Indian Academic Research Data Repository (IARDR) With INFLIBNET: A Futuristic Plan

Birender Pal ***Sanjay Kumar Singh***

Abstract

Data and Information are an integral component resulting to form the knowledge. In terms of research information management (e-resource), the Institutional Repository (IR) has played an immense role. Research data is another important asset of research which needs to be properly managed. The situations of research data management in European, American and Australian continents is very much advanced and are having a well research data management platform for their researcher. But when we study the research data management aspect basically in Indian context, then the situation is quite different than the other countries of the world. In Indian universities, a lot of research data has been collected by the researcher which is unutilized after the completion of research. This paper explores that what is the present status of research data management policy in Indian context. It also highlights on how India can have the research data management policy in the field of academics. It further reflects the proposed policy on how INFLIBNET centre can act as a national platform or nodal centre to build the anticipated IARDR. The paper at the end shows the merit and limitation of the proposed IARDR and the role of libraries.

Keywords: Data Repository, IARDR, INFLIBNET, NDSAP, NPE-2019, Open Access

1. Introduction

The present scenario of Indian higher education system has gone through a lot of development. Government of India is giving prime importance to the quality level of Research. UGC has already been assigned the task by Government of India to make the National Academic Depository (NAD) in order to store and provide access to the research degrees (UGC, 23/09/2019). UGC has further taken a serious note on the quality level of PhD thesis and as a result taken a forward step to check the quality level of Thesis that has been awarded for the last ten years (UGC, 22/09/2019). Here the role of Shodhganga repository will play a very significant role for the researcher who will undertake the task. Today, this task is possible because of the open access policy of depositing the thesis by UGC. Another instance that needs to be given primary importance is the research data in various forms that has been used by the researcher to complete his PhD Thesis. The time of every Researcher is very precious and before initiating the study he/she has to dedicate their time for literature search. Here the institutional repositories or the national repository plays a vital role. At the same time the aspect of research data is also an important aspect of research. Data collection by the researcher requires lot of time. The data which is quantitative in nature and permanent is observed to be collected by various

researchers. Sometimes the same data is collected by another researcher as primary data. So, the primary Research Data which is permanent in nature and is collected by the researcher needs to be preserved and provide access so that once it is published in the open access data repository then it becomes secondary in nature.

2. Institutional Repository to Data Repository

With the advancement in Artificial Intelligence, the role of data management will bring a smart change in the society. The open access movement has resulted for the creation of Institutional Repository in the different research level Institution of the world. Research institutions framed the policies for archiving of the research information in open access mode. Another step that has been taken by many research institutes is to understand equally the importance of research data. These institutes equally frame the guideline for open archiving of research data and are managing both the institutional repository as well as the data repository. Even in a few cases a single digital repository acts as a hub for both the institutional repository and data repository

2.1 Institutional Repository vs Data Repository

Institutional Repository (IR)	Data Repository (DR)
IR may basically contains the full text (entire work) as well as the data	DR contains the supplement data of full text
The purpose of IR in research is to manage the research information.	The purpose of DR in research is to manage the research data.

3. Types of Data Repository

Research Data Management and Research Institutional Repository Management are the two important repositories which need to be understood very carefully. A number of repositories use their institutional repositories to preserve both their full text research files as well as the research data in form of datasets. But many repositories are those repositories which fully concentrate only on the research data management aspect. Also there are institutions which maintain two different types of repositories for eg. University of North Texas (UNT) Libraries maintains two different repositories viz one for data repository and other for scholarly work (University of North Texas, 2019). So, based on the nature of different research repositories and its collections, following three types of data repositories are classified

- ❖ Data Repository (Datasets)
- ❖ Institutional Repository cum Data repository (both full text and datasets)
- ❖ Institutional to Data Repository(mapping of full text with dataset)

4. Research Data Repositories in the World

As per the data collected from Registry of Research data repository (https://www.re3data.org), a total 2399 numbers of data repositories are available in the registry as on date at the time of writing this paper (Berlin School of Library and Information Science and others). As per the registry, the maximum number of data repositories are from United States (1044) followed by Germany (370).

Table 1: Country wise Research Data Depository of the World Universities

Sl No.	Country	No of Research Data Repository of Universities
1.	Australia	09
2.	Austria	01
3.	Brazil	01
4.	Canada	01
5.	Canada International	05
6.	Chile	01
7.	China	01
8.	China, International	01
9.	Denmark, European Union	01
10.	Germany	07
11.	Germany, International	01
12.	Italy	02
13.	Lithuania	01
14.	Netherland	02
15.	New Zealand, International	01
16.	Norway	02
17.	Singapore, International	01
18.	South Africa	01
19.	Switzerland	01
20.	United Kingdom	26
21.	United Kingdom, International	03
22.	United States	10
23.	United States, Canada, International	16
24.	United States, International	03
25.	United States, Norway	01

Source: After Browsing 2399 data repositories of Registry of Research data repository (https://www.re3data.org)

After analyzing the 2399 data repository from the Registry of Research data repository, 99 numbers of repositories are those which are managed by the universities for its researcher in depositing and providing access to their research data. The Table 1 indicates that the trend of maintaining research data repository is basically prepared by American, European and Australian countries. The universities of this continent have realized the importance of research data repository. It is also observed that the trend of Asian countries is at a very nascent stage. None of the Indian universities research data repository is found to be registered in the Registry of Research data repository.

4.1 Policy on Research Data Management in Global context

Many Universities of the world are having a Research management policy in terms of managing the research data. The University of Liverpool in their data management policies guided a rule that the ownership of all research data of all academics or post graduate students will be the University of Liverpool (University of Liverpool, 2019). Further, the European commission on H2020 funded research states that the research publications and the research data must be publically uploaded on the open access repositories (European Commission, 2019). Following are the examples of three countries which have a national archive and discovery platform for research data.

1. Midas is the national data archive platform for research data of Lithuania working on the

European commission guideline on open access (Vilnius University, 2019).

2. Research Data Australia is another data discovery platform for research data of Australian researcher. (Monash University, 2019).

3. FRDR (Federated Research Data Repository) is another example where Canadian researcher can submit their research data (Portage and others, 2019)

5. Why India Need IARDR

Nishtha Anilkumar has carried a study on 15 national research/ academic institutes in India where it was found that research data management in libraries is at very nascent stage (Anilkumar, Nishtha, 2019). After verifying the data archiving status of these fifteen institutes, it was found that except ICSSR data service set up by INFLIBNET, none of the institute data repositories is registered with the Registry of Research data repository (re3data.org registry).

Further the result of Table: 1 reveals that out of fifty numbers of Indian data repositories registered in the registry, none of the repositories are from the Indian universities. This indicates that there is an important need to have an institutional policy for data repositories and its implementation at all the research institutes of India.

6. Indian National Data management policy

6.1 Policy of Government of India

The gazette notification of Government of India dated March 17, 2012 has published the National Data sharing and Accessibility Policy-2012 (NDSAP) according to which all the government funded non sensitive data is to be shared for better and effective decision making for future (Ministry of Science and Technology, Government of India, 2019). The data needs to be shared through the open government data platform in an open format so that it can be shared and access easily (Ministry of Science and Technology, Government of India, 2019). In the field of education, the draft National Educational Policy-2019 has proposed for National Repository of Educational Data (NRED) for sharing the educational information (MHRD, 2019). For the research purpose, even the draft NPE-2019 in its paragraph 19.2.4 has given prime importance to educational data at par with the policy of open data initiative keeping the data security at its prime concern (MHRD, 2019).

6.2 Policy of DST and DBT, GOI

In Indian context, the guidelines on open access of the Department of Science & Technology (DST) and Department of Biotechnology (DBT) -2014 clearly stated that the funded projects of the DBT and DST where the researcher published their research papers in different peer reviewed journals must have to upload the paper in the Institutional Repository (Ministry of Science and Technology, Government of India, 2019). But still the concept of research data management in the DST and DBT research projects guideline-2014 do not have any clear conception on the Research Data Management. Recently in July, 2019, DBT, Ministry of Science and Technology, Govt of India has published a draft framework for Biological Data Storage, Access and Sharing Policy of India (Ministry of Science and Technology, Government of India, 2019). This shows that the policy for data management especially the research data is going to become a reality very soon for the DBT funded projects and research.

6.3 UGC Policy of Open Access of Research

The NDSAP policy has its own limitation as it does not cover the self funded research where research data are collected by the researcher. Even, before the INFLIBNET Shodhganga programme, the full text PhD thesis of the researcher was confined within the four walls of their respective institute libraries. Later, the Union database of Thesis programme of INFLIBNET (IndCat) has shared the metadata description of the Thesis at the National level. In 2009, UGC has made compulsory for submitting the electronic copy of the complete Thesis of Researcher to the university and same should be uploaded in Shodhganga. But for the research data management of the universities at national level there is no such policy developed up till now. And as a result, research data (datasets) which is collected by the researcher are either preserved by the researcher in his own custody or sometimes deposits it to their institute which is very rare in Indian context. But when we understand the philosophy of Open Source, Open Access and Research Data Management (in open access) then one common concept prevails from all the three subjects is the concept of "Open" which means anyone can use, reuse and redistribute.

7. INFLIBNET and RESEARCH

INFLIBNET centre has been continuously engaged in providing a national platform to the research work and to the researcher through its different services like UGC Infonet Network connectivity (now merged with NKN), UGC Infonet Digital Library Consortium (now e-ShodhSindhu), Shodhganga and Shodhgangotri programme, In the later stage INFLIBNET centre has been entrusted by UGC to work on the areas on improving research be it the plagiarism software aspect or the UGC CARE list. INFLIBNET centre is a member in UGC-CARE list and also the director of INFLIBNET is a member of UGC Care Empowered Committee. Recently on 21st September, 2019, the Plagiarism Detection Software (PDS) "**ShodhShuddhi**"programme of INFLIBNET was launched to provide the plagiarism detection software to the research institutions (INFLIBNET, 2019). Moreover, the Vidwan database, the IRINS (Indian Research Information Network System) are the important programme of INFLIBNET which share and network the research platform to a national and global level in an organized way.

Table 2: INFLIBNET and its impact on Researcher

Sl No	Programme	Impact on Researcher
1.	Library Automation	Researcher access the record of respective library collection.
2.	IndCat	Researcher access the record of collections of other libraries.
3.	UGC Infonet Network connectivity	Researcher is introduced to the Internet.
4.	UGC Infonet Digital Library Consortium	Researcher gets the access to e-resource.
5.	Shodhganga and Shodhgangotri	Researcher improves and deposits the research work.
6.	VIDWAN	Connecting Researcher.

7.	IRINS	Researcher get the Research Information Management (RIM) service
8.	Research Project database	Researcher gains the information on funded research
9.	UGC CARE List	Researcher get the platform to publish research in fair publications
10.	ICSSR Data Service	Social Science Researcher get the access of data repository
11.	ShodhShuddhi	Researcher research capacity will be enhanced (predicted).

8. Research Data Management Policy (Proposed)

A need of the hour is to frame a Research Data Management policy in Academics. Here the role of MHRD is very important. Following are the proposed policy for the Academic Research Data Management Policy of India.

- MHRD may frame a centralized guideline through different bodies like UGC, AICTE for all higher educational institutions to collect and preserve the research data of the researcher.
- A national academic research data portal (IARDR) may be developed by INFIBNET centre.
- It will be mandatory for every researcher to deposit their research data to the university libraries. There may be two categories of data
  - a Personal information and sensitive data
  - b. Non-Sensitive data
- Type A data should be strictly restricted from uploading and access and Type B data can be free flow in open access mode for other researcher and public
- The IndCat, Shodhganga and the proposed IARDR should be mapped with each other. In this regard, the initiative is already taken by INFLIBNET to map IndCat and Shodhganga

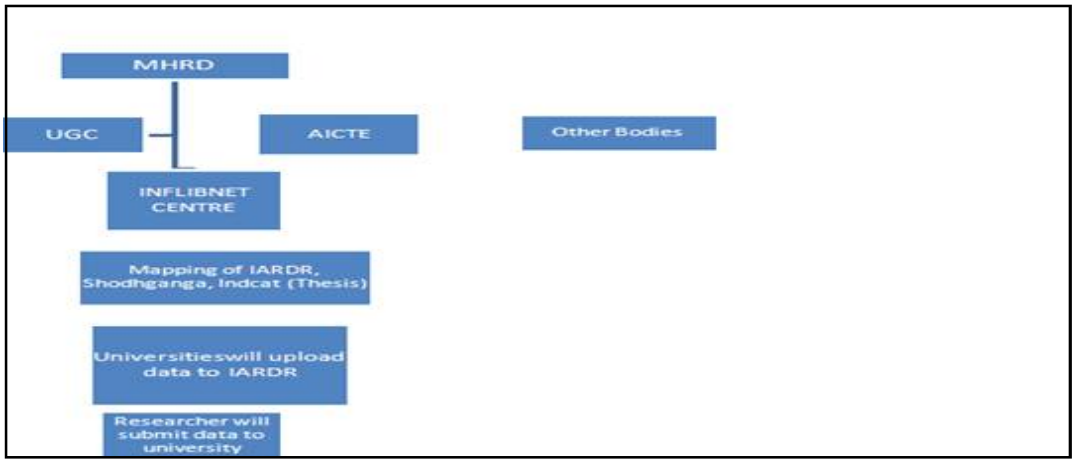

Fig 1: Layout of the Proposed IARDR

8.1 Benefit of National Academic Research Data Repository

- The proposed national academic research data repository will be a data conversion platform from big data to smart data with the integration of Artificial intelligence technology.
- Researcher may get the citation of not only for his research content but also for the research data he/she has collected.
- It may act as a supporting evidence of the PhD research work of the researcher.
- It may act as another parameter to check the quality level of the PhD thesis.
- It will save the time of the researcher from collecting the same data from the field.
- It will help the decision makers to take appropriate decision on a particular field based on the evidence.

8.2 Limitation of National Academic Research Data Repository

- The researcher may not share the research data without having a national research data depositing policy in the field of academics.
- if proper Research data is not collected at ground level then there will be chances that other research works which has used the research data from the repository may lose its quality at some level.
- Due to absence of mechanism of plagiarism check on research data, there will be chances of repeatedly depositing the same data.

8.3 Role of Libraries

Library professionals are very much well aware about the systematic organization of its library resources. Document management, Record management, e-resource management, assigning metadata all are different levels of works performed by the library professional. The open access and the Research have reengineered the role of libraries where the research output (e-resource) is managed with institutional repository and the research based data are further managed with research based data repository. In Shodhganga, the libraries are playing a very significant role. Libraries and Information centre of the different world universities where there is research data repository has been playing the leading role in collecting, preserving and providing access to the research data. Similar kind of role can be further expected from the Indian university libraries when initiating the task of research data management.

9. Conclusion

Today India is very much advanced in terms of technology which is further moving towards smart technology. The draft NPE-2019 has cited that data is the key fuel for Artificial Intelligence (AI) technologies (MHRD, 2019) and so when the AI technology will penetrate into Indian research then there will be gap in research data. So, Research data repository in the field of Higher education system of India is the need of the hour. The history of each and every services of INFLIBNET centre reveals the vision and mission of the organization. INFLIBNET and the university libraries are already connected with different services and today the bonding between the two is very strong. This opportunity can be utilized for the proposed IARDR

where the role of university libraries will be very crucial. As an IUC of UGC the INFLIBNET Centre is doing a great job in the academic and technological development for the academic fraternity in India.

References:

1. UGC. National Academic Repository. Available at https://www.ugc.ac.in/pdfnews/9536043_NAD-MHRD.pdf(Accessed on 22/09/2019).
2. UGC. Quality of PhD Thesis in Indian Universities. Available at https://www.ugc.ac.in/pdfnews/7172274_Public-Notice-Study-on-Quality-of-PhD-Theses-in-Univ.pdf(Accessed on 23/09/2019).
3. University of North Texas. Data RepositoriesServices. Available at https://library.unt.edu/services/data-repository/ (Accessed on 22/09/2019).
4. Berlin School of Library and Information Science,GFZ German Research Centre for Geosciences, Karlsruhe Institute of Technology Library and Purdue University Libraries. Registry of Research data repository Available at https://www.re3data.org/search?page=96(Accessed on 26/09/2019).
5. University of Liverpool.Research Data Management Policy.Available at https://www.liverpool.ac.uk/media/livacuk/computingservices/research-data-management/researchdatamanagementpolicy.pdf(Accessed on 26/09/2019).
6. European Commission. H2020 Programme: Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020. Available at https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf(Accessed on 26/09/2019).
7. Vilnius University. About MIDAS / Information about MIDAS projectAvailable at https://www.midas.lt/public-app.html#/about(Accessed on 26/09/2019).
8. Monash University. Research data Australia. Available at https://www.ands.org.au/online-services/research-data-australia(Accessed on 25/09/2019).
9. Portage, Compute Canadaand the Canadian Association of Research Libraries. About FRDR-Federated Research Data Repository (FRDR).Available at https://www.frdr.ca/docs/en/about/(Accessed on 25/09/2019).
10. Ailkumar, Nishita (2018). Research Data Management in India: A Pilot Study. EPJ Web of Conference. Available at https://www.researchgate.net/publication/326991769_Research_Data_Management_in_India_A_Pilot_Study/link/5c6d3a37a6fdcc404ebf3e1f/download (Accessed on 25/09/2019).
11. Anilkumar, Nishita (2018). EPJ Web of Conference
12. Ministry of Science and Technology, Government of India. National Data Sharing and Accessibility Policy(NDSAP-2012). Available at https://data.gov.in/sites/default/files/NDSAP.pdf (Accessed on 24/09/2019).

13. MHRD. Draft National Policy Education Policy. Available at https://mhrd.gov.in/sites/upload_files/mhrd/files/Draft_NEP_2019_EN_Revised.pdf (Accessed on 25/09/2019).

14. Ministry of Science and Technology, Government of India. DBT and DST Open Access Policy- Policy on open access to DBT and DST funded research. Available at https://dst.gov.in/sites/default/files/APPROVED%20OPEN%20ACCESS%20POLICY-DBT%26DST%2812.12.2014%29_1.pdf (Accessed on 25/09/2019).

15. Ministry of Science and Technology, Government of India. Biological Data Storage, Access And Sharing Policy of India. Available at https://dst.gov.in/sites/default/files/Draft%201%20-%20Biological%20Data%20Policy.pdf (Accessed on 25/09/2019).

16. UGC. UGC Guidelines for Shodhganga. Available at https://Shodhganga.inflibnet.ac.in/ (Accessed on 22/09/2019).

17. INFLIBNET. ShodhShuddhi. Available at http://pds.inflibnet.ac.in/ (Accessed on 26/09/2019).

About Authors

Dr. Birender Pal, Librarian, Dhakuakhana College, Dhakuakhana, Assam

Prof. Sanjay Kumar Singh, Head, Dept of Lib & Inf Sc., Gauhati University, Guwahati