# Development of Interoperable RDRs System in India

**A Conceptual Proposal**

Manu T R[1]* and Bhakti Gala[2]

[1]Central Library, Indian Institute of Technology Delhi

[2]School of Library and Information Science (SLIS), Central University of Gujarat, Gandhinagar*
Corresponding author email: manutr@library.iitd.ac.in

26th International Symposium on Electronic Theses and Dissertations ETD 2023, Gujarat, India 26-28 October 2023
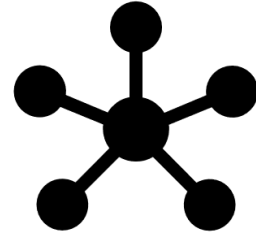
# Coverage

- Introduction

- Literature Review

- Research Data Repositories (RDR)

- Objectives of the Study

- Methods

- Pillars of RDM interoperability in India

- RDM Stakeholders in India

- RDRs in India

- Workflow of RDRs System in India

- Discussion

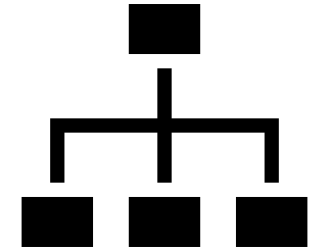- Conclusion

- Reference

2

# Introduction

Over 50+ research data repositories in India

Allows easy access & sharing
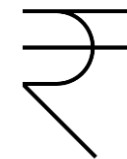
Enhances visibility of research data

Data management of an individual researcher/ institute
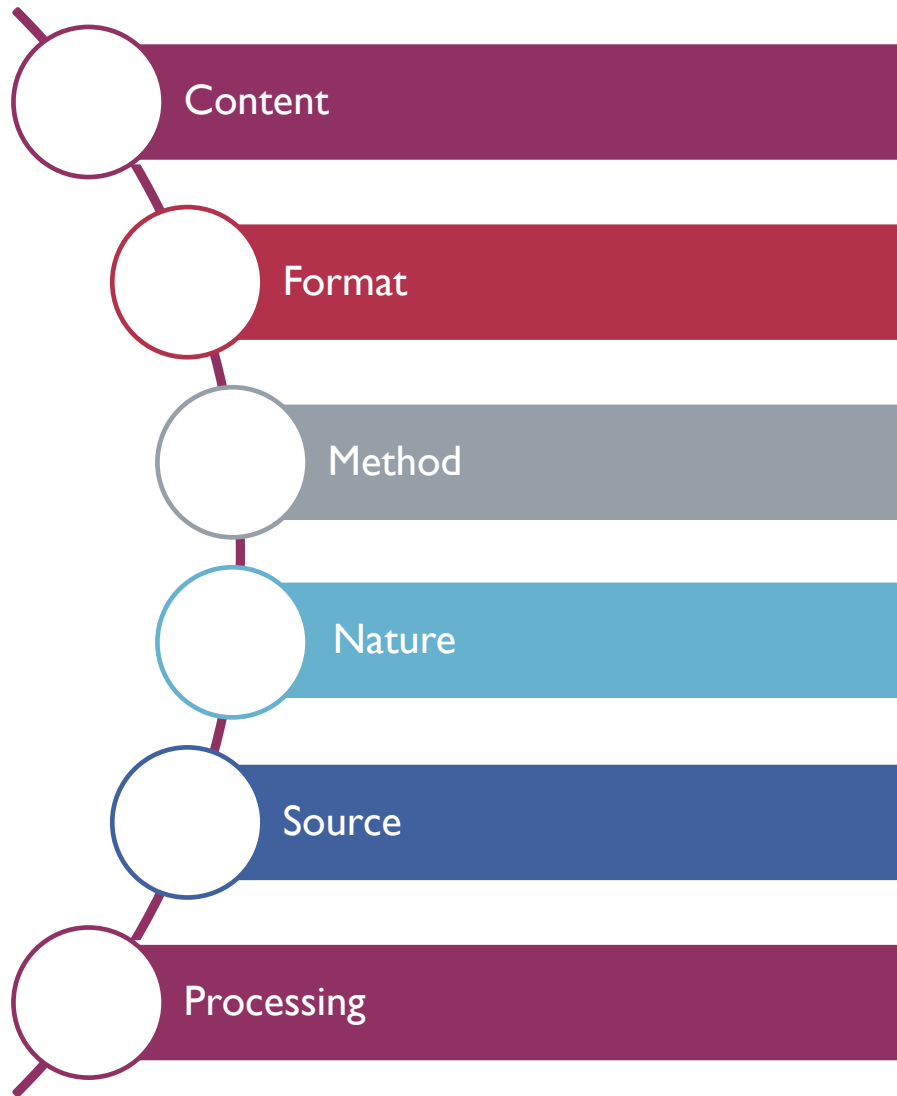
Value for reuse

Avoid duplication of research work

Funding agencies mandatory requirement

# Research Data Repositories (RDR)

- Platform for academic researchers to deposit, share, and access the research support data generated throughout the research lifecycle.

- It is collection of digital objects that are stored, managed, curated research data

- Ensuring the accessibility, preservation, and long-term availability of future access

- Linking between data and publications

- Full-filling regulatory requirements of the scholarly publishers and funding organization

- It can be divided into

  - *Institutional,*

  - *National,*

  - *International,*

  - *Disciplinary/multidisciplinary, and*

  - *Project-specific repositories*

4

# What are Research Data?

Content

Format

Method

Nature

Source

Processing



Some examples of the hundreds of different research data file types

Archived data
zip, rar, iso...

Audio
mp3, wav, aif...

Spreadsheet
csv, xls, tsv...

Documents
doc, pdf, odt...

Text
txt, rtf, bib...

Notebook
ipynb

3D graphic
obj, stl, ply

Molecular
cif, pdb, xyz

Image
jpg, png, svg...

Presentation
ppt, pptx, pptm

Visualisations
gephi, gexf

Geographical & map
keyhole, GIS, glf...

Code
python, r, java...

Video
mp4, mov, avi...

# Research Data Repositories

# Research Data Repositories (RDR) in India

- Over 50 Indian RDRs are listed in re3data.org

- Major subjects cover: *life science, medicine, science, technology, agriculture, humanities, and social sciences*

- ESSO-INCOIS, Hyderabad https://incois.gov.in/

- ICRISAT Dataverse Network https://dataverse.icrisat.org/

- ICSSR Data Service (INFLIBNET) http://www.icssrdataservice.in/

- Krishi (Knowledge-based Resource Information Systems Hub for Innovations in Agriculture) https://krishi.icar.gov.in/

- ESSO- Indian National Centre for Ocean Information Services https://incois.gov.in/portal/index.jsp

- Indian Space Science Data Center (ISSDC) https://www.issdc.gov.in/

- Kodaikanal Solar Observatory (KSO) https://www.iiap.res.in/?q=kodai.htm

- World Data Centre for Geomagnetism http://www.wdciig.res.in/

- Indian Biological Data Centre (IBDC) https://ibdc.rcb.res.in/

# Objective of the study

The objective of the study is to develop an interoperable RDR system in India. The focused objectives are as follows:

- To identify the significant stakeholders of RDR in India

- To identify the pillars and role of these stakeholders in the development of interoperable RDR systems in India.

- To propose a workflow and process of RDR systems in India.

# Methodology

Systematic Literature Review (SLR)

- Case studies

- Working papers

- Research reports

- Data management guides

- Similar conceptual frameworks were studied

- National research data services of UK, Australia, Europe, Canada, Netherlands, Germany etc.

# Findings 1: National Level Research Data Services (Major)


UK Data Service


Canadian Research Data Centre Network


Research Data Australia


Research Data Management - Science Europe


National Research Data Infrastructure Germany


Landelijk Coördinatiepunt Research Data Management (LCRDM) - Netherlands etc.

# Findings 2: RDR stakeholders in India…

- The development of RDR is a collaborative effort, and it requires the involvement of all support units to establish the RDM practices at the institute level to the national level.

- The major stakeholders of RDR involved:-

  - Who created the research data?

  - Who helps to create research data?

  - Who facilitates the management of research data?

  - Who uses the research data?

# Findings 2: RDR stakeholders in India

```
                        RDR
                  Stakeholders in
                      India
```

## Data Creator

- Researchers
- Research supervisor
- Faculties and scientific members
- Principal investigator
- Institutions
- Research funding agencies (India & International)

## Facilitator

- Institute leadership
- Academic departments/schools
- Research and development unit
- Libraries
- Library committees
- Research support unit
- IT support unit
- Administrative office
- Scholarly publishers
- Professional associations and networks

## Government at large

- Ministries of educations
- Ministry of science and technology
- Department of higher education
- Other ministries and higher education departments
- Government funding agencies
- Policy makers

## Users

- Students
- Researchers
- Project team
- Faculty members
- Scientists
- Government officials
- Private entities
- Journalist
- Citizen scientists

12

# Findings 3: Pillars of RDR interoperability in India…

- The RDM pillars need to work collaboratively and should have interoperability to establish sustained RDM services in India.
- Data creators and facilitators at the institutional level would have more responsibility
- Government funding agencies including the Ministry of Education and Higher Education Departments need to build ethical guidelines, policies and regulations.
- The National Research Data Services Centre by the Government of India is required to support the RDM initiatives by the institute and ministries and higher education departments.
- The National Research Data Services Centre will also be responsible for developing policies and imposing them to adopt and mandate to development of RDM services in the institutes in India.
- National research data services should develop the road map towards implementing national RDM services.
- Working with the RDM stakeholders to ensure research data is openly available to support innovation that benefits all research communities.

# Findings 3: Pillars of RDR interoperability in India

National Level
or Government
of India
Initiatives

Ministries &
Higher Education
Departments
Initiatives

Invidual
Institutions and
Organizations
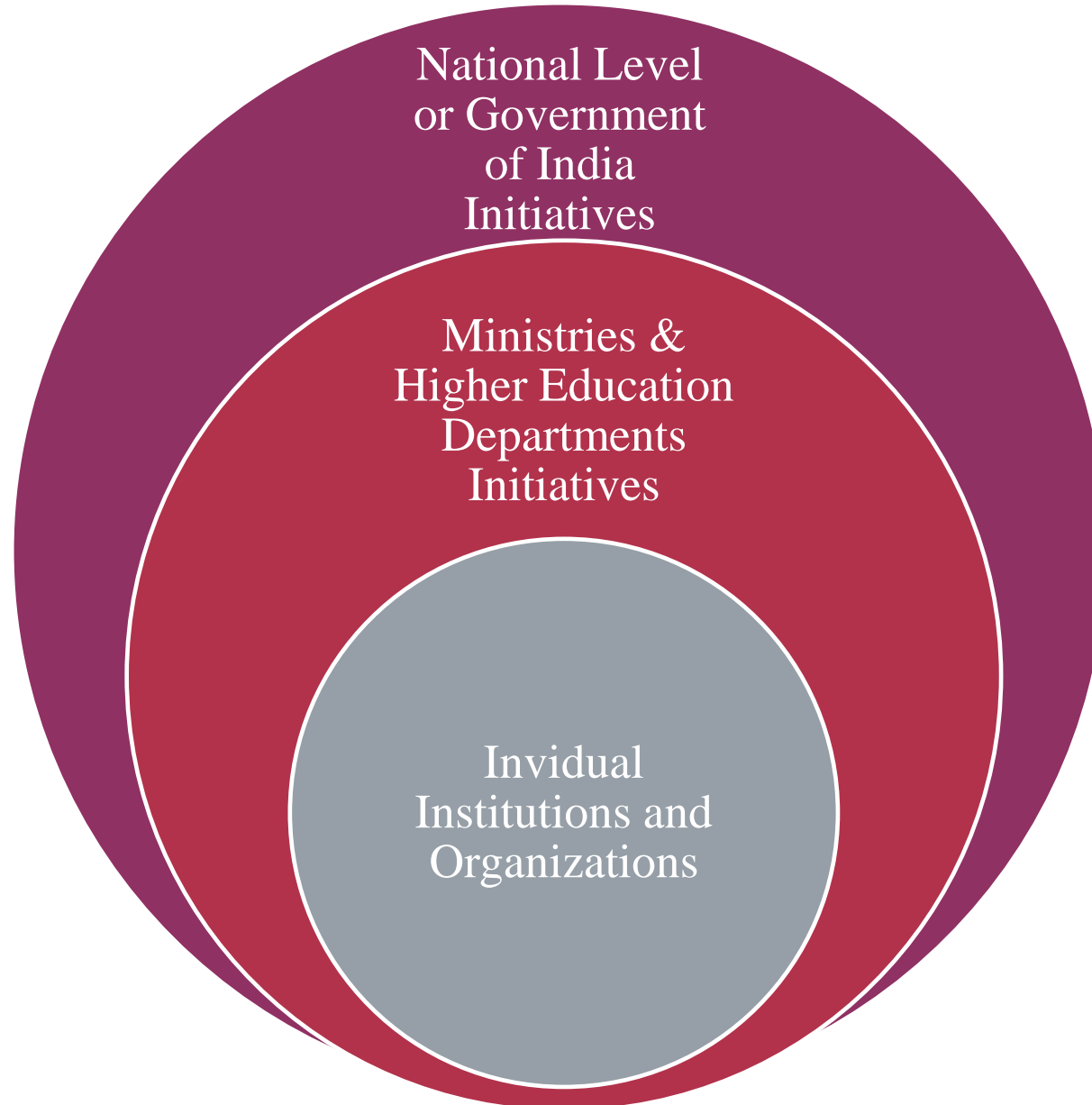
# Findings 4: Interoperable RDRs System in India…

- **Institutional level RDR**

    - Institutions have sole responsibility in building the OAI-PMH-compliant RDR in their respective institutions.

    - Development of policy compliance, guidelines and data quality control

    - Research data curation activities including the research dataset collection/receiving, cleaning, classification, metadata creation, assign license conditions,

    - Build trusted data repositories to publish research data at the institutional level.

    - Will provide the registered and restricted (limited) access within the campus.

    - Only processed data, published data, and analyzed data available for access with other researchers.

    - Confidential data and data which has the personal identifiable information will not be made available.

15

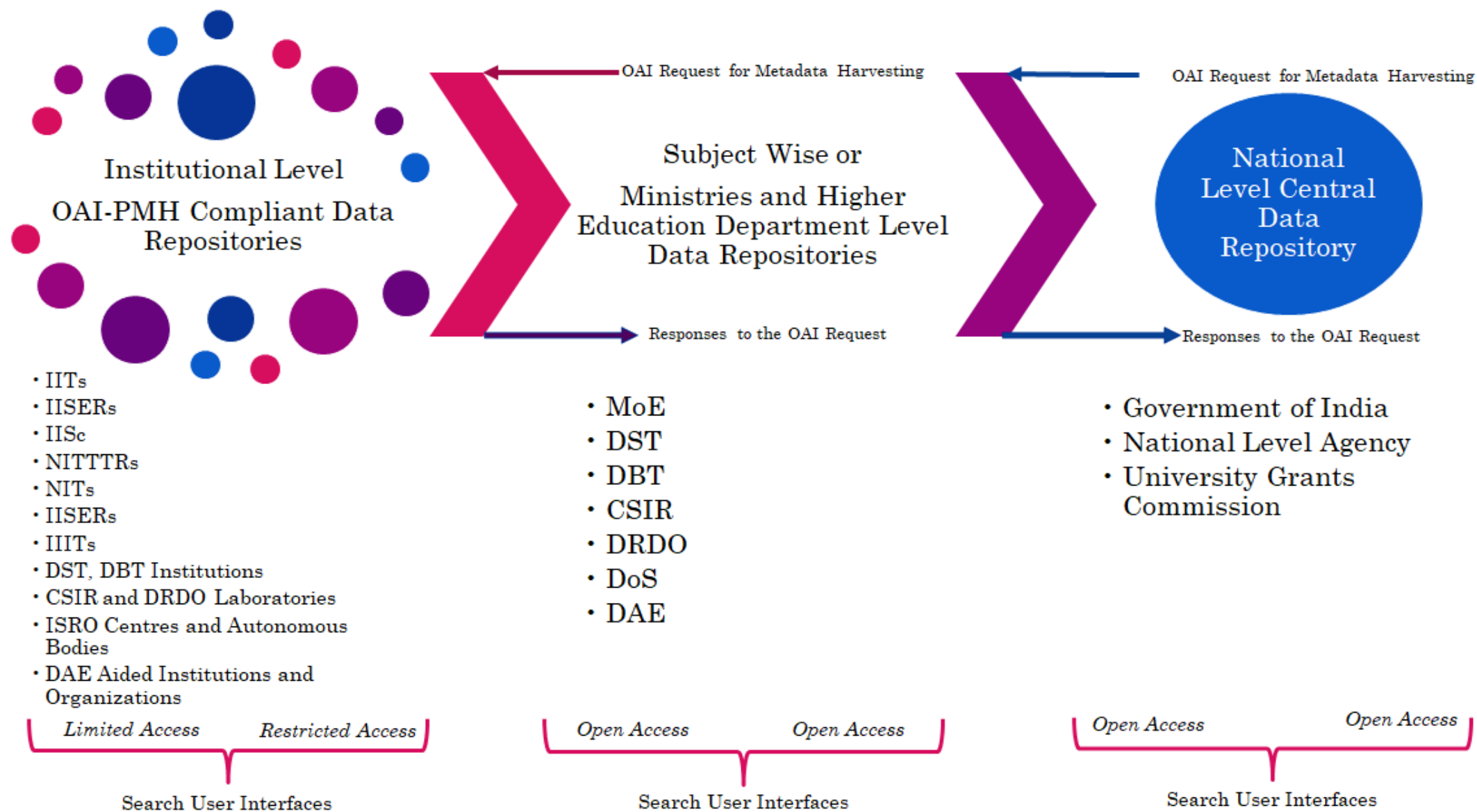# Findings 4: Interoperable RDRs System in India…

- **Disciplined or ministries and higher education department level RDR**

  - The second-level RDRs are hosted by a group of disciplined institutions or institutions funded by the Ministries of Education and Higher Education Departments, Government of India.

  - This level RDRs will host the research data generated by their all-respective institutions and researchers.

  - Research data will be harvested through the individual OAI-PMH-compliant institutional RDRs regularly.

  - These are value-added repositories along with institutional RDRs that enable the research communities to view these repositories to search and retrieve the research data by discipline wise.

  - However, users will have to follow the source URL of institutional RDR to download full datasets.

# Findings 4: Interoperable RDRs System in India…

- **National-level central RDR**
  - Centralized RDR will need to harvest the metadata of research data from the individual institutions and as well as disciplined or ministries and higher education level RDRs.
  - The nodal office, a national agency or network under the UGC, Government of India can host the central server for RDR.
  - The top-level academic unit/networking bodies may take responsibility for developing centralized RDR
  - It enables each institution and ministries / higher education departments will follow the strategic framework in developing the RDR and offering RDM services.
  - The users can access the research data of various institutional RDRs through central search interface without any registration, restriction, and limitation.
  - Users will follow the links to source of institutional RDR to access the full data set(s).

# Findings 4: Interoperable RDRs System in India



Institutional Level OAI-PMH Compliant Data Repositories

- IITs
- IISERs
- IISc
- NITTTRs
- NITs
- IISERs
- IIITs
- DST, DBT Institutions
- CSIR and DRDO Laboratories
- ISRO Centres and Autonomous Bodies
- DAE Aided Institutions and Organizations

Limited Access  Restricted Access

Search User Interfaces

OAI Request for Metadata Harvesting

Subject Wise or Ministries and Higher Education Department Level Data Repositories

Responses to the OAI Request

- MoE
- DST
- DBT
- CSIR
- DRDO
- DoS
- DAE

Open Access  Open Access

Search User Interfaces

OAI Request for Metadata Harvesting

National Level Central Data Repository

Responses to the OAI Request

- Government of India
- National Level Agency
- University Grants Commission

Open Access  Open Access

Search User Interfaces

18

# Summary

- India is lacking in establishing national-level RDM services and repositories like National research data services of other countries.

- *Individual institutions and organizations*; *ministries & higher education departments*; national level, and central level, government of India are major pillars in developing the interoperable RDR system in India.

- Centralized national RDR increases the academic and research prestige of the country and helps in disseminating research data worldwide.

- Research data will be harvested through the individual OAI-PMH-compliant institutional RDRs regularly.

- The proposed workflow in this research study may give way to developing national-level RDM services and repositories to enable more discoverable research data and enhance data usage.

# References (Selected)

- Allaway, R.J., et al., (2019). Engaging a community to enable disease-centric data sharing with the NF Data Portal. *Scientific Data, 6*(319). https://dx.doi.org/10.1038/s41597-019-0317-x

- Alsaad, A., O'Hara, K. & Carr, L. (2019). Institutional repositories as a data trust infrastructure. *WebSci'19 Companion.* Boston: ACM Digital Library.

- Amorim, R. C., Castro, J. A., Silva, J. R. & Ribeiro, C. (2017). A comparison of research data management platforms: architecture, flexible metadata and interoperability. *Universal Access in the Information Society, 16*(4), 851–862. https://dx.doi.org/10.1007/s10209-016-0475-y

- Anilkumar, N. (2018). Research data management in India: a pilot study. *EPJ Web of Conferences, 186*, 03002. https://dx.doi.org/10.1051/epjconf/201818603002

- Arguillas, F., Heslop, J. & Whyte, A. (2015). *Reviewing research data platform capabilities at Cornell Institute for Social and Economic Research* (*CISER*)*.* Edinburgh: A Digital Curation Centre Case Study.

- Baker, K. S. & Duerr, R. E. (2017). Research and the changing nature of data repositories. In L. R. Johnston, *Curating research data, volume 1: practical strategies for your digital repository.* Chicago: Association of College and Research Libraries.

- Barsky, E., Laliberté, L., Leahey, A., Trimble & Leanne. (2015). Collaborative research data curation services: a view from canada. In L. Johnston, *Curating Research Data, Volume 1: Practical Strategies for Your Digital Repository.* Chicago: Association of College and Research Libraries.

- Bhardwaj, R. K. (2019). Content analysis of indian RDRs: prospects and possibilities. *DESIDOC Journal of Library & Information Technology, 39*(6), 280-289. https:/idx.doi.org/10.14429/djlit.39.6.15137

- Bhoi, N. K. (2018). Mendeley Data repository as a platform for research data management. In B. Rautaray, D. K. Swain, & C. Swain, *Marching Beyond Libraries: Managerial Skills and Technological Competencies* (pp. 481–487). Bhubaneswar: Kalinga Institute of Industrial Technology.

# Thank You