i) training in research data management; (ii) technology structures; (iii) data management plans; (iv) relationships with stakeholders; (v) and the promotion of RDM by academic libraries.

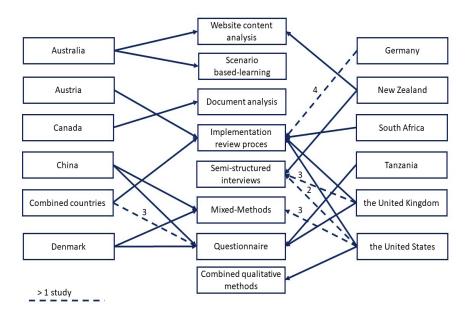


Figure 2. Features of the sample

#### 4.1. Training in research data management

Training seems to be a common point in the sample of studied papers. While training is necessary when a new topic is introduced to the library functions, it also presents a barrier to faculty members who are novices in having to deal with the topic. Training in RDM allows for the coverage of staff training needs, in this case research librarians, and allows them to acquire new skills. These new skills enable them to offer services in relation to data sharing, data reuse, DMPs and practical tools, and to provide support and assistance in RDM. The literature indicates that RDM training has been offered to staff and faculty members, as well as being offered prior to performing any actions in relation to RDM.

With regard to having a lack of skills, a study by Faniel and Connaway (2018) examined RDM experiences of 36 academic library professionals in the United

States (US) through the collection of qualitative data, in the form of 10 semi-structured interviews by telephone, and 3 focus groups with 26 participants. They further identify factors of influence, such as technical resources or researchers' perceptions about the library. In their findings, 31% indicated that they had no experience in RDM, and 46% had some experience. However, librarians observed this situation as an opportunity to promote the value of the library to their campus communities. They concluded that librarians had opportunities to be more efficient in sharing support through the research lifecycle.

With respect to training offered, Corall et al. (2013) performed a study with a total of (N=140) participants from academic libraries in four countries: Australia, Ireland, New Zealand, and the United Kingdom (UK). Their research was around the planned services, target audiences, service constraints and staff training needs. They found that the majority of respondents offered bibliometrics training (77.1% in Australia, the highest percentage), citation reports (88.9% in Ireland) and calculation of research impact (77.8% in Ireland). Concerning RDM, the levels were lower than for bibliometrics.

Nevertheless, having training in RDM provides research librarians with an opportunity to lead the change in their institutions. Moreover, librarians acquire RDM training from different places such as conferences or in-house training. This is the case of the study of Tenopir et al. (2014) who explored the RDM practises of librarians in US and Canadian academic research libraries, as well as the policies of such libraries offering research data services. They used a stratified random sample of library directors (N=351) with a response rate of 63% (N=221). Regarding their findings about training, 47% of librarians felt they had the opportunity to develop one type of RDS skills through attendance at conferences (62%), courses (53%) or training at the library (32%).

At the same time, librarians also participate in training in other higher education institutions to acquire RDM skills. In China, Nie et al. (2021) conducted a study of the Peking University Library (PKUL). In relation to skills training, the PKUL offered workshops to teach students how to identify data sources. Also, librarians participated in training offered by other universities to improve RDM skills. They concluded some critical points when implementing an RDM project including: needs assessment and collaboration, software must cover sustainability and stability criteria, communication among administrative and academic units is critical, and data curation programs need to be catered for librarians, researchers and students in a wide variety of formats. Similarly, in South Africa, the Cape Peninsula University of Technologies libraries developed a training plan for librarians during a three-year period to enable librarians to participate in RDM. The training plan covered a diversity of areas such as knowledge of relevant tools, data management planning, research methodologies or intellectual property rights (Chiware and Mather, 2015). Therefore, skills training needs to be part of the process to improve professional performance. For example, Cox et al. (2016) performed a study exploring the librarians' value in data management. They conducted 26-semistructured interviews with librarians directly responsible for RDM in their libraries. The study was confronted against 16 wicked problems found in the literature. In their findings, key issues were that RDM still had little information about some of its key aspects and was seen as a large-scale problem and more complex than most information problems. They concluded that the skills and attitudes of information professionals need to include the attributes in their training to help deal with wicked problems.

Likewise, Read et al. (2019) also reported a model for initiating RDM at six academic libraries from different institutions in the US. The participants from these six libraries had some of the following motivations: enhancing the visibility of the library with research data, gaining a better understanding of the research done in their corresponding institutions, and expanding their knowledge related to institutional needs around data management. The six participants' libraries were from University at Buffalo, the University of Delaware, Drexel University, Duquesne University, Stony Brook University, and Temple University. A pilot training program with eight educational modules was developed for librarian participants from these institutions. In these modules, participants were required to use two of three components: (1) a template and strategies for data interviews; (2) a teaching toolkit to teach an introductory RDM class; or (3) strategies for hosting a data class series. They concluded that it was necessary to build communities to support librarians in the process to update their skills in RDM (Read et al., 2019).

Similarly, Yoon and Schultz (2017) explored 185 academic library websites that promoted RDM services in the US. They performed a content analysis on libraries' websites to explore RDM services in academic libraries. They found that only 21.8% provided training on data sharing and data reuse within data management training. In another study, it was found that RDM support was not included in research services. For instance, in the US, an interview study was conducted with (N=10) librarians and (N=12) research integrity officers (RIO) recruited from private and public higher education institutions. Their goal was to analyse their job roles in their corresponding institutions concerning research data services. They found some common gaps in RDM across these institutions. In the case of RIO, researchers appeared to manage their own RDM needs after initial training. Contrarily, librarians could complement emerging practices and collaborate with supporting faculty, students, and staff. Another gap was that research services from other units were not embedded into research centres. Nevertheless, a common point among both groups was the need for better supporting cyberinfrastructure, the lack of adequate budget and staffing, and the need for additional training (Bishop et al., 2021).

There are also studies which report on how research libraries provide support and training to their communities. For instance, Strauch (2019, 2020) discusses how to provide practical support in RDM to researchers and students through the university library, at the Universität Hildesheim in Germany. The library supports the finding of personalised (ad-hoc) solutions or individual solutions for their research institutes and researchers. Moreover, the library offers training to researchers and students based on seminars, coffee lectures, or customised workshops. Through seminars on RDM, the library provides information and support, covering topics such as the types of data generated, where data comes from, how to manage the volume of the research data, the best infrastructure to use, or how to reuse the research data.

Similarly, Rehwald and Stegemann (2021) reported their experience at the Universität Duisburg-Essen (Germany). They explained the services offered to their research community. A typical central point was established having a network connection with all stakeholders with implications in RDM. They developed a road map of 4 stages. An initial stage was in analyzing their users' needs, a start-up process and development of the infrastructure and final deployment. They also reported the tasks of central point support for RDM. These services provide support and assistance in RDM such as training and seminars to researchers; assessment in a research project such as the creation of a data management plan, the use of specific tools by disciplines such as digital humanities, storage of structured data such as the use of a research data repository and research data publication such as the use of persistent identifiers.

# 4.2. Technology structure

The technology structure is an important compound in RDM. Questions such as data curation, data storage, data reuse, data sharing, interoperability or digital preservation are essential to offer RDM services in higher education institutions. However, without a technology structure and trained staff, there is a limit to what RDM services can be offered. This question affects both developing countries as well as developed countries. For example, in China, Huang et al. (2021) explored the status of Chinese university libraries and levels of development in RDM. Their research was based on the previous work of Cox et al. (2017, 2019) consisting of website analysis, university policies and services, an online questionnaire, and semi-structured interviews. They found that RDM in academic libraries in China was still in its infancy. Since some services were implemented, there were still some improvements needed. They found that the skills needed in academic libraries were research methods and data curation knowledge. They also found that a challenge was sharing data. They concluded that librarians adopted an attitude of waiting to follow the rest of the sector. Similarly, in Tanzania, a qualitative study was performed at Dodoma University. The study had (N=14) participants. It was found that researchers and stakeholders had a low level of awareness of RDM practices. Since the adoption of RDM practices from other institutions would be affordable, these practices need to be aligned with the university objectives, the researchers' culture and the context. Also, RDM involves departments such as library services, the ICT department, the research office, researchers and the executive management. They further discussed the different phases of implementing RDM at the university regarding policies and infrastructure, repository content development, management of active data and data selection and preservation (Mushi et al., 2020).

However, in developed countries, the situation seems far from optimal. For instance, it was found that in the UK among (N=82) institutions, (n=44, 53.8%) had technology infrastructure, (n=30, 37.5%) an institutional repository or (n=33, 41.3%) having access to external datasets (Corall et al., 2013). Also, in the US, Faniel and Connaway (2018), identified factors of influence, such as technical resources or researchers' perceptions about the library. Having services such as digital repositories was a factor of influence to support researchers. Another factor was having full-time RDM experts in the library to support RDM needs but, researchers' perceptions about expertise and librarians' services were not considered valuable.

In a study of the National Institutes of Health Library in the US, Kim and Syn (2021) interviewed four librarians who provide data services at the library. To perform the interviews, they created a two-model crosstab framework combining 3 categories of services from the Online Computer Library Center (OCLC) RDM services and 6 different levels of the data lifecycle detected from the existing different data lifecycles. In the case of the OCLC, the categories were: education services, expert services, and curation services. From the data lifecycle, they identified the following categories: data creation, data description, data storage, data sharing, and data preservation. They produced the crosstab framework to identify service gaps. In their results, they found that the collaboration operates at multiple levels involving different units such as the library units or external partners.

# 4.3. Data management plans

In recent years, data management plans have increasingly become a requirement for funding agencies, accreditation agencies and stakeholders. Thus, when researchers are deliberating on the development of new research projects, they need to elaborate a DMP which considers the context of their institution. At the University of Dodoma in Tanzania, the findings suggested a need to create an interdisciplinary team to generate a DMP template (Mushi et al., 2020). So, research librarians need to be trained in how to develop and implement DMPs which will support a broad range of researchers throughout the institution. Nevertheless, libraries are increasingly providing support for data management planning.

In the US, Tenopir et al. (2014) found that the provision of DMP as a research data services was offered by libraries (N=99) in a very limited number of them (n=26, 26,3%) and only 11.1% had plans to implement in the following two years. Also, in the US it was reported that the Victoria Commonwealth University (VCU) signed on as a DMPTool partner organization. This tool provides DMP templates that are customizable by partners and for different partner organizations. This partnership allowed VCU to promote the DMPTool among researchers in multiple areas and offer training to their communities (Henderson and Knott, 2015). In Australia, Cox et al. (2017) found that there was an increasing number of libraries that provided support in DPM (N=74, 25%) similar to the United Kingdom (N=47, 34%) in relation to former studies.

In Germany, Lindstadt and Smith (2019) explain how a founded consortium emerged from the Information Centre for Life Sciences (Informationszentrum Lebenswissenchaften ZB MED) to include life sciences with 20 German institutions, in order to provide a national infrastructure of research data, and to ensure the quality and interoperability through standards. This consortium also created PUBLISSO, a support service to assist researchers in publishing in Open Access Life sciences, which includes the following disciplines: medicine, biology, nutrition, agricultural sciences, and environmental sciences. Each of these mentioned different disciplines perform RDM in a different way, which should be considered inside the research life cycle. Within this structure, a library can offer comprehensive support services to researchers such as publication services, RDM or data findability enhancement. For instance, in the case of RDM on how to develop a data management plan (DMP), Humbolt-Universität-Berlin had created a guide for the German-speaking area on how to use DMPonline, a tool designed by the Digital Curation Center (DCC) in the United Kingdom.

In China, Huang et al., (2021) in a questionnaire answered by university libraries (N=42) found that as a strategic priority creating DMP support was lower than in western countries compared to former international studies (Cox et al., 2017; Cox et al., 2019). This also implies that DMP support to researchers is scarce and it is necessary to develop policies in the libraries to adopt good practices.

# 4.4. Relationship with stakeholders

Library stakeholders such as academic departments, need to collaborate with libraries in order to promote RDM awareness and RDM services. In relation to this, at unit level, there seems to be a need for academic libraries to foster collaborations across the university to include other administrative units, IT units, and the multitude of academic departments and research institutes across campus. These collaborations would allow for the provision of better support services to researchers. In addition, it is also necessary to collaborate and adopt agreements with external stakeholders such as technology providers to support RDM services in as many ways as possible. Because of the introduction of RDM services, and possible new services, in some research libraries new library positions have emerged such as "data curator", "data librarian", "data scientist", "data specialist", "data visualization librarian", "digital stewardship", "research data support", or "science librarian" (Ball, 2013; Bradley-Ridout, 2018; Corrall et al., 2013; Kim and Syn, 2021; Shipman and Tang, 2019; Tenopir et al., 2014). Also, at the university level, RDM policies need to be supported by the university management officers.

### 4.5. Promotion of RDM by academic libraries

In promoting RDM services among faculty and students it is necessary to install good practices in relation to data management, although this is not always possible. In New Zealand, Steiner (2015) explored Lincoln University, and their information literacy program on RDM, and interviewed six head research officers from six academic libraries. Steiner found that one of the main barriers in RDM is that researchers have different interests according to their disciplines, especially in areas such as data sharing. Another finding was that universities' enormous competition for funding and publications makes collaboration among academic libraries difficult. Steiner (2015) also compared the situation in Germany, arguing that a funding policy performed by the German Research Foundation (Deutschen Forschungsgemeinschaft or DFG) to promote open access was essential for the implementation of local developments.

Another way to explore RDM services in academic libraries is through performing content analysis on libraries' websites. Yoon and Schultz (2017) explored 185 academic library websites that promoted RDM services in the US. The research focused on service, information, education, and network. Through content analysis they showed that libraries offered data management (65%) and data service (17%), and that data curation was only used by a small number of libraries (3%). They also found that 60% had data deposit services. Data management was the second most offered service (41.1%), while data methodology (15.7%) was only provided by libraries with a data service unit. In relation to information, 80% of the libraries provided data management planning. Regarding education, they found that not all libraries had educational programmes in the form of workshops or lectures (34.5%). Only 21.8% provided training on data sharing and data reuse within data management training. Yoon and Schultz (2017) conclude that academic libraries must engage more actively to provide and develop more educational services for RDM (Yoon and Schultz, 2017). Another content analysis study was conducted by Shelly and Jackson (2018). They examined how 13 Australian universities assisted their researchers in RDM, by analysing their policies related to RDM or open access through content analysis. They found that 11 universities had introduced specific RDM policies, and 3 universities also introduced stand-alone open access policies. Their findings also indicate that researchers had different roles in all universities. For data storage, 11 universities provided internal storage options, and data uploading services. 9 universities recommend external data storage such as Research Data Australia to share and re-use data. RDM training for researchers was divided among libraries on one side, and e-research groups on the other. At the same time, it was not easy to find information on training courses on libraries' webpages. They concluded that there was no consistent approach to how universities addressed RDM services. While they found that the library played a primary role in managing RDM services in small universities, it was, however, unclear if libraries' staff had the adequate skills and experience to support researchers.

Similarly, in the US, Bardin et al. (2018) reported their experience in the Translational Research and Information Lab (TRAIL), an initiative founded at the University of Washington (UW) Health Services Library. In a triennial survey, the UW libraries found a recurring topic with regards to the needs of faculty members: services that support managing, archiving, and preserving research data. From these findings, they devised an initiative to better coordinate data and innovation services for clinical researchers through services offered by the UW, and then created partnerships with different units of the UW under six guiding principles: collaboration, quality, assessment, diversity, education, and access. This partnership with other units allowed the clinical research data management (CRDM) to expand services beyond patient records and literature review. These services include data visualisation, survey creation, bioinformatics consultation and the use of emerging technologies.

# 4.6. Limitations

This study has some limitations. While the authors retrieved as many papers as possible for the analysis, it should not be considered as an exhaustive analysis, as varying studies may also be missing from the sample. The authors observed that there are more case studies focused on one institution rather than research papers involving different institutions at the international or national level. Therefore, more research studies would enrich the literature and show best practices in RDM.

# 5. Discussion

In relation to the role of academic libraries in research data management (RQ1) the authors observed different strategies at different levels. First, the findings suggest that the library's role varies among countries. In western countries, it seems that there is a starting point where academic librarians are receiving training, and getting knowledge for the complete data lifecycle (Bishop et al., 2021, Cox et al., 2017; Yoon and Schultz, 2017). In other countries like China or Tanzania, there is still room for the development of RDM policies and specialised librarians to support researchers (Huang et al., 2021; Mushi et al., 2020).

Second, with regards to training, the literature seems to suggest that librarians do not always have adequate training for their professional development. There is a lack of adequate skills in the different topics which encompass RDM such as DMP, data preservation, data reuse or copyright issues (Bradley-Ridout, 2018; Kim and Syn, 2021; Shelly and Jackson, 2018). Also, it can be challenging for academic librarians and information professionals to acquire the necessary skills to support new roles in RDM (Cox and Pinfield, 2014). Although, the skills needed can be acquired through different actions such as scenario-based-learning (Searle, 2015) where a librarian is in a safer space dealing with advanced knowledge or also through informal training activities such as coffee lectures (Strauch, 2020), using prepared toolkits being part of an informal learning process (Read et al., 2019), workshops, MOOC's or data science courses to train librarians of all types such as academic, medical or government (Shipman and Tang, 2019).

Third, academic libraries can play a key role in their institutions under certain conditions. Researchers are drivers of RDM and libraries have to pay attention to their needs when developing new services. This implies acquiring feedback from the user's perspective and listening to users about library services (Bishop et al., 2021; Blumesberger, 2020; Huang et al., 2021; Pinfield et al., 2014). Thus, collaboration among departments and stakeholders would allow academic libraries to be more efficient in joining efforts for RDM services (Tenopir et al., 2014; Cox et al., 2016; Faniel and Connaway, 2018). When offering RDM services, it is also important to consider the context in which the library is placed. Data-intensive units such as meteorology or earth science do not have the same requirements as units which have less data intensive disciplines (Bishop et al., 2021; Huang et al., 2021; Steiner, 2015).

In relation to the promotion of RDM across institutions (RQ2), the findings show that libraries promote RDM differently. First, a popular approach to promote the services is through websites. That is, explaining the services they offer, but it seems that there is not always consistency on what is offered, probably due to a lack of staff. Also, websites need to be improved to provide key information on services and infrastructures to researchers and new users (Shelly and Jackson, 2018), with tailored information to address specialised communities (Lindstädt and Schmitz, 2019).

Second, another way to promote RDM is by offering guidance to researchers on how to reuse data within and outside the institution (Ball, 2013). However, the open data concept or "openness" is not established in the same manner in all countries (Huang et al., 2021; Kruse and Thestrup, 2014). This guidance also includes assisting researchers in writing DMPs, good data practices, and promoting the use of the FAIR principles. Nevertheless, it is essential that this knowledge exists among the library workforce (Bishop et al., 2021; Corrall et al., 2013; Hansen et al., 2021; Henderson and Knott, 2015).

A third way for promoting RDM services is to locate data services and have them as a centralised resource. Libraries have considerable experience in assessing users on different topics. In relation to RDM, data services such as data curation or having a technical structure to upload data was also observed in the literature. Nevertheless, technological structures need to be sustainable and cost-effective models, such as with shared repositories among different institutions. At the same time, repositories need to be developed with consideration of the specialised communities that require RDM services so as to ensure maximum diversity of RDM use cases (Blumesberger, 2020; Clements, 2013; Lindstädt and Schmitz, 2019; Steiner, 2015).

# 6. Conclusions

This systematic review examined the importance of RDM for academic libraries and the role they play in higher education institutions. There are two common points to be found in this review, being collaboration and training. Collaboration is necessary among libraries and their stakeholders to provide support in RDM to researchers and students. Also, collaborations are needed when it comes to the development of RDM services for a diversity of disciplines. Moreover, the literature shows that collaboration with other research institutions is beneficial in order to develop sustainable technological structures.

The literature indicates that faculty members and students need support and training in a wide variety of topics such as data curation, data sharing and data reuse as well as guidance on legal and ethical issues. The need for RDM services means that academic librarians must also be provided with support and training to enable them to deliver these services, as well as to enable them to provide training to faculty members who need these services. Moreover, their training must also consider the context of their institutions, and the diversity of researchers who require RDM services. Future data management training in library schools should be designed in such a way that it prepares future academic librarians for these new and emerging challenges. On the other hand, these new challenges also bring new opportunities for developing the role of academic libraries within the university ecosystem.

#### ACKNOWLEDGEMENTS

This work was supported by the Spanish Ministerio de Innovación, Ciencia y Universidades [grant ref RTI2018-094360-B-I00]. Authors thank the Ministry for funding the project.

# REFERENCES

- Abadal, E. (2021). Ciencia abierta: un modelo con piezas por encajar. *Arbor*, 197(799), a588. https://doi.org/10.3989/arbor.2021.799003.
- Ashiq, M.; M. H. Usmani and M. Naeem (2020). A systematic literature review on research data management practices and services. *Global Knowledge, Memory and Communication*, (ahead-of-print). https://doi.org/10.1108/GKMC-07-2020-0103.
- Ayris, P. (2017). Brexit: and its potential impact for open access in the UK. *Insights: the UKSG Journal*, 30, 1: 4–10. https://doi.org/10.1629/uksg.336.
- Ball, J. (2013). Research data management for libraries: getting started. *Insights*, 26, 3: 256–260. https://doi.org/10.1629/2048-7754.70.
- Bardyn et al. (2018). Bardyn, T.; E. Patridge; M. Moore and J. Koh. Health sciences libraries advancing collaborative clinical research data management in universities. *Journal of EScience Librarianship*, 7, 2. https://doi.org/10.7191/jeslib.2018.1130.
- Bishop et al. (2021). Bishop, B. W.; H. R. Collier; A. M. Orehek and M. Ihli. Potential roles for science librarians in research data management: a gap analysis. *Issues in Science and Technology Librarianship*, 98. https://doi.org/10.29173/ISTL2602.
- Blumesberger, S. (2020). Repositorien als Tools für ein umfassendes Forschungsdatenmanagement: am Beispiel von PHAIDRA an der Universitätsbibliothek Wien. *Bibliothek Forschung und Praxis*, 44, 3: 503–511. https://doi.org/10.1515/bfp-2020-2026.
- Boté, J.-J. and M. Térmens (2019). Reusing data: technical and ethical challenges. *DE-SIDOC Journal of Library and Information Technology*, 39, 6: 329–337. https://doi.org/10.14429/djlit.39.6.14807.
- Bradley-Ridout, G. (2018). Preferred but not required: examining research data management roles in health science librarian positions. *Journal of the Canadian Health Libraries Association*, 39, 3: 138–145. https://doi.org/10.29173/JCHLA29368.
- Chiware, E. R. T. and Z. Mathe (2015). Academic libraries' role in research data management services: a South African perspective. South African Journal of Libraries and Information Science, 81, 2, Article 2. https://doi.org/10.7553/81-2-1563.

- Chu et al. (2022). Chu, C. M.; J. Raju; C. Cunningham; J. Ji; V. Ortíz-Repiso Jiménez; A. Slavic; A. M. Talavera-Ibarra and S. Zakaria. *IFLA Guidelines for Professional LIS Education Programmes*. [cited: 2022–09–25]. Available from: https://repository.ifla.org/handle/123456789/1987
- Clements, A. (2013). Research information meets research data management ... in the library? *Insights*, 26, 3: 298–304. https://doi.org/10.1629/2048-7754.99.
- Corrall, S.; M. A. Kennan and W. Afzal (2013). Bibliometrics and research data management services: emerging trends in library support for research. *Library Trends*, 61, 3: 636–674. https://doi.org/10.1353/lib.2013.0005.
- Cox, A. M. and S. Pinfield (2014). 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.
- Cox et al. (2017). Cox, A. M.; M. A. Kennan; L. Lyon and S. Pinfield. 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: 2182–2200. https://doi.org/10.1002/asi.23781.
- Cox et al. (2019). Cox, A. M.; M. A. Kennan; L. Lyon; S. Pinfield and L. Sbaffi. Maturing research data services and the transformation of academic libraries. *Journal of Documentation*, 75, 6: 1432–1462. https://doi.org/10.1108/JD-12-2018-0211.
- Cox, A. M.; S. Pinfield and J. Smith (2016). Moving a brick building: UK libraries coping with research data management as a 'wicked' problem. *Journal of Librarianship* and Information Science, 48, 1: 3–17. https://doi.org/10.1177/0961000614533717.
- European Commission. (2019). *Open Science*. [cited: 2022–08–12]. Available from: https://ec.europa.eu/info/sites/default/files/research\_and\_innovation/knowledge\_ publications\_tools\_and\_data/documents/ec\_rtd\_factsheet-open-science\_2019.pdf
- Faniel, I. M. and L. S. Connaway (2018). Librarians' perspectives on the factors influencing research data management programs. *College and Research Libraries*, 79, 1: 100–119. https://doi.org/10.5860/crl.79.1.100.
- Hansen et al. (2021). Hansen, J. S.; S. Gadegaard; K. K. Hansen; A. V. Larsen; S. Møller; G. S. Thomsen and K. F. Holmstrand. Research data management challenges in citizen science projects and recommendations for library support services: a scoping review and case study. *Data Science Journal*, 20, 1: 25. https://doi.org/10.5334/dsj-2021-025.
- Healy et al. (2022). Healy, S.; H. Byrne; K. Schmid; N. Bingham; O. Holownia; M. Kurzmeier and R. Jansma. Skills, tools, and knowledge ecologies in Web Archive research, *WARCnet Special Reports*, August 2022. Aarhus, Denmark: WARCnet.
- Henderson, M. E. and T. L. Knott (2015). Starting a research data management program based in a university library. *Medical Reference Services Quarterly*, 34, 1: 47–59. https://doi.org/10.1080/02763869.2015.986783.

- Huang, Y.; A. M. Cox and L. Sbaffi (2021). Research data management policy and practice in Chinese university libraries. *Journal of the Association for Information Science and Technology*, 72, 4: 493–506. https://doi.org/10.1002/asi.24413.
- IFLA (2021). UNESCO agrees Open Science recommendation: implications for libraries. [cited: 2022–08–10]. Available from: https://www.ifla.org/news/unesco-agrees-open-science-recommendation-implications-for-libraries/
- JISC. (2021). *Research data management toolkit*. [cited: 2022–08–12]. Available from: Jisc. https://www.jisc.ac.uk/guides/rdm-toolkit
- Kim, S. and S. Y. Syn (2021). Practical considerations for a library's research data management services: the case of the National Institutes of Health Library. *Journal of the Medical Library Association*, 109, 3: 450–458. https://doi.org/10.5195/jmla.2021.995.
- Komiyama, Y. and K. Yamaji (2017). Nationwide research data management service of Japan in the Open Science era. IN: 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI). (Pp 129–133). https://doi.org/10.1109/IIAI-AAI.2017.144.
- Kruse, F. and J. B. Thestrup (2014). Research libraries' new role in research data management: current trends and visions in Denmark. *LIBER Quarterly: The Journal of the Association of European Research Libraries*, 23, 4: 310–335. https://doi.org/10.18352/lq.9173.
- Lefebvre, A. and M. Spruit (2021). Laboratory forensics for Open Science readiness: an investigative approach to research data management. *Information Systems Frontiers*. https://doi.org/10.1007/s10796-021-10165-1.
- Lindstädt, B. and J. Schmitz (2019). Das Management von Forschungsdaten als Handlungsfeld wissenschaftlicher Bibliotheken: Forschungsunterstützung am Beispiel ZB MED – Informationszentrum Lebenswissenschaften. *Bibliothek Forschung und Praxis*, 43, 1: 42–48. https://doi.org/10.1515/bfp-2019-2006.
- Muellenbach, J. M. (2021). A Pilot to initiate research data management services within academic libraries helps librarians to learn about, engage with, and enhance skills within their research communities. *Evidence Based Library and Information Practice*, 16, 1: 104–106. https://doi.org/10.18438/eblip29879.
- Mushi, G. E.; H. Pienaar and M. van Deventer (2020). Identifying and implementing relevant research data management services for the library at the University of Dodoma, Tanzania. *Data Science Journal*, 19, 1: 1. https://doi.org/10.5334/dsj-2020-001.
- Nie, H.; P. Luo and P. Fu (2021). Research data management implementation at Peking University Library: foster and promote Open Science and Open Data. *Data Intelli*gence, 3, 1: 189–204. https://doi.org/10.1162/dint\_a\_00088.
- Page et al. (2021). Page M. J.; J. E. McKenzie; P. M. Bossuyt; I. Boutron; T. C. Hoffmann; C. D. Mulrow; L. Shamseer et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021; 372:n71. doi: 10.1136/bmj.n71.

- Pinfield, S.; A. M. Cox and J. Smith, J. (2014). Research data management and libraries: relationships, activities, drivers and influences. *PLOS ONE*, 9, 12: e114734. https://doi.org/10.1371/journal.pone.0114734.
- Read et al. (2019). Read, K. B.; J. Koos; R. S. Miller; C. F. Miller; G. A. Phillips; L. Scheinfeld and A. Surkis. A model for initiating research data management services at academic libraries. *Journal of the Medical Library Association*, 107, 3: 432–441. https://doi.org/10.5195/jmla.2019.545.
- Read, K. and A. Surkis (2018). *Research data management teaching toolkit [Data set]*. *figshare*. https://doi.org/10.6084/m9.figshare.5042998.v6.
- Rehwald, S. and J. Stegemann (2021). Roadmap zur Servicestelle für Forschungsdatenmanagement am Beispiel der Universitätsbibliothek Duisburg-Essen. *Information* – *Wissenschaft & Praxis*, 72, 4: 194–203. https://doi.org/10.1515/iwp-2021-2161.
- Rice, R. (2019). Supporting research data management and Open Science in academic libraries: a data librarian's view. VOEB-Mitteilungen, 72, 2: 263–273. https://doi.org/10.31263/voebm.v72i2.3303.
- Searle, S. (2015). Using scenarios in introductory research data management workshops for library staff. *D-Lib Magazine*, 21, 11/12. https://doi.org/10.1045/november2015-searle.
- Shelly, M. and M. Jackson (2018). Research data management compliance: is there a bigger role for university libraries? *Journal of the Australian Library and Information Association*, 67, 4: 394–410. https://doi.org/10.1080/24750158.2018.1536690.
- Shipman, J. P. and R. Tang (2019). The collaborative creation of a Research Data Management Librarian Academy (RDMLA). *Information Services and Use*, 39, 3: 243– 247. https://doi.org/10.3233/ISU-190050.
- Slavnic, Z. (2017). Research and data-sharing policy in Sweden: neoliberal courses, forces and discourses. *Prometheus*, 35, 4: 249–266. https://doi.org/10.1080/08109028.2018.1499542.
- Steiner, K. (2015). Forschungsdatenmanagement und Informationskompetenz: Neue Entwicklungen an Hochschulbibliotheken Neuseelands = Research data management and information literacy: new developments at New Zealand university libraries. Information – Wissenschaft & Praxis, 66, 4: 230–236. https://doi.org/10.1515/iwp-2015-0040.
- Strauch, A. (2019). Forschungsdatenmanagement an der Stiftung Universität Hildesheim: praktische Unterstützung für Forschende und Studierende durch die Universitätsbibliothek. Information – Wissenschaft & Praxis, 70, 5/6: 259–263. https://doi.org/10.1515/iwp-2019-2052.
- Strauch, A. (2020). "To begin, at the beginning [...]": Bibliotheken als "Player" im professionellen Forschungsdatenmanagement. Universitätsbibliothek Hildesheim. *Bibliothek Forschung und Praxis*, 44, 2: 166–169. https://doi.org/10.1515/bfp-2020-0025.

- Tenopir et al. (2014). Tenopir, C.; R. J. Sandusky; S. Allard and B. Birch. Research data management services in academic research libraries and perceptions of librarians. *Library and Information Science Research*, 36, 2: 84–90. https://doi.org/10.1016/j.lisr.2013.11.003.
- Traub, I. D.; B. S. Solis and P. Budroni (2016). Forschungsdaten und zeitgemäße Aufarbeitung durch Policies – 2. Internationaler LEARN Workshop zum Thema, Forschungsdatenmanagement<sup>6</sup> (Wien, 6. April 2016). *Mitteilungen der Vereinigung* Österreichischer *Bibliothekarinnen und Bibliothekare*, 69, 1: 142–150. https://doi.org/10.31263/voebm.v69i1.1406.
- Truter, V. Z. (2021). Research data management and sharing practices in the digital humanities with a focus on publisher support: a case study in the field of Web archive studies: MA thesis. Faculty of the Humanities, University of Cape Town. [cited: 2022–08–10]. Available from: https://open.uct.ac.za/handle/11427/36196
- UK Data Archive. (2022). *Research Data Management*. [cited: 2022–08–12]. Available from: https://www.data-archive.ac.uk/managing-data/
- UNESCO. (2021). UNESCO recommendations for Open Science. [cited: 2022–08–10]. Available from: https://unesdoc.unesco.org/ark:/48223/pf0000379949.locale=en
- WARCnet. (s. a.). *Working Groups Aarhus University*. [cited: 2022–08–10]. Available from: https://cc.au.dk/en/warcnet/working-groups
- Wilkinson et al. (2016). Wilkinson, M. D.; M. Dumontier; Ij. J. Aalbersberg; G. Appleton; M. Axton; A. Baak; N. Blomberg et al. The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3, 1: 160018. https://doi.org/10.1038/sdata.2016.18.
- Wu, M. and X. Chen (2016). Library service design based on the needs of chemistry research data management and sharing survey. *Proceedings of the Association for Information Science and Technology*, 53, 1: 1–4. https://doi.org/10.1002/pra2.2016.14505301137.
- Yoon, A. and T. Schultz (2017). Research data management services in academic libraries in the US: a content analysis of libraries' websites. College & Research Libraries. https://doi.org/10.5860/crl.78.7.920.