Research Data Management: A proposed framework to boost research in Higher Educational Institutes

Bhojaraju Gunjal\(^1\) and Panorea Gaitanou\(^2\)

Abstract

This paper attempts to present a brief overview of several Research Data Management (RDM) issues and a detailed literature review regarding the RDM aspects adopted in libraries globally. Furthermore, it will describe several tendencies concerning the management of repository tools for research data, as well as the challenges in implementing the RDM. The proper planned training and skill development for all stakeholders by mentors to train both staff and users are some of the issues that need to be considered to enhance the RDM process. An effort will be also made to present the suitable policies and workflows along with the adoption of best practices in RDM, so as to boost the research process in an organisation. This study will showcase the implementation of RDM processes in the Higher Educational Institute of India, referring particularly to the Central Library @ NIT Rourkela in Odisha, India with a proposed framework. Finally, this study will also propose an area of opportunities that can boost research activities in the Institute.

Keywords:
Research Data Management, Knowledge Organisation Systems, Data Sharing Policy, Information Retrieval, Higher Educational Institutes

Introduction

The advent of new technologies along with the development of several research data management (RDM) tools has led to a great revolution in the automation and digitization in libraries, which aim to provide innovative value added services to their patrons. At the same time, the adoption of various policy frameworks for managing data and workflow systems along with other Knowledge Organisation Systems (KOS) such as taxonomies, ontologies etc. that enable research data interoperability and enhance information retrieval, pose challenges to the information professionals within the library context.

The importance of RDM is increasingly recognized by several organizations and institutions around the world, as it plays a crucial role in the documentation, curation and preservation of research data. Therefore, it is natural that libraries are considered as a critical stakeholder in the RDM landscape. Their role is highly related to the following: RDM policy development, advocacy and awareness, patrons training, advisory services, data repository development etc. [1].

The management of research data has recently emerged as a strategic priority for universities [2], as they are trying to figure out the ways that research data should be supported, either in terms of advice and training or infrastructure for storage, sharing and curation [3]. 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” [3]. Globally, there are several cases for libraries participating in research projects, in order to enhance research data retrieval and dissemination. Their main aim is to present practices regarding the research data management as part of the research work. Also, an important target for libraries is to gather information and learn about research data processes. The paper presents an overview of the literature review regarding RDM, as well as issues regarding its management and curation. Next, the paper refers to several aspects regarding research data processes and techniques. Finally, it showcases the implementation of RDM in the Technical Institute of India, referring particularly to the Central Library @ NIT Rourkela, Odisha, India with a proposed framework.

**Literature Review**

There are several efforts highlighting different data policies, at several levels, which aim to create best practices regarding the curation of research data. Globally, there are a lot of cases of libraries that participate in research projects, such as:

- The Research Data Pilot Project at the Helsinki University Library [4] that aims to gather information about the current practices and issues of researchers while collecting, using and storing research data and mainly explore all the possible ways in which the library could contribute to the data management process.
- The DataRes Project [5], a project funded by the Institute for Museum and Library Services (IMLS), which examines the perceptions of library professionals faced with supporting federal funding agency mandates for research data management plans, etc.
- The re3data.org Registry Making Research Data Repositories [6], a well-known project, which indexes research data repositories and offers researchers, funding organizations, libraries and publishers an overview of the heterogeneous research data repository landscape, etc.
- The Leeds Building Capacity Project (LBC) [7], which aims to apply the outputs and outcomes from existing JISC projects, services and intelligence to enhance research, learning and teaching at Leeds.
- The ADMIRAL Project [8] at Oxford University that developed a pilot data management infrastructure for life science researchers.
- The FISHnet Project [9], which developed a platform for research data curation and sharing in freshwater biology.
- The Infrastructure for Integration in Structural Sciences (I2S2) Project [10], funded under the Research Data Management Infrastructure strand of the JISC’s Managing Research Data Programme, aims to understand and identify the requirements for a data-driven research infrastructure in the Structural Sciences and focuses particularly on the domain of Chemistry.
- The Sudamih project at Oxford University [11] developed a ‘database-as-a-service’ platform to support humanities researchers.
- The joint German/Brazilian project on Mangrove Dynamics and Management (MADAM) [12] aims to contribute significantly to a sustainable resource management, based on research and continuous long-term survey.

The paper also reports researches carried out to survey in detail the ways in which libraries are currently involved in RDM and the extent to which the development of RDM services is a strategic
priority for them. Pinfield et al. (2014) analyzed the study on RDM and libraries through interviews of library staff from different UK institutions [13]. Their research showed that although libraries play a significant role in RDM, there is uncertainty and variation in the relationship with other stakeholders, such as IT services and research support offices. Delserone [14], Henty [15] and Lewis [16] set out the case for library involvement. Corrall and Cox [17] and Verbaan and Sen [18] further discussed the range of possible roles. Lyon [19] identified a number of opportunities for libraries and a lot of challenges in developing the capacity and capabilities to carry out RDM as well. Pryor [20] summarises these opportunities and challenges by talking about “the re-purposed librarian”, while Procter et al. [21] put libraries in a wider institutional context, emphasizing on their need to collaborate with IT services and academic staff. Finally, it worth’s mentioning SERCIDA [22], a strategic project for supporting the cooperation and exchange of knowledge between the EU countries associated within the Council of European Social Sciences Data Archives (CESSDA) and the Western Balkan Countries (WBC) in the field of social science data archiving.

Moreover, some of the data services that cater to RDM are: the Norwegian Centre for Research Data (NSD) [23], one of the world’s largest archives for research data of its kind that provides data to researchers and students in Norway and abroad. Also, the Social Science Data Archive (SSDA) [24] at University of Ljubljana offers access to data for social science analysis, with emphasis on problems related to Slovenian society. Similarly, the UK Data Archive (UKD) [25], UK’s largest collection of digital research data in the social sciences and humanities. At last, the Swedish National Data Service (SND) [26] is an infrastructure for Swedish research within the humanities, social sciences and health sciences.

Research Data Management: An Overview

Research Data Management (RDM) concerns the organisation of data, from its entry to the research cycle through the dissemination and archiving of valuable results. It aims to ensure reliable verification of results and permits new and innovative research built on existing information [2]. The University of Leicester refers to the RDM process, as follows [27]:

- **Create** data and plan for its use,
- **Organise**, structure and name data,
- **Keep** it – make it secure, provide access, store and back it up,
- **Find** information resources and **share** with collaborators and more broadly, publish and get cited.

Table 1 illustrates some of the most important aspects of RDM.

<table>
<thead>
<tr>
<th>Aspects of RDM</th>
<th>Description - URLs</th>
</tr>
</thead>
<tbody>
<tr>
<td>RDM Basics and Overview</td>
<td>• A step-by-step guide to RDM by JISC. - <a href="https://www.jisc.ac.uk/guides/research-data-management">https://www.jisc.ac.uk/guides/research-data-management</a></td>
</tr>
<tr>
<td></td>
<td>• IASSIST Resources - <a href="http://www.iassistdata.org/resources">http://www.iassistdata.org/resources</a></td>
</tr>
<tr>
<td></td>
<td>• University of British Columbia, Canada - <a href="http://researchdata.library.ubc.ca">http://researchdata.library.ubc.ca</a></td>
</tr>
<tr>
<td></td>
<td>• Research Data Management (RDM) Roadmap - University of Edinburgh, UK -</td>
</tr>
</tbody>
</table>

| **Funders’ Data Plan Requirements and Policies** | Digital Curation Centre (DCC)  
- [http://www.dcc.ac.uk/resources/data-management-plans/funders-requirements](http://www.dcc.ac.uk/resources/data-management-plans/funders-requirements)  
| **Data Management Plans (DMP)** |  
- Funding bodies require grant-holders to develop and implement Data Management and Sharing Plans - [http://www.dcc.ac.uk/resources/data-management-plans](http://www.dcc.ac.uk/resources/data-management-plans) |
| **Policies and Guidelines** |  
- Institutional data policies in UK are listed by Digital Curation Centre (DCC) website to help institutes to manage research data effectively - [http://www.dcc.ac.uk/resources/policy-and-legal/institutional-data-policies](http://www.dcc.ac.uk/resources/policy-and-legal/institutional-data-policies)  
- OSTP Data Access Plan, University of Michigan - [http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/ostp.html](http://www.icpsr.umich.edu.icpsrweb/content/datamanagement/ostp.html)  
- OpenAIRE - Open Access Policies in the European Union - [https://www.openaire.eu/oa-policies-mandates](https://www.openaire.eu/oa-policies-mandates) |
| **Tools and Technologies** |  
- DCC provides tools and applications list – [http://www.dcc.ac.uk/resources/tools-and-applications](http://www.dcc.ac.uk/resources/tools-and-applications)  
- SSDA - [http://nesstar2.adp.fdv.uni-lj.si/webview](http://nesstar2.adp.fdv.uni-lj.si/webview) |
| **Future Directions of RDM** |  
- Directions in research data management for UK universities outline a vision of where the sector should aim to be in five years’ time. Five key areas for actions are (i) Policy development and implementation; (ii) Skills and capabilities; (iii) Infrastructure and interoperability; (iv) Incentives for researchers and support stakeholders and (v) Business case and sustainability - [http://repository.jisc.ac.uk/5951/4/JR0034_RDM_report_200315_v5.pdf](http://repository.jisc.ac.uk/5951/4/JR0034_RDM_report_200315_v5.pdf) |
| **Training Programs on RDM** |  
- MANTRA - a free online course at University of Edinburgh (UoE), UK - [http://datalib.edina.ac.uk/mantra](http://datalib.edina.ac.uk/mantra)  
- Melbourne_MANTRA - an online training program in RDM based on UoE - [http://library.unimelb.edu.au/digitalscholarship/training_and_outreach/mantra2](http://library.unimelb.edu.au/digitalscholarship/training_and_outreach/mantra2) |
| **Best Practices for RDM** |  
| Registry and Directory | • The registry of research data repositories shows what repositories exist in a subject discipline - [http://www.re3data.org](http://www.re3data.org)  
• Registry of Open Access Repositories (ROAR) hosted at the University of Southampton, UK - [http://roar.eprints.org](http://roar.eprints.org)  
• OpenDOAR is an authoritative directory of academic open access repositories [http://www.opendoar.org](http://www.opendoar.org) |
| Open Access Policies and Tools | • Tools for research professionals for open access (OA) archiving - [https://www.jisc.ac.uk/sherpa](https://www.jisc.ac.uk/sherpa)  
• Registry of Open Access Repository Mandates and Policies (ROARMAP) - [http://roarmap.eprints.org/](http://roarmap.eprints.org/)  
• Open Research Data Platform, Switzerland - [http://openresearchdata.ch](http://openresearchdata.ch) |
| RDM Status | • RDM status around the World - [http://service.re3data.org/search?query=&countries[]=ind](http://service.re3data.org/search?query=&countries[]=ind) |
| Data Services | • Norwegian Centre for Research Data - [http://www.nsd.uib.no/nsd/english/index.html](http://www.nsd.uib.no/nsd/english/index.html)  
• Arhiv družboslovnih podatkov (Slovenian Social Science Data Archives) - [http://www.adp.fdv.uni-lj.si](http://www.adp.fdv.uni-lj.si)  
• UK Data Archive - [http://data-archive.ac.uk](http://data-archive.ac.uk)  
• Swedish National Data Service - [http://snd.gu.se](http://snd.gu.se) |

**Table 1: Different aspects of RDM**

**RDM in Institutes around the world**

Some of the HEIs have already successfully implemented RDM in their services around the world, such as the University of Edinburgh, which developed an RDM policy and a roadmap, which reports on the plans, achievements and challenges faced, while there was an attempt to bring the University of Edinburgh RDM Roadmap to fruition [33]. Similarly, Hodson and Jones [34] provide in their blog the “Seven rules of successful research data management in universities”. Moreover, Jones et al [35] provide a guide in which they propose ways to develop Research Data Management Services in HEIs. More specifically, the purpose of this guide is to help institutions understand the key aims and issues associated with planning and implementing research data management (RDM) services. In addition, the Tufts University has implemented a Research Data Management System (RDMS). As mentioned by Professor Diane L. Souvaine, Project Sponsor and Vice Provost for Research, “RDMS will not only ensure that Tufts is compliant with federal regulations and guidelines but also provide a strategy and the tools that support a much needed University-wide data management service towards enhancing research across the Tufts community”[36]. Several other universities, such as University of Virginia [37], the National Coordination Point Research Data Management, which facilitates a national strategy for research data management in the Netherlands [38], Cornell University [39]...
In addition, some RDM related efforts in India are the following: the KRISHI Research Data Management [41], a Knowledge based Resources Information Systems Hub for Innovations in agriculture and an initiative of Indian Council of Agricultural Research (ICAR) and Dora and Kumar [42], present an implementation at Indian Institute of Management, Ahmedabad which can be used for creating and preserving research data.

RDM @ NITR: A proposed framework to boost research

The Institute was founded as the Regional Engineering College, Rourkela (REC) on August 15th 1961 under the joint management of the Government of India and Government of Orissa for imparting quality education in the field of Engineering and Technology. The college was designated, on June 26th 2002, as the National Institute of Technology, a fully autonomous Institute supported by the Ministry of Human Resource Development, Government of India [29]. The Biju Patnaik Central Library (BPCL) is the central library of the NITR and caters particularly to the educational and research needs of the academic community, while its resources are consulted by scholars from all over the country. More information regarding the BPCL can be seen at [30].

Managing research data is an important aspect of the scholarly process and for many libraries, such as BPCL, relates to new roles and responsibilities they will potentially assume in fulfilling their traditional mission of collecting and sustaining the scholarly record. A plan on organising, storing and accessing such contents is accomplished in the BPCL with the use and implementation of appropriate tools and techniques, the adoption of well defined policies, the proper training for the stakeholders, etc. Within this framework, in BPLC a variety of services emerge, which mostly aim at supporting scholars’ research data management needs. The main role of the Library should be to educate scholars on the benefits of managing research data, besides providing repositories where data can be stored, preserved and accessed. Therefore, we can see that the Library moves to researcher-focused services. Figure 1 illustrates all the types of research data and services supported by the library staff.
One of the most important databases maintained in the library is the DSpace@NITR [31] for the Scholarly publications (see Figure 2). This database includes mainly Book Chapters, Journal Articles and Conference Proceedings. Similarly, a separated repository that also implements DSpace is created to store all official key documents generated at NITR. Moreover, EThesis@NITR [32] is the institutional repository, which includes the Electronic Theses & Dissertations from the academic fraternity (see Figure 3).

Figure 2. DSpace @ NITR
Proposal to boost research @ NITR

With respect to RDM, the following proposal is suggested to boost research at NITR. As shown in Figure 1, there is an ample of opportunities in the Library to establish a data repository like other repositories implemented to store and access all scholarly output of the NITR fraternity. From the Sponsored Projects section, data related to all projects may be acquired and made available to the users by framing appropriate policies. This data needs to be organised by framing and adopting an appropriate data policy as presented below.

Figure 3. EThesis @ NITR

Figure 4: Lifecycle of Research Data and its management
Research data generated from the R&D activities and Sponsored Projects needs to be planned covering RDM services, plans, opportunities or challenges faced at the institution. This will provide more opportunities in several ways, such as (i) by minimizing the repetition of works by enhancing the re-use of the existing data; (ii) by providing opportunities for research collaboration; (iii) by extending research projects and so on. With the adoption of such data policies and strategies, this study proposes to adopt RDM.

Finally, Figure 5 illustrates the proposed framework to implement RDM at NITR.

‘Data Sources’ contain data from different sources, generated from various research and project data, therefore it refers mainly to the scholarly output from the university’s community. ‘Data Collection’ is the data input from various sources, which are collected from different sections, such as academics, registry, research cell or from users directly need to be collected for its proper storage. In ‘Data Storage’ all data collected with the help of the library staff and computer centre are stored in various platforms with the implementation of various tools, such as DSpace, ePrints, etc. The data can be organised by adapting a proper metadata schema to enable easy access. ‘Data Access’: The stored data can be made easily accessible by the internal and external users through various platforms. With the use of a proper data management plan and appropriate tools, the implementation of RDM will yield an effective usage of the system.

The following section will present several of the issues and challenges most of the institutions face, including NITR as well. Benefits of such opportunities are considered and implemented at NITR and adoption of a few more aspects is under progress to suit the needs of the Institute.
Issues and Challenges in RDM
As already mentioned, the management of research data is a major challenge for Higher Educational institutions (HEI). Huge volumes of born-digital data are being produced and need to be managed accordingly for proper use by the academic fraternity. Cox and Pinfield (2013) framed a library-oriented model of institutional RDM, which helps to clarify the different issues involved in the RDM challenge by identifying various layers of activity, multiple stakeholders and a large number of factors that influence the implementation of any programme [3]. Halbert [28] suggested the following factors that may hinder the progress of the RDM process:

- Lack of Funding
- Lack of Organizational Structures
- Lack of Professional Preparation
- Lack of Priority among Researchers
- Lack of Institutional Mandates.

In addition, there are several challenges to be taken care of by libraries in order to adopt RDM:

- Adoption of policies, such as Archival Policies, embargo, access rights, etc.
- User training of library staff
- Training for stakeholders
- Support from stakeholders, such as technical expertise and academic fraternity
- Voluntary submission of publications to the repositories
- Data curation
- Database creation
- Compliance of policies with funding agencies
- Upgradation of features and software
- Culture change
- Change management.

Opportunities to Boost Research in HEIs
Despite the aforesaid challenges, there are possible opportunities that can help to overcome and boost research in HEIs in many ways. Some of these elements that can make difference and boost research in HEIs are the following:

a. Policy framework: The adoption of policies such as access rights, etc. is very much essential for HEIs for a successful RDM.

b. Adoption of KOS Aspects: The adoption of various policy frameworks for managing data and workflow systems along with other Knowledge Organisation Systems (KOS), such as taxonomies, ontologies etc. that enable the interoperability of research data and enhance information retrieval (IR) and these tools will enhance the user experience as well.

c. Mapping of experts and data/content: With the proper use of techniques, the mapping of experts with the stored data/content will help educational institutions to conduct research to progress in the next level at higher rate.

d. Awareness and re-use of data/content: With the proper credits to the source, the researchers can make re-use of the existing data to enhance awareness of such resources at early stage enhancing at the same time the research progress.
e. **Discovery and analysis tools:** These tools can help researchers to discover and identify the required research data from various resources including social media. They also showcase the value of their research from various platforms presenting the ways being utilised across the globe.

f. **Research Support Service by the library:** The library should provide research support services to their scholars and faculty in a much more effective way to boost the research activities in HEIs. Along with the provided library services, the library staff should be capable enough to help researchers in providing orientation about latest tools and techniques of research, customised resources in the area of research, etc.

**Conclusions**

Research data management is a complex issue involving multiple activities carried out by various actors addressing a range of drivers and influenced by a large set of factors [13]. Currently, it seems that there is a tendency that libraries are moving towards developing new institutional RDM policies and services considering it as an important part of their future role. Institutional policy development is required as a basis for coordinated action on data management [2]. The paper presented as a case study the Central Library @ NIT Rourkela in Odisha, India, which illustrated in brief the types of benefits to be gained by managing the research data.

**References**


10. I2S2 Pilot Implementation. Available at: http://www.sdu.tees.ac.uk/lbc/index.php?p=display-project-details&project_id=33&search=yes&tag=11
23. NSD. Available at: http://www.nsd.uib.no/nsd/english/index.html
24. SSDA. Available at: http://www.adp.fdv.uni-lj.si/eng
25. UKD. Available at: http://data-archive.ac.uk
26. SND. Available at: https://snd.gu.se/en
27. University of Leicester. What is Research Data Management. Available at: http://www2.le.ac.uk/services/research-data/rdm/what-is-rdm
29. NITR. Available at http://nitrkl.ac.in
30. BPCL. Available at http://library.nitrkl.ac.in
31. DSpace@NITR. Available at http://dspace.nitrkl.ac.in/dspace
32. EThesis@NITR. Available at http://ethesis.nitrkl.ac.in


---

**End-notes**

1 Dr. Bhojaraju Gunjal is the Head & Deputy Librarian at Central Library of National Institute of Technology, Rourkela, Odisha, India, e-mail: bhojaraju.g@gmail.com

2 Dr. Panorea Gaitanou is member of the Database & Information Systems Group (DBIS), Department of Archives, Library Science and Museum Studies, Ionian University, Corfu, Greece, e-mail: rgaitanou@ionio.gr