International Journal of Information Science and Management
Vol. 24, No. 1, 2026, 133-154
DOI: https://doi.org/10.22034/ijism.2026.2024173.1419 / DOR: https://dor.isc.ac/dor/20.1001.1.20088302.2026.24.1.7.8

Original Research

A Framework for Developing Research Data Governance in Higher Education Institutes: The Case of the University of Qom

Mehdi Akhlaghi
MSc in Knowledge and Information Science, Department of Knowledge and Information Science, Faculty of Literature and Human Sciences, University of Qom, Iran.
m.akhlaghi@qom.ac.ir / mehdiakhlaghi462@gmail.com
ORCID iD: https://orcid.org/0000-0002-5407-0560

Yaghoob Norouzi
Professor, Knowledge and Information Science, Department of Knowledge and Information Science, Faculty of Literature and Human Sciences, University of Qom, Iran.
y.norouzi@qom.ac.ir / ynorouzi@gmail.com
ORCID iD: https://orcid.org/0000-0003-3030-7647

Ali Taei Zadeh
Assistant Prof., Department of Computer Engineering. Faculty of Engineering, University of Qom, Iran.
a.taeizadeh@qom.ac.ir / alitaee@gmail.com
ORCID iD: https://orcid.org/0000-0001-8691-3975

Behrooz Rasuli
Assistant Prof., Knowledge and Information Science, Society and Information Research Department, Iranian Research Institute for Information Science and Technology (IranDoc), Iran.
Corresponding Author: asuli@irandoc.ac.ir
ORCID iD: https://orcid.org/0000-0001-6091-6967

Received: 02 March 2024
Reviewed: 20 May 2024
Accepted: 30 August 2025

Abstract

Research data are vital assets, and higher education institutes (HEIs) are advised to establish spaces for sharing such data. Consequently, the governance of research data has become a significant concern for HEIs. However, the coverage and content of research data governance (RDG) remain unclear. This study aims to identify dimensions and components of an RDG and propose a comprehensive framework in this regard. To identify key dimensions of the RDG, 20 documents on data management and RDG from HEIs worldwide were selected. These documents were coded using MAXQDA software. Subsequently, in-depth interviews were conducted with seven experts in the field. The interviews were transcribed and coded in MAXQDA. The results of document analysis and expert opinions were combined to identify dimensions and components of a comprehensive framework for RDG within HEIs. The research indicates that a comprehensive framework for RDG should address the organizational, legal, and technical aspects of research data management. The framework for RDG comprises five main sections: introduction, approval, principles and policies, roles and responsibilities, and data management. Research data constitute valuable institutional assets that require proper governance. The framework for RDG aims to effectively and efficiently manage these data through policy formulation and delineation of roles and responsibilities within HEIs. To harness the potential benefits of these assets, an RDG policy is essential for HEIs. This is the first study to identify the dimensions and key components of a framework

for RDG. While extensive research has been conducted on data management/governance, a structured policy in this domain is neglected. This study attempts to bridge this gap.

Keywords: Open Research, Open Science, Open Data, Institutional Repository, Information Policy, Laws, Regulations, University of Qom, Data Governance, Higher Education Institutes.

Introduction

Research data is one of the most crucial assets for higher education institutions (HEIs) globally. These data help fulfill institutions' missions in academic research and community development. Consequently, HEIs try to collect, organize, and disseminate this data in proper and effective ways to enable future reuse and impact.

Research data encompasses scientific records collected or generated during research that serve as the foundation for arguments, discussions, or computations (European Commission, 2013, cited in Marlina et al., 2022). Experiment results, measurements, fieldwork observations, surveys, interview recordings, and photographs are examples of research data (Marlina et al., 2022). Researchers are generating increasing volumes of research data during their studies, which must be managed for immediate and long-term (re)use. The value of data becomes apparent when it is utilized. Open science encourages researchers and HEIs to make knowledge, including research data, broadly accessible with minimal restrictions. Sharing research data accelerates scientific progress and innovation. Therefore, HEIs and other data creators need a policy to govern their research data.

Therefore, data governance has emerged as an important area of study over the past decade, gaining increasing importance across various sectors, including higher education (Al-Ruithe et al., 2019; Hassani et al., 2025; Nai et al., 2025). Organizations, including HEIs, are recognizing research data as a strategic asset that underpins decision-making, operational efficiency, and innovation. Consequently, establishing robust research data governance (RDG) frameworks is seen as essential for managing, safeguarding, and leveraging data effectively (Bernardo et al., 2024; Sweeney, 2018).

RDG involves making decisions about data quality, access, ethical and security implications, and overall research credibility (Kouper et al., 2020). RDG addresses questions of who governs or decides about research data. Researchers often make decisions relying on organizational norms, procedures, and disciplinary guidelines. This freedom in decision-making enhances growth but can also lead to diverse and sometimes questionable practices (Kouper et al., 2020).

Having policies for RDG in HEIs holds significant value across various dimensions. Legally, they establish clear responsibilities and obligations for all stakeholders regarding research data and data subjects, including compliance with regulations such as the General Data Protection Regulation and the Data Protection Act 2018, thereby reducing the risk of legal issues or reputational harm. Academically, these policies help preserve research data as identifiable, retrievable, and reusable assets, promoting transparency, scrutiny, and debate within the research community while upholding high research standards. Ethically, such policies safeguard the privacy, rights, and well-being of all participants, researchers, and others involved in research activities, mitigating potential harm and ensuring ethical conduct throughout the research process (BIMMInstitute, 2021).

While some HEIs may adopt existing data governance models or develop tailored

approaches, these models may not be directly applicable to RDG due to factors such as the evolving nature of data, complex research networks, and varying access models (Kouper et al., 2020). As such, there remains a critical need for a comprehensive framework specifically designed to address the unique challenges of RDG within HEIs.

RDG is a multifaceted endeavor that requires careful consideration and strategic planning. As research data continue to expand in volume and complexity, HEIs face the challenge of effectively managing, preserving, and sharing these invaluable assets. The imperative to establish a structured framework for RDG arises from the recognition of the pivotal role research data play in advancing academic knowledge, facilitating collaboration, and fostering innovation.

This research aims to discover the key aspects and elements of a framework for governing research data. Such a framework can help HEIs develop a policy to govern their research data. The research is motivated by the increasing importance and complexity of research data in the era of big data and open science. The research will contribute to the existing knowledge and practice of research data management and governance in HEIs.

Literature Review

The growing recognition of research data as a strategic asset has prompted a significant body of academic and professional literature on data management and governance. While the concept of data governance is well established in the corporate and public sectors, its application within the unique context of HEIs, particularly for research data, remains an evolving field. A systematic review of the academic literature on data governance from 2007 to 2024 highlights its emergence as an important area of study. Early work often drew on enterprise data management, but recent studies have begun to address the specific complexities of the academic environment (Hassani et al., 2025).

Research on data governance in HEIs often falls into several key categories, including the development of governance models, the analysis of institutional policies, and case studies of practical implementation. One notable study, for instance, provides a holistic framework and best practices for data governance in the context of intelligent water systems, a model that, while outside the academic realm, emphasizes the importance of a structured approach to managing complex data sets (Dadiala & Sinha, 2025). Another study proposed a goal-oriented conceptual modeling approach for data governance, enabling organizations to focus on data-related objectives and address challenges while protecting their data assets (Sothilingam et al., 2021). These examples, while not specific to research data in HEIs, illustrate the broader academic effort to develop flexible governance models adaptable to specific organizational needs.

Within the higher education sector, one study (Zhang et al., 2025) examined the institutional evolution of university data governance in China, using the lens of historical institutionalism. This research identifies three distinct phases of evolution: functional departmentalism, cross-departmental collaborative governance, and a more modern focus on data openness and application. The authors' analytical framework highlights the roles of institutional factors, power dynamics, and "path dependence" in shaping governance models. The study concludes that to advance data governance, universities must break from this path dependence and reform institutional frameworks to meet evolving demands.

Similarly, a case study from an academic health-care organization, the University of Iowa,

provides practical insights into establishing data governance for sharing and accessing real-world clinical data for research. The study describes the evolution of their data governance program and the development of an external data-sharing process, noting a significant improvement in data accessibility. This case study emphasizes the importance of data literacy, cross-office collaboration, and structured workflows. However, it also highlights implementation challenges, such as new bottlenecks and the need for ongoing researcher education (Davis et al., 2025). This demonstrates that even successful frameworks can create new complexities that require ongoing management.

In the broader context of research data, there is a recognized need for a research-centric data governance system that can bridge the gap between institutional policies and discipline-specific requirements. Such a system aims to give researchers control over data ownership and align it with their research objectives. This aligns with the argument that RDG involves making decisions about data quality, access, ethical and security implications, and overall research credibility. However, the freedom in researchers' decision-making can also lead to diverse and sometimes questionable practices. A study on data governance in the education sector further underscores the need for a comprehensive literature review and for organizations to design strategies that stimulate positive responses to the adoption of RDG practices (Javadi et al., 2023).

Some studies have focused on the challenges of RDG and management. Patel (2016) identified significant issues in research data management, including copyright, data-use licenses, misinterpretation of data, data security, data privacy, and mental health issues. Marlina and Purwandari (2019) indicated that weaknesses in research data management services include a lack of skills, insufficient research data policies, unsupported infrastructure, and insufficient incentives. Other notable issues in research data management include data privacy, data security, and policy compliance. Ashiq et al. (2020) highlighted challenges, including data storage, copyright-related issues, limited organizational support, a shortage of skilled data personnel, financial constraints, data-sharing concerns, data misinterpretation, and data loss. Other significant challenges relate to technological issues, particularly in software, hardware, rapid changes in information technology, and increased tool costs.

Kouper et al. (2020) asserted that current RDG models do not address research objectives, social and cultural norms, dispute resolution, and dilemmas in the sharing and use of research data. They propose moving towards an RDG pattern that employs conceptual frameworks and views data through the lens of shared knowledge. Marlina et al. (2022) identified factors and indicators of research data management models. The research findings introduce a proposed model comprising four categories: technology, organization, environment, and individuals. The technology category encompasses hardware, software, and security. The organizational category identifies five internal factors: policy, management support, strategy, structure, and organizational culture. The ‘individuals’ category includes four factors: knowledge, skills, situational awareness, and training. The environmental category involves governance regulations and financial supplier policies.

Cupoli et al. (2014) introduced the DAMA International's DMBOK2 framework. This framework defines 11 knowledge areas for data management. Expertise is categorized into knowledge areas, each of which may include multiple topics. The 11 knowledge areas and some of their topics include data governance (central to the framework), interaction and alignment between knowledge areas through data governance processes, and other areas such as data

architecture, data modeling and design, data storage and operations, data security, data integration and interaction, documents and content, references and primary data, data warehousing and business intelligence, metadata, and data quality.

Bhansali (2013) defined a three-tier model commonly adopted by most organizations. At the top of the model is the data governance council or governance committee responsible for data management decision-making. This council is composed of various specialists who handle policy-making and includes both business and technical stakeholders. The next level is the data governance office, which coordinates data governance (strategic) and stewardship (tactical) activities. At the lowest level are data stewardship teams, which may be located across various parts of the organization. Continuous interaction exists among these three levels of the data governance structure.

A critical analysis of existing data governance frameworks, particularly in the context of research data within HEIs, reveals several limitations. Many frameworks lack a holistic or universally accepted approach, making standardization challenging (Odebrecht, 2025). Historically, data governance initiatives were often IT-driven, leading to rigid and fragmented processes that lacked broader organizational buy-in (Abraham et al., 2019). Within HEIs, researchers usually encounter fragmented responsibility and ambiguous accountability for research data management. Existing institutional support frameworks are often seen as insufficient or disconnected from the practical needs of individual researchers, hindering effective data management and data ownership. Furthermore, data quality and integrity issues are persistent across health data and other sectors, directly impacting decision-making accuracy and operational efficiency, and many institutions lack overall data standards and robust verification processes. Specifically, there is a lack of literature explicitly outlining the implementation and practical experiences of RDG within Research Performing Organizations (RPOs), limiting the availability of best practices tailored for academic environments.

Moreover, traditional data governance approaches struggle to cope with the explosion in volume, variety, and velocity of modern big data, alongside the emergence of new technologies such as Artificial Intelligence (AI), cloud computing, and the Internet of Things (IoT). These advancements introduce further complexities related to security, privacy, ethical use, and data quality that current frameworks may not adequately address. The explicit integration of moral concerns, human rights, and broader societal values into data governance frameworks is often overlooked, despite the potential for data misuse, bias, and the exacerbation of inequalities, especially when dealing with sensitive data such as genomic information. This highlights a crucial gap in current research, indicating a need for more contextually informed brain data governance frameworks and empirical studies on how stakeholders view these issues. Such comprehensive, adaptable, and research-centric data governance frameworks are necessary to move beyond generic models, address unique ethical, legal, organizational, and technical challenges in HEIs, and ultimately enhance the value derived from research data.

Materials and Methods

This study aims to explore the essential elements of a framework for governing research data in HEIs. This study falls under the type of research with an exploratory purpose. Exploratory research is the initial investigation that tries to propose a new hypothesis or theoretical concept in a specific scientific domain. Exploratory research is valuable and practical when the topic has not been previously studied for a particular sample or group, and

existing theories cannot adequately explain it (Creswell, 2014).

The background of research data management and RDG in the field of management was investigated. Subsequently, through the utilization of the "Google" search engine with the term "Research Data Governance," the top twenty documents (Table 1) related to research data management and RDG from HEIs around the world were selected. Notably, because RDG is emerging, most retrieved documents were related to research data management. Given the shared elements between these two domains, these documents were used to identify the dimensions of the RDG framework. It is worth mentioning that these documents were retrieved in June 2022.

Table 1
A list of the top 20 documents related to research data management and RDG from HEIs

HEIs	Document Title	URL
BIMM Institute	Research Data Governance Policy	https://www.bimm.ac.uk/document/research-data-governance-policy/?v=1686737694
Florida State University	Research Data Management POLICY	https://regulations.fsu.edu/sites/g/files/upcbnu486/files/policies/research/FSU%20Policy%207A-26.pdf
Missouri University of Science and Technology	Policy - Scholars' Mine Research Data Management	https://scholarsmine.mst.edu/cgi/viewcontent.cgi?article=1006&context=scpro_guidelines
Monash University	Research Data Management Policy	https://www.monash.edu/library/researchers/data-collection-management/about
Rice University	Research Data Management	https://policy.rice.edu/308
University of Leeds	Research data management policy	https://library.leeds.ac.uk/info/14062/research_data_management/68/research_data_management_policy
The University of Edinburgh	Research Data Management Policy	https://www.ed.ac.uk/information-services/about/policies-and-regulations/research-data-policy
The University of York	Research Data Management Policy	https://www.york.ac.uk/about/departments/support-and-admin/information-services/information-policy/index/research-data-management-policy/
Robert Gordon University	Research Governance and Integrity Policy	https://www.rgu.ac.uk/research/research-governance/research-integrity-and-ethics
Tilburg University	Research Data Management Regulation	https://www.tilburguniversity.edu/sites/default/files/download/Research%20Data%20Management%20Regulation%20%28version%20April2020%29.pdf
The Universitat Autònoma de Barcelona	Institutional Policy on Open Access to Research Data	https://ddd.uab.cat/pub/docins/2020/222172/Institutional_Policy_on_Open_Access_to_Research_Data_at_UAB_a2020m7.pdf
University of Cambridge	Research Data Management Policy Framework	https://www.data.cam.ac.uk/university-policy

HEIs	Document Title	URL
University OF CAPE TOWN	Research Data Management Policy	https://www.uct.ac.za/research-support-hub/research-data/managing-research-data
The University of Liverpool	Research Data Management Policy	https://www.liverpool.ac.uk/library/research-data-management/
University of Melbourne	Research Data Management Policy	https://policy.unimelb.edu.au/MPF1242/
University of Salford	Research Data Management Policy	https://www.salford.ac.uk/research/research-integrity/research-data-management
University of Southampton	Research Data Management Policy	https://www.southampton.ac.uk/about/governance/regulations-policies/policies/research-data-management#:~:text=Good%20research%20practice%20requires%20research,funder%2C%20sponsors%20and%20the%20University
Utrecht University	University policy framework for research data	https://www.uu.nl/sites/default/files/university_policy_framework_for_research_data_utrecht_university_-_january_2016.pdf
Al-Quds University	Research Data Management Policy	https://dspace.qou.edu/contents/OSOLPolicies.pdf
Vanderbilt University	Research Data Governance Framework	https://www.vanderbilt.edu/datagovernance/data-governance-committee/

The selected documents were then coded in MAXQDA, enabling the identification of the components of the RDG framework. In the subsequent phase, this research conducted interviews with seven experts, including one policymaker, one researcher in data governance, and five Library and Information Science experts. The experts were purposively selected based on their specific expertise and roles relevant to the research topic. The goal was to ensure a diverse and comprehensive range of perspectives. The selection criteria included:

- Policymaker: An individual with a direct role in creating and implementing institutional policies related to research or data, providing a high-level, practical perspective on governance challenges.
- Data Governance Researcher: An academic with a background in the field of governance, offering a theoretical and research-based understanding of governance frameworks.
- Library and Information Science Experts: Five individuals with deep expertise in managing information and data, representing the practical and technical side of research data management and preservation.

An interview can be structured, semi-structured, or unstructured and can be conducted face-to-face, over the telephone, online, or in writing (Gibson & Brown, 2009). The interviews were semi-structured. The interview questions were structured based on an analysis of 20 documents on research data management and RDG from HEIs around the world. The conducted interviews were transcribed and encoded through MAXQDA. Data analysis occurred in two phases: coding and pattern identification. In the coding phase, codes were assigned to each section of

the data. In the subsequent phase, these codes were organized into larger categories, constituting the pattern recognition process. In the next step, the results from document analysis and expert perspectives were merged to identify the dimensions and components of a comprehensive framework for RDG within HEIs.

According to Creswell (2014), researchers' familiarity with the research topic and their investment of ample time in conducting research are strategies for enhancing the credibility of qualitative research. Having protocols for conducting fieldwork, document study, and observation for data collection and analysis improves the reliability of the research. Another method for increasing reliability involved leveraging "MAXQDA" for data analysis. Implementing textual data, using an interview protocol, taking notes during document study, and conducting observations also contributed to increased reliability in this study.

Results

The results of document analysis and expert opinions are combined to shape the dimensions and components of a comprehensive RDG framework for higher education institutions. This framework is presented in Figure 1.

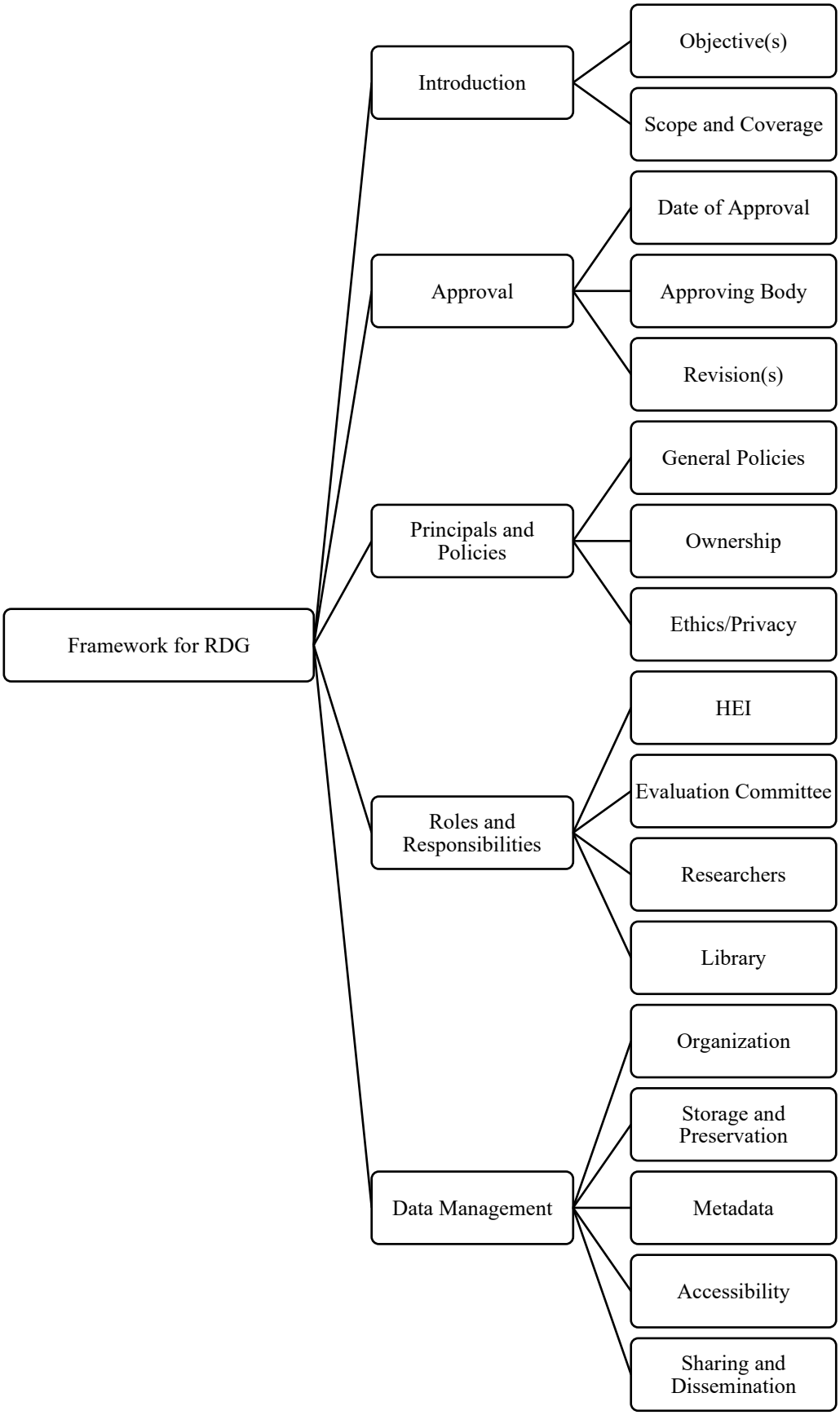


Figure 1: Dimensions and components of the RDG framework

As depicted in Figure 1, the RDG framework comprises five key dimensions: Introduction, Approval, Principles and Policies, Roles and Responsibilities, and Data Management. Each of these dimensions consists of various components.

Introduction

The Introduction section serves as the foundational element of the RDG framework (Figure 2). Its primary purpose is to articulate the framework's objectives, scope, and coverage. This section makes it clear that the framework's overarching goal is to ensure that all research data are managed effectively throughout their entire lifecycle. It clarifies the responsibilities of both the institution and individual researchers, aiming to promote data integrity, enable future reuse, and ensure compliance with all relevant legal, ethical, and funder requirements. By explicitly defining who the framework applies to-including all faculty, students, and staff-and its lack of geographical limits, it establishes a clear and comprehensive mandate for data governance across the university. The introduction essentially sets the stage by defining the "what" and "why" of the framework before delving into the "how."

Figure 2: RDG framework: Introduction

Objective

The RDG framework delineates the responsibilities of the university and researchers for managing research data. It ensures the preservation, maintenance, and accessibility of research data for sharing and reuse, in compliance with legal, ethical, and funding-agency requirements. One of the documents articulates the objective of the framework as follows: "1- Ensure that researchers manage their data effectively, enabling them to: Demonstrate the integrity of their research. Preserve eligible data for reuse within and beyond the university (as appropriate). Comply with ethical, legal, funder, and other requirements in relation to data and data management. 2- Ensure that University Departments understand their responsibilities in relation to governance of research data management in their departments 3- Ensure that the University responsibilities for research data management are clearly articulated" (The University of York, p. 1).

Scope and coverage

Document analysis indicates that these frameworks encompass the entire lifecycle of research data and include all research activities. For instance, "This framework applies to all persons who generate or utilize Research Data at the University. This group includes students, faculty, staff, researchers, contractors, visitors, and others who access the University's Research Data or computing, storage, and network facilities. It applies in all locations where the University conducts its activities without geographical limits, subject to applicable local laws and regulations." (Vanderbilt University, p. 1).

Approval and relevance

The Approval and Relevance section establishes the framework's institutional authority and validity (Figure 3). It specifies the official body responsible for issuing and sanctioning the framework, namely the Research Deputy of the institution. This delegation of authority ensures that the framework has the necessary institutional backing to be enforced across all departments and research activities. Additionally, this section outlines the process for periodic review, with a typical cycle ranging from 1 to 5 years. This periodic review is crucial for ensuring the framework remains current, relevant, and adaptable to new technologies, evolving legal requirements, and changes in research practices. By formalizing the approval process and setting a review schedule, this section provides a framework that is both officially legitimate and long-term sustainable.

Figure 3: RDG framework: Approval and relevance

Principles/policies

The Principles and Policies section lays out the core philosophical and operational tenets that guide all aspects of research data governance (Figure 4). It defines fundamental principles that view research data as a valuable institutional asset, emphasizing the need to manage it,

ensure it is fit for purpose, and make it readily accessible, while also remaining secure and compliant with regulations. This section also addresses critical policies on ownership and ethical/privacy issues. It clarifies that, in most cases, the university retains ownership of the research data, even if the researcher leaves the institution. This policy is vital for ensuring long-term institutional preservation and control. Furthermore, it explicitly mandates that all data management adhere to the highest ethical standards, protecting the rights and privacy of research subjects and ensuring compliance with regulations such as the GDPR and the Data Protection Act.

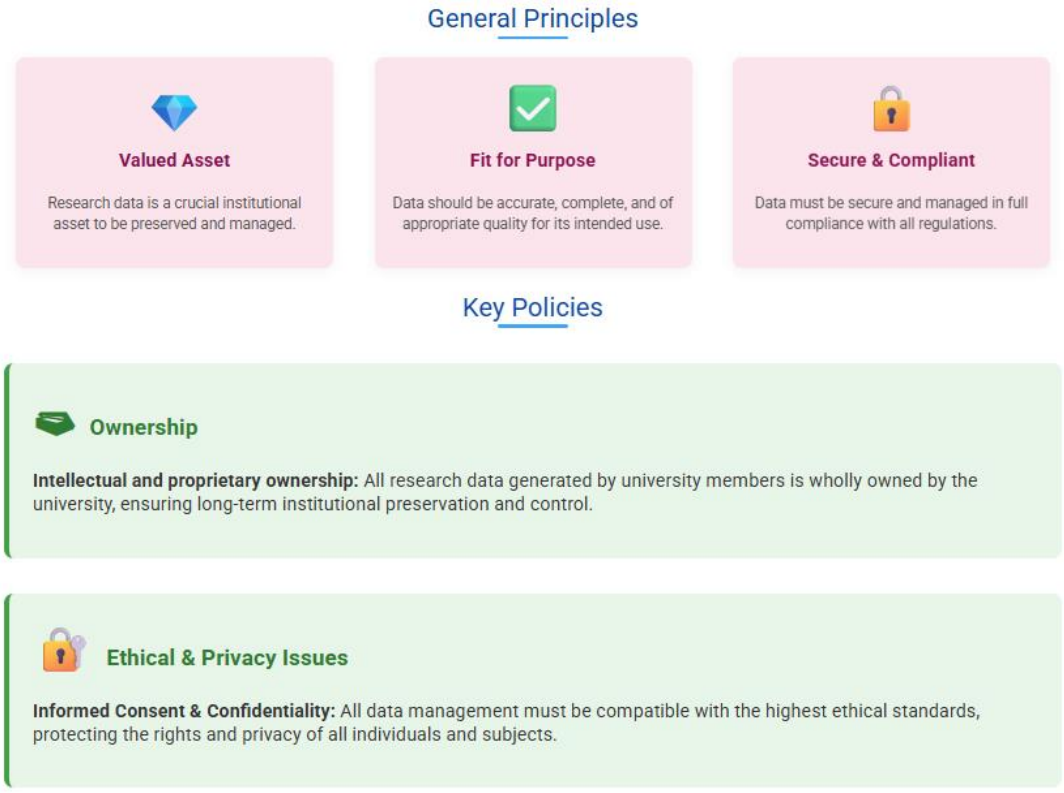


Figure 4: RDG Framework: Principles and policies

General principles

In this section, the fundamental principles of the framework are defined from the institution's perspective and objectives. One document states its principles as follows: "1- Research Data is a valued asset […] 2- Research Data is managed: Research Data should be appropriately managed (i.e., collected, stored, protected, and used) throughout its life cycle […] 3- Research Data is fit for purpose: Research Data should be as accurate and complete as possible (given the constraints of the data and how it was collected), at the appropriate quality for its primary purpose and all other known legitimate uses […] 4- Research Data is accessible, comparable, and reusable […] 5- Research Data is secure and compliant with regulations [...] " (Vanderbilt University, pp. 2-3).

Ownership (physical and intellectual)

An essential dimension in the RDG framework is the discussion of intellectual and proprietary ownership, which involves legal and contractual issues related to research data. Data

ownership should be determined in accordance with institutional regulations and agreements with third parties before commencing a research project. Whether intellectual ownership belongs to the researcher or the university is a consideration. Nevertheless, most of the reviewed documents treat research data as an asset owned by the university.

"All research data generated by UoL academics or research postgraduate students is wholly owned by the University (or funder) and remains with the University if the academic leaves the institution. Where taught students generate data under the supervision of an academic member of staff as part of a wider funded project, the data will be wholly owned by the University (or funder) as above. Unless it is a condition of a grant or contract, exclusive rights to research data must not be assigned, licensed, or otherwise transferred to external parties." (University of Liverpool, p. 1).

Ethical/privacy issues

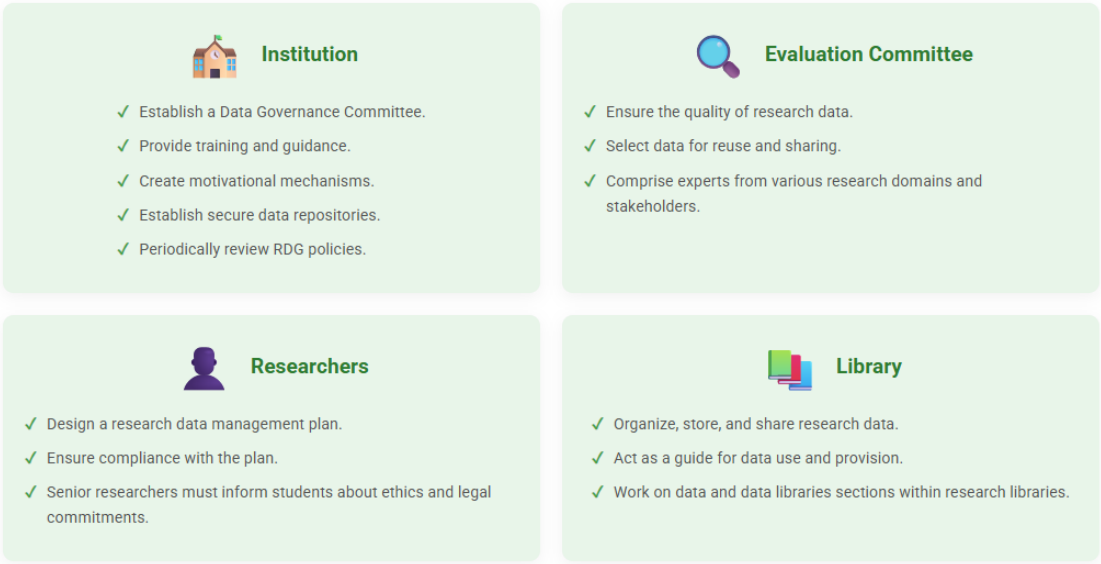
In any research data management program, legal requirements, ethical issues, and privacy considerations must be taken into account. " […] The management of research data should be compatible with the University's commitment to the highest ethical standards in research, protecting the rights, dignity, health, safety, and privacy of the community, including research subjects, and with its commitment to the welfare of animals and the integrity of the environment. […] " (Monash University, p. 1). "In determining access rights for research data, the legitimate interests of the subjects of research data must be protected. The rights of citizen scientists and the public to access publicly funded research should also be consideredas well. When open access to datasets is not legal or ethical, information governance restrictions on access and use must be applied as necessary. […]" (The University of Edinburgh, p2). "[...] The study must ensure that the participants have given their 'informed consent'. This must inform participants about how research data will be stored, preserved, and used in the long term, and how confidentiality will be maintained. The 'informed consent' should enable the data collected to be made available to other researchers […]" (University of Liverpool, p2).

Roles and responsibilities

The Roles and Responsibilities section clearly delineates the duties of all stakeholders involved in the research data lifecycle (Figure 5). It defines who is accountable for what, preventing ambiguity and ensuring effective implementation of the framework. Key roles are outlined for the institution, which is responsible for establishing a data governance committee, providing training, and maintaining secure data repositories. The framework also identifies the need for an Evaluation Committee to assess data quality and select data for sharing, composed of experts from various research domains. Crucially, the section details the responsibilities of researchers, who must create and follow data management plans, and supervisors, who must ensure their students are informed. It also recognizes the vital role of the library in organizing, storing, and providing access to research data.

In the context of defining roles and responsibilities, one of the interviewees states: "[...] Data governance determines these, for example, who is responsible for qualitative analysis or the quality of data; these are specified in data governance. Who has decision-making authority

in this area? Who is responsible for data quality assurance? Its processes are defined there [...]"



(Interviewee 7, pos 24).

Figure 5: RDG Framework: Roles and responsibilities

Institution

By examining documents and consulting experts' opinions, several institutional tasks can be identified: establishing a data governance committee, providing training and guidance, offering recommendations to researchers, creating motivational mechanisms for research data management, establishing secure data repositories, updating hardware and software resources, maintaining and ensuring public access to data as needed, and periodic review of RDG policies and frameworks. One document states the institution's tasks as follows: "1- Developing infrastructure and training to promote best practice in data management amongst its researchers. 2- Supporting researchers to plan for data management and write data management plans for grant applications. 3- Managing a dedicated website guiding the University's researchers in good data management practice, including data deposition and related metadata description, and good data governance in compliance with relevant legal and ethical obligations. 4- Maintaining an institutional metadata catalogue of research datasets for publicly funded research in line with funder requirements. 5- Providing Digital Object Indicators (DOIs) for datasets deposited at the University as per the DOI policy." (University of Southampton, p. 4).

Evaluation committee

One of the objectives of RDG is to ensure quality data and the selection of data for reuse and sharing. From the perspective of experts, research data assessors can be assigned an evaluative role. Since research is conducted across various fields, a committee comprising experts from each field and relevant stakeholders is required to implement the RDG framework. "[...] If there is any selection, it should be determined by a committee that includes various stakeholders. The research domain, the research deputy, the research manager, and the library head should be present. The domain that collects its data: one person, two people, or three expert people who understand the value of that data and can talk about it. This is the time that helps us select the data. If we are going to select the data, how should we do so? These differ

depending on the domain." (Interviewee 5, Pos 23).

"[...] Well, research data is definitely in files. Part of data quality assessment or data assurance can be formal discussions. For example, our file is correct, not damaged. These are formal discussions about the data. Another part is the data's content. Naturally, the data producers themselves and their peers can assess its quality [...]" (Interviewee 7, Pos 22).

Researchers

At the beginning of each research project, a research data management plan must be designed, outlining how data will be collected, stored, and reused. Researchers are obligated to ensure compliance with the research data management plan within all relevant legal, ethical, and supervisory frameworks. Senior researchers and supervisors must ensure that students are adequately informed about research ethics, effective data management, and legal commitments.

Library

The library can play a role in organizing, storing, and sharing research data. One of the interviewees expresses: "[...] For instance, there should be a section that guides individuals on how we used to lend books, how to use the reference section, how to use data, and how to provide data. Now, in the digital environment, it has its own rules and regulations, and we need to consider them. But as for where it can fit, well, globally, as I mentioned, we now have sections for data and data libraries. This means that in research libraries, they assign sections to work on these data [...]" (Interviewee 1, Pos 15).

Data management dimensions

The Data Management Dimensions section focuses on the technical and logistical aspects of the framework (Figure 6). This section provides a detailed guide to handling research data in practice. It outlines the role of an Organizational Custodian—typically a repository service—that provides the infrastructure for storing and accessing data. It also specifies requirements for storage and preservation, including adherence to best practices, secure infrastructure, and defined retention periods (often 10 years). A significant component of this section is the emphasis on metadata, requiring that sufficient, high-quality metadata be recorded to make data discoverable and reusable. Finally, it addresses accessibility and sharing policies, mandating that data management plans specify when and how data will be made available. It also provides clear guidelines for when data sharing may be restricted, such as in cases of intellectual property concerns or compromised privacy.

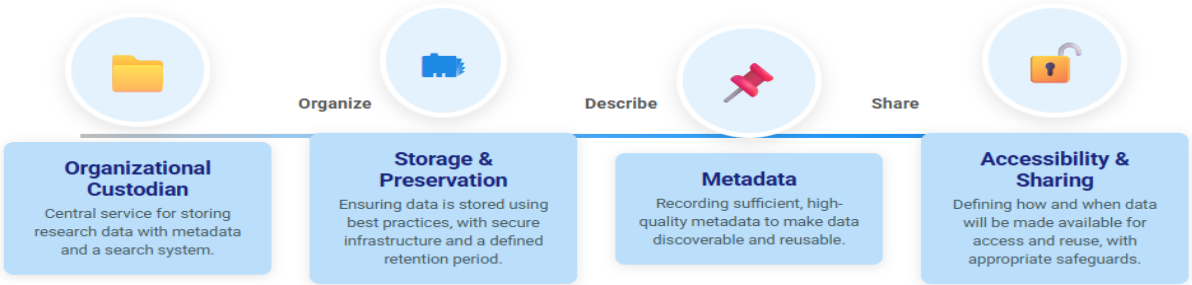


Figure 6: RDG Framework: Data management dimensions

Organizational custodian

Organizational custodian is a service that allows researchers to store research data along with appropriate metadata and provides a search system for accessing the data. The repositories or third-party data archives: 1- provide facilities to enable research data stored in the repository to comply with the Research Data Management Regulation. 2- Use open standards for access and metadata wherever possible. 3- are preferably certified. 4- are technically and organizationally sustainable" (Tilburg University, p. 3).

Storage and preservation

"Research data must be stored in a manner that is consistent with accepted best practices in the digital preservation community. This will include both technical infrastructure (hardware, software, network access, data backup, facilities, maintenance, etc.) and ongoing preservation management activities. […]" (Missouri University of Science and Technology, p. 5). "[…] Routine published data should be retained for an appropriate period consistent with standard practices or requirements within a discipline, funding agency, or department, or for a fixed period of time-[…]" (Florida State University, p2). A significant number of documents have specified a ten-year retention period in the absence of legal requirements.

Metadata

Another topic addressed in RDG is metadata. To describe and organize this data, metadata is required to enable the implementation of referencing research data.

"[…] Sufficient metadata should be recorded and made openly available to render the research data discoverable for effective reuse. The metadata standard applied should be appropriate, rich, of high quality, and comply with the discipline's norms and good practice, using a recognised standard where such exists. A data citation is required to provide due recognition to the data provider and should include a persistent, machine-actionable, globally unique, and widely used identifier for the data. […]" (University of Cape Town, p. 10). "[...] To ensure its metadata, standards must be defined. For example, let's assume that these data are in what format? For instance, how extensive is the population of this research? What has been the time span? What items have been investigated, and where have the findings been reflected? [...]" (Interviewee 3, Pos 8).

Accessibility

"Research data management plans must specify how and when research data will be made available for access and reuse. Generally, accompanying research data should be made available as soon as findings are published, under appropriate safeguards when necessary. A Data Protection Impact Assessment is required whenever data about individuals is used. Research outputs that convey findings from research data should include a citation or data access statement that clearly provides readers with a route to obtain the data. When the researcher does not hold the intellectual property rights or have permission to make derived data available, they must provide a citation or data access statement in relevant publications to the source of data." (The University of Edinburgh, p2).

Sharing

"All research data produced by UoL researchers should be made openly available with as few restrictions as possible, promptly. Personal research data may be made available if it is

anonymised and there is informed consent to share it. Sharing of research data will be restricted when: 1- the rights of individual researchers/and or subjects would be compromised. 2- The research has an Intellectual Property opportunity arising from it, in alignment with the Intellectual Property Policy. 3- It includes secondary data that you may have had the right to use as part of your research, but do not have the right to share with others. " (University of Liverpool, p. 3).

Discussion

Research data is information collected during the research process. Research data can be quantitative or qualitative. Due to variations across subject areas, these data exhibit distinct structures. The heterogeneity of research data and the differing standards across fields are noteworthy concerns in research data management. To implement a framework for RDG, an initial point must be established: all these data must be collected digitally and delivered to the institutional repository.

An issue arising in the management of such data is that they are collected for specific research purposes in different studies and may not be applicable or of sufficient quality for reuse in other research or scientific endeavors. Given these circumstances, should data be collected from all research endeavors? In response, preserving all this data for a specified period can help validate the credibility and authenticity of research, potentially deterring scientific misconduct and enhancing research quality. However, controlling all research data and detecting fraud when a researcher has personally generated the data can be challenging. Nonetheless, through organizational management and governance of these data, steps can be taken to reduce such occurrences.

One of the goals of research data management is the sharing of these data. Sharing data increases trust, reuse, citations, transparency, and efficiency (Patel, 2016). However, it should be noted that data can be reused only if they are of high quality. Data quality involves assessing a dataset's suitability for addressing specific organizational needs. The definition of data quality varies to some extent based on requirements. Researchers might not necessarily need abundant data; they require appropriate data. Small data values can be as valuable as large ones. In many instances, no relevant data might be available due to the inability to locate or access it, or because it is nonexistent (Semeler, Pinto, & Rozados, 2017, as cited in Koltay, 2020).

Data quality can be measured in two dimensions: formal and content-related. In the formal dimension, the technical aspects of data collection and preservation are discussed, while in the content-related dimension, the data must adhere to quality criteria specific to each field. Solomonides (2019) outlines data quality criteria as data accuracy, data credibility, data reliability, data timeliness, and data completeness, and research data should conform to these criteria. Subject specialists from different fields and information technology experts should assess data according to these criteria and field-specific considerations.

Another aspect of data collection and reuse that poses a challenge is data misinterpretation, as Patel (2016) noted. Data misinterpretation occurs when the methodology, conditions, and limitations that influenced data collection are unavailable, potentially leading to unreliable results despite valid data. A solution to this issue is metadata that describes these data. This metadata should include fields that convey the research community, time range, and data collection methods.

Governance implies policy-making, process determination, and the definition of roles and

responsibilities within an organization to ensure individuals are accountable for their decisions. Governance enhances data quality. The RDG framework encompasses the entire data lifecycle in research activities, involving data management planning, data collection, data analysis, data dissemination and sharing, data preservation, and data reuse (Briney, Kristin, 2015, cited in Marlina & Purwandari, 2019). Responsibility for governance lies at the upper echelons of the institution. Given the various stakeholders in research data discussions, a committee consisting of senior managers and data stakeholders is needed to ensure that adopted policies align with organizational visions and to monitor the implementation of RDG. The COBIT governance framework defines roles using the RACI chart, including responsible, accountable, consult, and inform (IT Governance Institute, 2005, cited in Wende, 2007). This committee should consider legal, ethical, and financial supplier requirements, as well as other research data-related obligations. This committee could consist of research vice-presidents, research directors, IT vice-presidents, library directors, and individuals familiar with research outputs and data management, and open science issues.

The matter of ownership, whether it be physical or intellectual, in the context of research data constitutes one of the challenging discussions within this framework. Experts in this domain have notably focused on this issue. The findings of Patel (2016), Marlina and Purwandari (2019), and Majidi et al. (2017) have elucidated data ownership, intellectual property rights, and privacy as significant aspects in the discourse on research data management. Data ownership, in a broader legal context, encompasses various facets. When stakeholders and financial contributors serve as third parties in research endeavors, contractual arrangements among them determine data ownership. However, the discourse on data ownership takes on a different dimension in research projects funded by the public budget. In such cases, provision must be made to allow public data sharing if no sharing restrictions exist. Marlina and Purwandari (2019) highlight that due to the absence of national and organizational policies, data collected in research projects funded by public budgets are typically retained by researchers with limited access. Nevertheless, the rights of researchers, as those who have primarily invested effort in the research process, should be upheld.

Within the framework proposed in this study, a distinction is made between physical and intellectual dimensions of ownership to balance institutional and individual rights. While the physical ownership of research data-the raw files and storage-often defaults to the university (especially when using public infrastructure), this framework argues that intellectual ownership should ideally rest with the researcher. This distinction recognizes the researcher’s creative labor. It ensures their right to be credited through data referencing), even when the institution maintains legal "ownership" for compliance or funding purposes (Patel, 2016), even when the institution maintains legal "ownership" for compliance or funding purposes.

Another legal and ethical matter within the framework is the discussion of ethical considerations and privacy concerns in research data. It is worth noting that just as ethical considerations are pivotal in research, they are also crucial in the context of research data. The RDG framework, through the establishment of requisite policies and strategies, must ensure the protection of privacy for participating individuals. Moreover, it should assure that the security quotient of the data maintains that the research data will not be misused, thus preserving the rights of both researchers and participating individuals. Additionally, specific research projects are subject to confidentiality conditions, which can be based on national issues or safeguarding the rights of financial contributors to research. Therefore, data disclosure in these projects could

compromise the interests of data stakeholders. To ensure information security and protect data privacy and confidentiality, the involvement of information technology experts skilled in encryption or data anonymization is imperative. Thus, information technology specialists play a crucial role as participants in the policy-making and execution discussions within this framework.

One of the institution's primary responsibilities is to establish an RDG committee and provide the necessary technical, legal, and organizational infrastructure for RDG. Furthermore, the institution is tasked with the training of researchers. Some experts have identified a lack of familiarity among researchers with discussions of RDG and management as one of the challenges in implementing this framework. The institution could develop mechanisms to encourage researchers in this regard. Marlina and Purwandari (2019) view the absence of incentives as one of the institution's shortcomings in providing research data management services.

The evaluation committee, composed of experts from various fields familiar with the research data management process and technical specialists in data, should undertake the examination of the formal and content aspects of research data. Given the heterogeneous nature of research data, it is necessary to engage domain specialists to assess and monitor its quality. These individuals should be well-versed in the requirements of research data management. This committee can serve as a reference to guide individuals on research data management, ensure effective framework implementation, identify strengths and weaknesses, and transfer these insights to the governance committee for framework enhancement.

Collected data, if not organized, will be of little use. Libraries, in the context of organization and dissemination, can play a fundamental role. However, librarians need to train in data science to provide essential data librarian services. Regarding the role of researchers, they should establish a specific plan aligned with institutional guidelines for managing their data throughout the research lifecycle. The interaction of researchers with the execution of this framework is a multifaceted discussion encompassing various issues. A data governance plan drives change within an organization. By soliciting participation and feedback, valuable insights can be gained for the effective implementation of data governance. Moreover, understanding the factors that influence individuals' perception of data governance aids in its practical implementation (Stephens, 2018).

The Dimension of Data Management Encompasses Organizational Repository, Storage and Preservation, Metadata, Access, and Sharing. The organizational repository is the institution's repository, where researchers' data are collected. This repository fulfills an essential need for organizational management of research data. Information technology assumes responsibility for the infrastructure that stores, processes, and reports data. These infrastructures must incorporate measures to prevent unauthorized use of data (Cheong & Chang, 2007).

Storage and preservation of research data are conducted in accordance with the institution's established policies. These institutional policies determine whether all data should be stored or only data from specific research should be retained. Additionally, the duration for which this data is retained is determined based on these policies and the requirements of research funders. Once the organization's storage and preservation policies are defined, the evaluation committee can adhere to them to select which data to store. The committee can determine the retention period for different data sets and identify data sets with little retention value.

To adequately describe research data and facilitate easy access, appropriate metadata is

indispensable. These metadata fields should describe the data structure and collection methodologies. Access to data is contingent upon the institution's policies and the rights of research data stakeholders. The institution's policies are established and overseen by the governance committee to make data accessible. Moreover, data access considerations must incorporate data confidentiality and privacy.

Sharing high-quality data not only promotes scientific advancement but also saves time and resources. In the context of data sharing, commercial organizations often face greater constraints due to competitive pressures than higher education institutions. The purpose of research in universities is to utilize their findings for societal growth and development. This objective is achieved by sharing research results, unless there are limitations on sharing. Research data are no exception to this principle. Sharing high-quality data is a significant outcome of implementing the RDG framework.

Conclusion

RDG within HEIs facilitates the storage, preservation, access, and sharing of substantial amounts of research data. These data, if not adequately managed over time, are susceptible to loss, and the endeavor to reproduce them can be time-consuming, costly, and in some cases, infeasible. Effective research data management ensures data integrity and coherence, enabling these data to contribute to the validation of research findings and to be reused in subsequent studies. To achieve efficient and effective management of these data assets and prevent the waste of this valuable resource, establishing an RDG framework is imperative. This framework should address various dimensions and issues involved in the management of research data.

However, implementing such a framework is not without its challenges. A primary obstacle is institutional inertia and "path dependence," where established norms and departmental silos can resist new, cross-functional policies. Additionally, researcher resistance may arise due to concerns about a perceived loss of autonomy over their data or a lack of understanding of the framework's benefits. The financial and logistical investment in technical infrastructure and ongoing training for researchers and staff also poses a significant challenge for many institutions.

These challenges, however, can be addressed by the framework itself. By clearly delineating roles and responsibilities for all stakeholders, the framework mitigates confusion and establishes accountability. The institutional commitment to providing training and guidance, along with the creation of motivational mechanisms, can help overcome researcher resistance. Furthermore, the framework's focus on defining clear policies for data ownership, ethics, and security provides the necessary structure to reform existing practices. By addressing these practical and cultural challenges head-on, the proposed framework can be effectively adopted, ensuring the long-term integrity and value of an institution's research data assets.

References

Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, 49, 424–438. https://doi.org/10.1016/j.ijinfomgt.2019.07.008

Al-Ruithe, M., Benkhelifa, E., & Hameed, K. (2019). A systematic literature review of data governance and cloud data governance. *Personal and Ubiquitous Computing*, 23(5-6), 839–859. https://doi.org/10.1007/s00779-017-1104-3

Ashiq, M., Usmani, M. H., & Naeem, M. (2020). A systematic literature review on research data management practices and services. *Global Knowledge, Memory and Communication*, 71(8/9), 649–671. https://doi.org/10.1108/GKMC-07-2020-0103

Bernardo, B. M. V., Mamede, H. S., Barroso, J. M. P., & dos Santos, V. M. P. D. (2024). Data governance & quality management—Innovation and breakthroughs across different *Journal of Innovation and Knowledge*, 9(4), 100598. https://doi.org/10.1016/j.jik.2024.100598

Bhansali, N. (2013). *Data governance: Creating value from information assets.* Auerbach Publications.

BIMMInstitute (2021). *Research data governance policy*. Retrieved from https://www.bimm.ac.uk/research-data-governance-plan

Cheong, L., & Chang, V. (2007). The need for data governance: *A case study. ACIS 2007 Proceedings - 18th Australasian Conference on Information Systems* (pp. 999-1008). Australia. Retrieved from https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1110&context=acis2007

Creswell, J. W. (2014). *Qualitative, quantitative, and mixed methods approaches in the UK*. Sage Publications.

Cupoli, P., Earley, S., & Henderson, D. (2014). *Dama-dmbok2 Framework*. Retrieved from https://www.academia.edu/39943510/DAMA_DMBOK2_Framework

Dadiala, R., & Sinha, S. (2025). Data governance and modeling practices for intelligent water systems at a sewershed scale. *Water Practice and Technology*, 20(5), 1321–1337. https://doi.org/10.2166/wpt.2025.076

Davis, H. A., Kerkman, D., Hoberg, A. A., Countryman, M., Beaver, W., Bybee, K., Blum, J. M., & Knosp, B. M. (2025). Establishing data governance for sharing and access to real-world data: A case study. *JAMIA Open*, *8*(3), ooaf041. https://doi.org/10.1093/jamiaopen/ooaf041

Gibson, W., & Brown, A. (2009). *Working with qualitative data*. Sage Publications. https://doi.org/10.4135/9780857029041

Hassani, H., Huang, X., & MacFeely, S. (2025). Mapping the evolution of data governance scientific research. *Data and Policy*, 7, e51. https://doi.org/10.1017/dap.2025.10014

Javadi, M., Chung, M. S., Fard, N. G., & Miskat, M. (2023, December). A data governance literature review in education sector. In *2023 International Conference on Computational Science and Computational Intelligence (CSCI)* (pp. 1662-1669). IEEE.

Koltay, T. (2020). Quality of open research data: Values, convergences and governance. *Information*, 11(4), 175. https://doi.org/10.3390/info11040175

Kouper, I., Raymond, A., & Giroux, S. (2020). An exploratory study of research data governance in the U.S. *Open Information Science*, 4(1), 122–142. https://doi.org/10.1515/opis-2020-0010

Majidi, A., Naghshineh, N., Ismaili Ghivi, M. R., & Hashemi, M. (2017). Study of the foundation, models, and issues of research data curation and management in scientific and academic environments. *Human Information Interaction*, 4(2), 31–57. Retrieved from http://hii.khu.ac.ir/article-1-2598-fa.html [in Persian]

Marlina, E., Hidayanto, A. N., & Purwandari, B. (2022). Towards a model of research data management readiness in the Indonesian context: An investigation of factors and indicators through the fuzzy Delphi method. *Library & Information Science Research*, 44(1), 101141. https://doi.org/https://doi.org/10.1016/j.lisr.2022.101141

Marlina, E., & Purwandari, B. (2019). Strategy for research data management services in Indonesia. *Procedia Computer Science*, 161, 788–796. https://doi.org/https://doi.org/10.1016/j.procs.2019.11.184

Nai, S., Rifai, A., & Sadiq, A. (2025). Data governance, key insights, strategic challenges, and future imperatives. In *Data Governance, DevSecOps, and Advancements in Modern Software* (pp. 1–15). IGI Global. https://doi.org/10.4018/979-8-3373-0365-9.ch001

Odebrecht, C. (2025). Research data governance. The need for a system of cross-organisational responsibility for the researcher's data domain. *Data Science Journal*, 24, 12, 1-8. https://doi.org/10.5334/dsj-2025-012

Patel, D. (2016). Research data management: A conceptual framework. *Library Review*, 65(4-5), 226–241. https://doi.org/10.1108/LR-01-2016-0001

Solomonides, A. (2019). Research data governance, roles, and infrastructure. In *Clinical research informatics* (pp. 291-310). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-98779-8_14

Sothilingam, R., Pant, V., Shahrin, N., & Eric, S. K. (2021). Towards a goal-oriented modeling approach for data governance. In *PoEM (Forum)* (pp. 69-77). Retrieved from https://ceur-ws.org/Vol-3045/paper08.pdf

Stephens, J. J. (2018). *Data governance importance and effectiveness: Health system employee perception*. Doctoral Dissertation. School of Health Sciences, Central Michigan University. USA. Retrieved from https://www.proquest.com/openview/a80278e979fd9c9cc0f99805241401ea/1?pq-origsite=gscholar&cbl=18750

Sweeney, K. J. (2018). Re-Imagining data governance. In *Handbook of research on knowledge management for contemporary business environments* (pp. 251–278). IGI Global. https://doi.org/10.4018/978-1-5225-3725-0.ch014

Wende, K. (2007). A model for data governance–organising accountabilities for data quality management. ACIS 2007 proceedings. Paper *presented at the 18th Australasian Conference on Information Systems* (417-425). Australia. Retrieved from https://www.alexandria.unisg.ch/server/api/core/bitstreams/2fc24932-c72f-4405-9960-fda5bfd8d52e/content

Zhang, D., Song, B., Geng, H., Chen, Y., & Liu, H. (2025). The institutional evolution of chinese university data governance: An analytical framework based on historical institutionalism. *Education Sciences*, 15(7), 891. https://doi.org/10.3390/educsci15070891