The study aimed to explore the Research Data Management (RDM) practices of university faculty members through qualitative research design. The data was collected through semi-structured, in-depth interviews from purposively selected ten faculty members from the University of Punjab (PU). The study discovered some significant factors including RDM and curation practices, the amount of research data produced, the support needed for data curation and their willingness to share it. In addition, the study explored issues the researchers face with regards to RDM. The findings reveal that respondents need assistance regarding storage and security of data, improving the quality of backup, support for storage and preservation. They agreed with a need for a central repository of the University.

Keywords: Research Data Management; Academics RDM skills; Research Data Management Services; University of the Punjab

INTRODUCTION

Data management is an active process by which digital resources remain discoverable, easy to use and accessible for a longer time (Procter, Halfpenny, & Voss, 2014; Steinhart, 2014). RDM is defined by Cox and Verban (2018) as creating, finding, organizing, storing, sharing and preserving data within any research process. Whyte and Tedds (2011) explained RDM as, “the organization of data, from its entry to the research cycle through the dissemination and archiving of valuable results”. Cox and Pinfield (2014) stated that, “RDM consists of a number of different activities and processes associated with the data lifecycle, involving the design and creation of data, storage, security, preservation, retrieval, sharing, and reuse, all taking into account technical capabilities, ethical considerations, legal issues and governance frameworks”.

The issue of RDM has got the attention of research scholars, data scientists, library and information scientists. The studies are being conducted on different but
related aspects by computer scientists, data scientists, library and information professional and so forth. The RDM practices of researchers and academics and their data literacy skills have also emerged as a trending research area. For example, Kim (2015) investigated RDM practices and perceptions of data sharing of university researchers in Korea and found that the majority participants perceived research data sharing useful and the participants usually shared data with researchers they know. Some agreed that public funded data should be open to the public, while others partially agreed with it. Similarly, Henty et al. (2008) reported that the researchers created digital data in the life cycle of projects. The researchers used software for storing and backing up their data. Most of the researchers had no RDM plans but were willing to share data. Another case study, Van Tuyl and Michalek (2015) investigated the RDM practices of Carnegie Mellon University faculty members and revealed a need for the expansion of data management services, support for long-term data preservation, discoverability, assistance in backup, storage and preservation of data.

On the one hand, higher education and research in Pakistan have been growing fast and a huge amount of data in digital form is being produced with hardly available its record at university’s level. On the other hand, university libraries, specifically of the HEC high ranked universities’ have been re-envisioning their services in order to meet the evolving needs of the academic scholars. Under this background, the present study was designed to investigate RDM practices of faculty members through qualitative design using interview as the method to get understanding of their practices, taking sample from the University of the Punjab. This is a first baseline qualitative study on the subject.

**LITERATURE REVIEW**

The researcher used ‘faculty practices of RDM’, ‘researcher’s practices of RDM’, the ‘challenges of data for faculty’ as phrases for searching database ProQuest and Emerald, Science Direct and Google Scholar. The review of the literature has established that libraries involved in RDM are currently working in the areas of advocacy and policy development while supporting creation of a new systems. Some of the studies have explored data preservation and withholding practices of researchers in the developed world and discussed perspectives of RDM practices benefits of reuse. For instance, Aydinoglu, Dogan, and Taskin (2017) found that Turkish researchers were aware of the benefits of data management, were willing to share their research data, and had decent preservation habits. They
expressed that they lacked the technical skills and knowledge needed for RDM. In addition, no institutionalized support (staff, training, software, and hardware) was provided to researchers. Berman (2017) explored the RDM practices of scientists of the University of Vermont (USA) and identified four major areas of research data services where researchers need assistance: infrastructure, metadata, data analysis and statistical support, and informational research data services. Shen (2015) found that while analyzing the RDM reuse practices of faculty researchers that potential data were lost right after the original work was done. Akers and Doty (2013) conducted a survey of Emory University USA in order to understand disciplinary differences of faculty practices and perspective of RDM to guide department of library services to support the management of data at university. They categorized faculty members into four research domains that included: arts and humanities, social sciences, medical sciences, and basic sciences and found significant differences among the four research domains for RDM.

Ward et al. (2011) noted that many researchers of the University of Cambridge and Glasgow organized their research data in an ad hoc fashion, and found difficulties in retrieval, storage, backup, sharing and re-use. They recommended a bottom-up approach to support researchers with step-by-step development of RDM practices by understanding the researchers’ needs and perspectives.

As far as data sharing is concerned, Koltay (2015) pointed out that the researchers would be in a position to share their data if they were rewarded for data citations. Drachen et al. (2016) further reported that data citations also increased the number of citations to papers. It is the biggest challenge that researchers perceive that librarians do not understand the complications of data and may not play any role for its orderly preservation (Read et al., 2015).

Librarians play a pivotal role in supporting researchers regarding RDM services as Rice and Southall (2016) emphasized on supporting data literacy, and stressed that librarians often arranged programmes of information literacy or literature searching for students and research scholars. They can easily add concepts of data management and data reuse. Librarians may give instructions in using standalone or online reference management tools such as Endnote, Reference Manager and Zotero; such tools may be the best methods of controlling data throughout research projects. Delasalle (2013) argued that researchers and administrators need to work together to address challenges of RDM such as accessibility of data for future use.
An effective RDM policy is necessary to initiate the development of RDS in an organization. Authors have recently pointed out the importance of forming policy. For example, Searle (2015) reported that the Griffith University, Australia has enforced a policy regarding RDM and directed researchers to follow highest standards while managing research data. The standards included: policies and requirements of funding agencies, technical protocols, legislation and expectations of border community. Carlson et al. (2015) mentioned that an outreach support programme regarding RDM was arranged for the faculty members of Purdue University, USA. Many junior faculty members were targeted creating data management within their research labs for the first time. A series of workshops were arranged that included topics such as: data management issues, data literacy, background information on RDM and knowledge of RDM tools. The scientists do not make their data available to other researchers for various reasons that include: insufficient time, initial and short-term parts of research data, non-availability of support from their organizations. There are also significant differences and approaches in data management practices based on primary funding agency, subject discipline, age, work focus, and world region (Tenopir et al., 2011).

Whitmire, Boock, and Sutton (2015) revealed that the Oregon State University (OSU) researchers were generating a wide variety of data types. They discovered that the faculty was not using campus data storage infrastructure, and maintained their own storage servers. Many faculty members were creating metadata but there was a need to provide support to improve metadata standards.

Tenopir, Birch, and Allard (2012) reported that data-intensive research created challenges for scholars and demanded an RDM skill set that was not part of the scientist’s education and for which they had no time for training. There are several initiatives, such as NSF’s DataNet program that have been working to address the need for this skill set.

The literature review has established that this area has been studied from various angles by the researchers from various countries, but has not by Pakistani information professionals. Only one study by Ameen and Rafiq (2017) investigated research data literacy skills of scientists belonging to four universities of Pakistan. It reported that the respondents showed a lack of skills in the use of standard metadata. They saw a major role of university libraries in guiding them in the use of standards in data management.
Research Objective

The main objective of the study was to explore the RDM practices of faculty members. The following research questions were asked to achieve the objective:

1. What kind of research projects academics are engaged with, and what amount and type of data they produce?
2. What are their data saving and backup practices?
3. What is their opinion about sharing data?
4. What kind of data management support they required?
5. What are the major issues and challenges they face in RDM?

METHODOLOGY

In order to meet the objectives of the study the qualitative data were collected through semi-structured, in-depth interviews of 10 faculty members due to a limited availability of the academics. The sample was purposively selected on the bases of their research publication record (Professors of Emeritus, Professors, Associate Professors, Assistant Professors and Lecturers having more than 15 impact factor papers in their academic carrier). They belonged to the disciplines of Earth Sciences, Biochemistry, Microbiology, Chemical Engineering, Mathematics, Administrative Science, Botany, Zoology and Medical. Since the availability of the faculty members were limited.

Based on previous literature available regarding practices of researchers about research data, an interview-guide of seven key themes was designed to conduct interviews. The interview was focused on these major elements: a) an overview of faculty members’ research activities; b) types of research projects, data created; c) management practices; and d) willingness to share data for reuse. The sample was selected purposively from HEC’s top 2nd ranked “University of the Punjab”. The interviews were audio recorded after getting participants permission. The transcripts were analyzed using thematic approach. The anonymity of the interviewees has to be maintained. The study is based on the part of qualitative data collected during 2017 by the first author for his doctoral study.

RESULTS

The following section presents the findings according the frequency of the themes occurrences in tables followed by the analysis, interpretations and discussion. The quotations from the interview extracts have been included to enrich the interpretation.
RQ 1. Faculty Engagement in Different Projects, Amount and Types of Data Produced

The interview started with asking about the ongoing projects. The analysis shows that they were engaged in both supervising students’ research projects for higher degrees and their own projects. Table 1 shows the opinion of faculty members.

Table 1

<table>
<thead>
<tr>
<th>Researchers’ Engagement in the Projects</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Projects of M. Phil, Ph. D, and research projects of commercial organizations were also carried out</td>
<td>4</td>
</tr>
<tr>
<td>Involved only in student research of Ph. D and personal research papers</td>
<td>4</td>
</tr>
<tr>
<td>There are foreign projects going in our labs</td>
<td>2</td>
</tr>
</tbody>
</table>

It shows that four of the respondents were involved only in M. Phil and Ph.D. students research and in their own research activities. Similarly, four of them were engaged in commercial projects such as paint, mineral exploration, food industry and health issues. Two of them were involved in foreign projects also. All faculty members had a good understanding perception of RDM and curation services. The interviewees remarked that they were involved in students’ research projects from BS to Ph.D. level, many research projects from industry and other organizations were also carried out. Microbiological tests were conducted for various commercial organizations and after completion a large amount of data was generated. According to them data were not saved and properly curated for reuse. One interviewee pointed out that “sometimes the organizations are not ready to share their research data with other competitive organizations (R#5). The researcher observed that the respondents were keenly interested that there should be a research data repository which can take care of their produced data.

A respondent remarked that “when he retired from his service, he had a big room full of research data including his field reports, lab testing notes and reports, slides of microscopic studies of rocks and minerals. Since it was not properly arranged and majority was in print format, it was really pity that he lost most of the research data” (R#2).
Aydinoglu et al. (2017) reported that Turkish researchers were aware of the benefits of data management, were willing to share their research data, and had decent preservation habits, they expressed that they lacked the technical skills and knowledge needed for RDM. In addition, no institutionalized support (staff, training, software, and hardware) was provided to researchers.

Data Saving

The researchers were asked regarding the types of data and strategy they usually used to save their research data. The question was divided in two parts: a) the kind of data they were generating; and b) the saving devices they were using for this data. It is apparent from Table 2 that they were producing various types of data and using different devices for saving it.

<table>
<thead>
<tr>
<th>Content Analysis of the Responses</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>We have both qualitative and quantitative data including: films, slides, text, graphs, product itself</td>
<td>4</td>
</tr>
<tr>
<td>A large number of numerical data, slides of thin section, photographs, power point presentations, lab note books.</td>
<td>6</td>
</tr>
<tr>
<td>Research data was saved through personal computers and laptops</td>
<td>9</td>
</tr>
<tr>
<td>Use hard disk, file servers, USB, Google Clouds, emails, Google sharing and Dropbox</td>
<td>8</td>
</tr>
</tbody>
</table>

The respondents replied that they generated different types of data such as lab notes, text documents, pictures, films, PowerPoint slides, microscopic slides films and pictures, images. Most of the respondents (8) used personal computers and laptops to save the data in soft form on Hard disks, USB, CDs, Google Cloud, Google Sharing, Drop Box and Research Gate. In some cases, these types of data and their analysis were arranged and managed in print format.

The researchers did not follow any standards or metadata structure for archiving data and no proper software was being used to retrieve data for sharing purposes. In addition, they were not provided any kind of support under which they may learn how to save research data right from the beginning of the project with the aim that it would be saved for future use. One respondent replied, “I save my research data in Drop Box. They offer up to 5 GB free space and then charge
nominal amount for extended space” (R#7). Similarly, Smith II (2014) found that 65% researchers did not use standards, best practices and guidelines.

Averkamp, Gu, and Rogers (2014) and Steinhart et al. (2012) reported that researchers generated different types and sizes of data but were not concerned about the metadata. Alexogiannopoulos, McKenney, and Pickton (2010) reported that they were anxious about the storage of their data. Most of the researchers used Microsoft software for creating documents and spreadsheets and created .doc and .xls file types; similarly, .jpeg for image files for saving research data at the University of Northampton.

RQ 2. Data Saving and Backup Practices of Faculty

The interviewees were asked how much data they usually produce and save during a life-cycle of a project or in one year.

Table 3

<table>
<thead>
<tr>
<th>Data Generated During Life-cycle of Projects</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Have no idea in GB or TB</td>
<td>2</td>
</tr>
<tr>
<td>No standard followed</td>
<td>6</td>
</tr>
<tr>
<td>Approximately 6000 pages’ research data produced in all projects per year</td>
<td>1</td>
</tr>
<tr>
<td>40 GB data per year having external hard discs for backup</td>
<td>1</td>
</tr>
<tr>
<td>Use external hard disk for backup procedure</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3 reveals that six of the respondents would take backup of data on CDs, Google cloud, Drop Box and two were using external hard disk. Six reported that no metadata standard was followed while saving or preserving data. They were not maintaining a precise idea about their data in GB or TB produced in a year or in different projects.

A professor told that he would preserve data in the form of printed pages which included graphs, lab note books findings, and other test results and almost 6000 pages were saved in a year. The researcher observed while taking the interview of a professor that thousands of reports and files of data were piled around him in bundles placed in shelves in print form, which needed scanning if to be preserved for future. It was also found that almost all researchers saved their data both in soft and print form, however, senior faculty had most of the data in print form, which needed to be scanned for curation. This shows their high time association with print form. However, those who save data in soft form use external hard disks for backup. The faculty members used several devices and
different channels for backup of the research data. It seemed necessary to know about the information regarding the body of data. The interviewees were asked to provide information about their data backup procedure.

Table 4

<table>
<thead>
<tr>
<th>Saving and backup practices of faculty members</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data is saved on Personal Computers, Laptops and file servers</td>
<td>9</td>
</tr>
<tr>
<td>Backup of data is stored on Hard Disk, USB, Google Cloud, Dropbox</td>
<td>6</td>
</tr>
<tr>
<td>Data is saved in print format</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4 displays that mostly faculty members used personal computers, laptops and departmental file servers for saving their research data. Six of the respondents told that they save their data on USBs, Hard Disk, Google Cloud, Dropbox and Research Gate for backup procedure. Some research data are also available in print format. Another interviewee responded, “they usually save their research data of projects in soft and print format, but do not use any kind of software or follow any standard for saving research data. He told that they have no support for that. So they usually take backup on hard drives and CDs” (R#1).

The qualitative analysis revealed that researchers did not follow standards or metadata structure for archiving data and no proper software is used to retrieve data for sharing purposes. One respondent replied “I save my research data at Dropbox. They offer up to 5 GB free space and then charge nominal amount for extended space. The benefit is that you have your data with you all the time” (R#7).

Similarly, the issue has also been discussed by Smith II (2014), who conducted a study for the use of standards, best practices and guidelines to manage research data of the Florida State University, USA. He found that 65% researchers do not use standards, best practices and guidelines. The study validates the findings of previous studies Averkamp et al. (2014) and Steinhart (2014) who underlined that the researchers generated different types and sizes of data but were not concerned about the metadata standards.

RQ 3. Willingness About Sharing Data

The next question asked was about their willingness in sharing their research data with other researchers and scholars through a common repository. Table 5 depicts different opinions of respondents who intended to share their data with central library.
Table 5

<table>
<thead>
<tr>
<th>Contents Analysis of the Responses</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consent should be taken from researchers</td>
<td>5</td>
</tr>
<tr>
<td>Data should be open for all as it is available on Google</td>
<td>2</td>
</tr>
<tr>
<td>Fear of data misuse</td>
<td>1</td>
</tr>
<tr>
<td>Some data has security concerns</td>
<td>1</td>
</tr>
<tr>
<td>We would not share data</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 5 shows their positive attitude towards willingness of sharing data, however, if they were guaranteed security of data. Furthermore, they shared their fears about misuse of labs data. A few respondents said there should be an option for all faculty members whether they want to provide access to their data openly or with some conditions. Two replied that data should be open for all as it has been done on Google. Three of the respondents were not ready to share their data and they stressed that if possible a repository should be initiated in their department.

A few mentioned that there was no culture of sharing data in their respective departments. They however share their raw data with students during their research projects. The interviewees said that secrecy was very important in their research activities; they could not share their data before submitting their research projects and reports. It indicated that they would share data after published research based on that data. One respondent argued that they had sensitive lab data based on observation, so if anyone gets that it can easily be manipulated and misused by others.

A respondent remarked that “nowadays the importance of research data has increased so it should not be wasted after research process. Its access should be open, but some data must be restricted. He explained some researchers are worried about the security of their data, if they were ensured that their data would be safer in central library, they would definitely share it. They must know who would use their data and would cite them also” (R#9).

MacMillan (2014) also reported that issues of security and control play an important role in RDM services. There may be a tension between funder and publisher requirements for data sharing and the technical and cultural barriers that inhabit such sharing.

Tenopir et al. (2017) expressed that the attitudes, behavior and willingness of researchers for data sharing are the basis of managing research data by themselves and by other stakeholders involved in the processes and RDS.
RQ 4. Help Required in Managing Data

It was asked if they needed some data literacy trainings to improve their research data curation practices and strategies, and how do they see the role of library in this regard.

Table 6

<table>
<thead>
<tr>
<th>Help Required Regarding RDM</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A special librarian is required who knows the procedure and may visit departments to guide us</td>
<td>6</td>
</tr>
<tr>
<td>A support person was required</td>
<td>4</td>
</tr>
<tr>
<td>Support required for metadata standards and proper software</td>
<td>3</td>
</tr>
<tr>
<td>Department may provide a computer literate support person</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 6 shows that all of them showed a need to have help in enhancing their data literacy skills and gave suggestions in this regard. Six of them urged that there should be skillful library professional to guide them about better options of data saving. However, others said that guidance about research data saving is not the responsibility of the library professionals only, instead, the computer literate staff from our own department can also take care of it.

A respondent stated that every Ph.D. student was issued a lab notebook and supervisors regularly check how a student records observations and readings of the different experiments. When asked whether data in these notebooks was saved somewhere in the digital archive? The answer was a ‘no’.

The respondents admitted that managing research data has become a technical job and that most of us have data in print forms such as reports, diaries, lab. notebooks. It should be scanned on certain standards which could be saved. In addition, they needed some support in the form of training as to how they should manage their research data in soft format, which can be uploaded to repositories easily. They were also of the view that since the data properly saved for future use required certain skills, they felt a need to get some sort of facilitation from trained library staff in order to learn new techniques regarding data preservation, treatment and curation.

This was a conversation which explored the perception of faculty about having a research data management repository in the central library. They were further required to tell that if they agreed with this kind of facility then what should
be the policy regarding sharing and curation, mechanism of access provision and data security. If they did not want to share their research data, then they were asked to state the reasons behind their apprehensions in this regard.

**Central Repository**

The researchers’ views on establishing a central repository for research data preservation in central library of every university is as follows:

<table>
<thead>
<tr>
<th>Contents Analysis of the Responses</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nothing can be more beneficial than creating a repository in central library</td>
<td>8</td>
</tr>
<tr>
<td>There should be a repository and data also be checked through Turnitin before uploading there</td>
<td>2</td>
</tr>
<tr>
<td>A data repository required in Department</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 7 shows that all of the researchers gave a positive response to this question and said that there should be a central repository. They were willing to share their research data under certain conditions. There was also an opinion from the respondents that “research data must be checked through Turnitin (a similarity checking software to curtail plagiarism) before uploading to the central repository” (R#5). The respondents appreciated the idea of establishing a research data repository in the central library. They were so anxious about it that one respondent replied that “if this has been established before, he would not have lost his 50 years’ research data” (R#1). It is the need of the time to initiate an RDM curation center in the central library. However, it appears quite complicated as many of the senior professors had their research data in print or handwritten form, which needed conversion before being uploaded to a repository.

**RQ 5. Issues and Challenges Faced by Academicians Regarding RDM**

In the end, the respondents were asked to share the issues they had faced in curing their research data at a personal level.
Table 8

Challenges of Researchers for RDM

<table>
<thead>
<tr>
<th>Contents Analysis of the Responses</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of skills required to use metadata and standard software</td>
<td>5</td>
</tr>
<tr>
<td>Lack of awareness about who will save, and manage data for sharing</td>
<td>2</td>
</tr>
<tr>
<td>A focal person should be there visit for training</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 8 shows that a majority of the respondents revealed a need for training in using metadata and standard software. Three of them stressed that a focal person may be from the central library or from their own department who should take care of research data. They mentioned that help from their university or department was not extended in this regard. Some respondents stated that nobody knew about who was responsible for data preservation. There must be a person who looks after these technical matters. They have computers but do not have access to specialized software which was necessary to keep the data. All the respondents were of the view that they need some sort of support on how to save data, provide metadata to data and use some software for this purpose.

To highlight the library services to provide support and training to researchers, Si et al. (2015) conducted a survey of 87 best libraries of the world’s top ranked universities and presented that only 57.5% of those provided RDM services including research data, data curation and storage, guidelines and training. Similarly, Corrall, Kennan, and Afzal (2013) surveyed 140 libraries of Australia, New Zealand, and UK and highlighted the low level of engagement of libraries in providing RDS support and training.

CONCLUSION

The findings establish that the faculty members have been generating data in digital form, and the faculties of the University of the Punjab hold a vast amount of research data in printed or handwritten form. The analysis reveals that use and demand of research data management services is prevalent among faculty members of the University. They urged a need for learning the skills related to making RDM plans in a standardized form. They wanted support in long-term preservation, reuse, and improved backup storage quality of data. The major issues identified were: lack of support staff having adequate knowledge of computer skills. Almost all participants agreed to preserve research data in a common repository to reuse and share. There was a need for assistance regarding storage and security of data and there was also a strong response indicating that
researchers required significant funding for this purpose. The respondents also showed interest in the types of services they might need in supporting RDM.

This study opens a window for the future research work in the field of RDM practices of faculty members in the universities of Pakistan while creating an awareness regarding RDM practices. Also, it will help librarians to understand the issues of faculty members in this regard and plan to design special services to them. Moreover, the study will be helpful for university administration to recognize the need for creating a data policy and repository in order to implement research data management and its practices.

Limitations of the Study

The sample consisted of the faculty members from the University of the Punjab only. Due to limited time and resources available for the study, it was hard to have a large and extensive qualitative survey. Hence, the findings may be indicative of the relevant faculties practice and further studies on their data literacy skills and practices are needed. RDM practices of researchers of various disciplines should be carried out to determine their practices.

REFERENCES


Delasalle, J. (2013). Research Data Management at the University of Warwick: recent steps towards a joined-up approach at a UK university. LIBREAS. Library Ideas(23).


Smith II, P. L. (2014). *Exploring the data management and curation (DMC) practices of scientists in research labs within a research university.* The Florida State University.


