RESEARCH DATA GOVERNANCE FOR RESEARCH PERFORMING ORGANIZATIONS IN MALAYSIA: A MODIFIED DELPHI APPROACH

NORZELATUN RODHIAH BINTI HAZMI

FACULTY OF ARTS AND SOCIAL SCIENCES
UNIVERSITI MALAYA
KUALA LUMPUR

2024

RESEARCH DATA GOVERNANCE FOR RESEARCH PERFORMING ORGANIZATIONS IN MALAYSIA: A MODIFIED DELPHI APPROACH

NORZELATUN RODHIAH BINTI HAZMI

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF ARTS AND SOCIAL SCIENCES
UNIVERSITI MALAYA
KUALA LUMPUR

2024

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **NORZELATUN RODHIAH BT HAZMI**

Matric No: **s2038059**

Name of Degree: **DOCTOR OF PHILOSOPHY**

Title of Project Thesis ("this Work"):

RESEARCH DATA GOVERNANCE FOR RESEARCH PERFORMING ORGANIZATIONS IN MALAYSIA: A MODIFIED DELPHI APPROACH

Field of Study: **LIBRARY AND INFORMATION SCIENCE**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the Universiti Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date: 10 December 2024

Subscribed and solemnly declared before,

Witness's Signature Date: 10 December 2024

Name:

Designation:

RESEARCH DATA GOVERNANCE FOR RESEARCH PERFORMING ORGANIZATIONS IN MALAYSIA: A MODIFIED DELPHI APPROACH

ABSTRACT

The rapid growth of research data and its influence on research practices have led to an increased recognition of the importance of research data governance (RDG) worldwide. Nevertheless, a lack of literature explicitly delineates RDG implementation, and practices undertaken by research performing organizations (RPOs). This study sought to fill the existing gap in the literature by exploring the implementation and practices of RDG by RPOs within a post-positivist research paradigm. To achieve this, two methods were employed: a content analysis of RDG/Research Data Management (RDM) policy documents through desk research, and a four-round modified Delphi method, involving systematic solicitation and collection of feedback from a pool of experts comprising research data practitioners. This was accomplished through a series of designed sequential surveys formulated based on the findings of desk research. These surveys used in the Delphi study were tailored to systematically explore RDG activities, focusing on governance tasks, areas of involvement, and decision domains, as well as the roles responsible for governing research data within RPOs in Malaysia. Additionally, it sought to understand the experiences and practices of research data practitioners in Malaysia RPOs in terms of governing research data. The study employs descriptive and inferential statistics, encompassing percentages, interquartile range, mean and median scores, mean differences, t-values, and significance levels obtained from the paired-sample t-tests. The analysis of responses unveiled a significant consensus among the experts on 106 out of 119 RDG task statements across various roles, nineteen (19) areas, and eight (8) decision domains deemed significant to RDG implementation within RPOs. The gap analysis between actual implementation and perceived importance reveals that while certain tasks

are considered highly important, they may not be implemented to the same extent, as all mean differences show negative values. Besides the RDG activities, key players, structural positions, and ideal nomenclatures of identified RDG roles were determined. The aim of the study was to develop an RDG framework based on the findings obtained. By providing an RDG framework that can be used as a set of best practices, this study can assist RPO leaders in considering implementing RDG and its efforts in their organizations.

Keywords: Research data governance; research data management; data stewardship; research data; research organizations

SDGs: Quality Education; Industry, Innovation, and Infrastructure; Peace, Justice, and Strong Institutions; Partnerships for the Goals

TADBIR URUS DATA PENYELIDIKAN UNTUK ORGANISASI PELAKSANA PENYELIDIKAN DI MALAYSIA: PENDEKATAN DELPHI YANG DIMODIFIKASI

ABSTRAK

Pertumbuhan pesat data penyelidikan dan kesannya terhadap amalan penyelidikan telah membawa kepada pengiktirafan yang semakin meningkat mengenai kepentingan tadbir urus data penyelidikan (TUDP) di seluruh dunia. Walau bagaimanapun, kekurangan literatur yang secara eksplisit menjelaskan pelaksanaan TUDP dan amalan yang dijalankan oleh organisasi pelaksana penyelidikan (OPP). Kajian ini bertujuan untuk mengisi jurang yang sedia ada dalam literatur dengan meneroka pelaksanaan dan amalan TUDP oleh OPP dalam paradigma penyelidikan pasca-positivisme. Untuk mencapai matlamat ini, dua kaedah telah digunakan: analisis kandungan dokumen dasar TUDP/Pengurusan Data Penyelidikan (PDP) melalui kajian meja, dan kaedah Delphi yang dimodifikasi sebanyak empat pusingan, yang melibatkan pengumpulan maklum balas secara sistematik daripada sekumpulan pakar yang terdiri daripada pengamal data penyelidikan. Ini dicapai melalui satu siri tinjauan berurutan yang direka dengan teliti berdasarkan penemuan kajian meja. Tinjauan ini digunakan dalam kajian Delphi yang dimodifikasi untuk meneroka aktiviti TUDP secara sistematik, dengan memberi penekanan kepada tugas tadbir urus, bidang penglibatan, dan domain keputusan, serta peranan yang bertanggungjawab untuk mengurus data penyelidikan dalam OPP di Malaysia. Selain itu, kajian ini bertujuan untuk memahami pengalaman dan amalan pengamal data penyelidikan di kalangan OPP di Malaysia berkaitan dengan tadbir urus data penyelidikan. Kajian ini menggunakan statistik deskriptif dan inferensial, merangkumi peratusan, julat interkuartil, skor min dan median, perbezaan min, nilai-t, dan tahap kepentingan yang diperoleh daripada ujian t-sampel berpasangan. Analisis tindak balas mendedahkan konsensus yang signifikan dalam kalangan pakar terhadap 106

daripada 119 kenyataan tugas TUDP merentasi pelbagai peranan, sembilan belas (19) bidang, dan lapan (8) domain keputusan yang dianggap penting untuk pelaksanaan TUDP dalam OPP. Analisis jurang antara pelaksanaan sebenar dan kepentingan yang dirasakan menunjukkan bahawa walaupun tugas-tugas tertentu dianggap sangat penting, ia tidak dilaksanakan pada tahap yang sama, kerana semua perbezaan min menunjukkan nilai negatif. Selain aktiviti TUDP, pemain utama, kedudukan struktur, dan nomenklatur yang ideal bagi setiap peranan telah dikenal pasti. Tujuan kajian ini adalah untuk membangunkan rangka kerja TUDP berdasarkan penemuan yang diperoleh. Dengan menyediakan rangka kerja TUDP yang boleh digunakan sebagai satu set amalan terbaik, kajian ini dapat membantu pucuk pimpinan OPP dalam mempertimbangkan pelaksanaan TUDP dan usaha-usahanya dalam organisasi mereka.

Kata kunci: Tadbir urus data penyelidikan; pengurusan data penyelidikan; penyelenggaraan data; data penyelidikan; organisasi pelaksana penyelidikan

***SDGs*:** Pendidikan Berkualiti; Industri, Inovasi, dan Infrastruktur; Keamanan, Keadilan, dan Institusi Kukuh; Perkongsian untuk Matlamat

Dedication

To my late father, who always encouraged me to embark on this wonderful journey...

To the adults in the family: Firdausi, Mak, Mama, Papa, Nurul, Eyla, Ida, Amad, Amir, Kak Sara, Fahmi, thank you for everything. Love you all so much!

To my beloved children, let this achievement inspire you to always pursue knowledge without boundaries or limits, and to never surrender!

To the RPOs (particularly in Malaysia), this research is dedicated to you. With RDM being a fresh concept for many, I hope that the tailored framework serves as a valuable guide for you to either embark or refine existing RDG practices within your respective organization.

ACKNOWLEDGEMENTS

In the name of Allah, the most beneficent, the most merciful.

Alhamdulillah..

All praise and grace be to Allah, the Lord of all mankind. Without His divine permission and guidance, this long-held dream would never have come to fruition. It is through His mercy, wisdom, and blessings that I have been able to reach this milestone. I am deeply grateful for the strength and perseverance He has granted me throughout this journey.

I would also like to express my deepest gratitude to those who have supported me throughout the journey of completing this dissertation. Without their guidance, encouragement, and unwavering support, this achievement would not have been possible.

First and foremost, I would like to thank my parents. *Mak (arwah)Ayah*, your unconditional love, sacrifices, belief, and prayers for me have been my greatest source of strength. You have always encouraged me to pursue my dreams and have supported me in every possible way. This accomplishment is as much yours as it is mine. To my parents-in law, thank you for your prayers, support, and kind wishes.

To my beloved husband, thank you for your constant support, patience, understanding, and du'a. Your encouragement and faith in me kept me motivated, even during the most challenging times. Your love and companionship have been my anchor, and I am incredibly grateful for your presence in my life.

To my wonderful children, thank you for your love and for bringing so much joy into my life. Your smiles, hugs, and laughter have been a constant source of motivation and happiness. You have taught me the true meaning of perseverance and have been my inspiration to keep going.

I extend my heartfelt appreciation to my family, whose collective support has been invaluable. To my siblings, thank you for being my cheerleaders and for always believing in me. Your words of encouragement have been a source of inspiration. To my extended family, thank you for your prayers and unwavering support.

Most importantly, I would like to express my deepest gratitude to my supervisors. *Jazakillahu khairan kathira* Professor Dr. Abrizah Abdullah, this thesis would have not been completed without your advice and unfailing patience. Your mentorship has been a crucial factor in my academic growth, and I am profoundly grateful for the opportunity to learn from you. Your guidance, wisdom, and expertise have been instrumental in shaping this dissertation. To my co-supervisor Associate Professor Dr. Yanti Idaya Aspura Mohd Khalid, thank you for your constructive criticism, and for always pushing me to strive for excellence. As someone once told me: “Without good supervisors, you are nothing and your thesis is nothing”.

I would also like to extend my sincere appreciation to the Ministry of Higher Education for providing me with the HLP scholarship that made this academic pursuit possible. Additionally, I am deeply grateful to IIUM and Dar al-Hikmah Library for granting me study leave with full salary. Your support has been crucial in enabling me to focus on my studies without the burden of financial stress and allowed me to dedicate the necessary time and effort to complete this dissertation without compromising my financial stability.

To my PhD peers, thank you for being my pillars of strength. Your encouragement, laughter, and brotherhood have made this journey enjoyable and less challenging. Your support has been invaluable, and I cherish each one of you for the unique role you have played in my study life.

To each and every faculty member, thank you for your support and motivation throughout this journey. Your encouragement and belief in my potential have been instrumental in my success.

I would like to express my sincere gratitude to the examiners who were involved at various stages of this process, including the proposal defense, candidature defense, and viva. I am truly fortunate to have had such esteemed scholars as Professor Stephen Pinfield, Professor Stephen Mutula, and Professor Kiran Kaur as my examiners during the viva. To all of the examiners, I would like to express my heartfelt appreciation: 'Your invaluable feedback and insightful comments have greatly enriched the quality of this research. I truly appreciate your willingness to dedicate your time and expertise to this process'.

Finally, I would like to thank everyone who contributed to this dissertation, whether directly or indirectly. Your support, in its many forms, has been deeply appreciated. Thank you all for being part of this incredible journey!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ASM	:	Academy of Sciences Malaysia
DMP	:	Data Management Plan
FAIR	:	Findable, Accessible, Interoperable, and Reusable
MOSP	:	Malaysia Open Science Platform
OPP	:	Organisasi Pelaksana Penyelidikan
RDG	:	Research data governance
RDM	:	Research data management
RPO	:	Research performing organization
TUDP	:	Tadbir urus data penyelidikan
ToT	:	Training of Trainers

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

"And do not mix the truth with falsehood or conceal the truth while you know [it]".
(Surah Al-Baqarah, 2:42)

1.1 Introduction of Study

Research data is seen as a valuable asset for research performing organizations (RPOs). Recognizing the intrinsic value of research data, it becomes imperative for RPOs to establish frameworks for research data governance (RDG), ensuring the comprehensive protection of research data. The establishment of effective RDG not only safeguards the integrity and confidentiality of research data but also fosters a conducive environment for sustained research excellence within the organization (Al-Ruithe et al., 2016; Liu et al., 2020). This acknowledgment underscores the crucial role that RDG plays in promoting responsible and secure handling of research data, thereby enhancing the overall credibility and impact of research activities within RPOs.

Considering the aforementioned, the purpose of this study was to examine how RDG tasks are implemented in Malaysia RPOs, the significance of each task, and to identify the RDG structure that is ideal for these organizations to ensure the quality and accessibility of research data during RDG implementation and practices. This chapter provides context for the research by outlining the subject, its purpose, aims and objectives, research questions, problem statement, and research gaps. Additionally, it discusses the research methodology, scope, and significance of the study, establishing a comprehensive foundation for the subsequent exploration of RDG in Malaysia RPOs.

1.2 Background of Study

The study emphasizes the comprehensive nature of data management by integrating the eleven aspects of Data Management Body of Knowledge (DMBoK), which covers data governance, data modeling & design, data storage & operations, data security, data integration & interoperability, document & content management, reference & master

data, data warehousing & business intelligence, metadata, data quality, and data architecture (DAMA International, 2017). In this body of knowledge, data governance is the highest level of planning activity above all the other ten components, as shown in Figure 1.1 (DAMA International, 2017; Wang et al., 2018). The choice of DAMA's DMBoK framework is justified by its holistic approach to data management, ensuring that all critical aspects are systematically addressed. This integration fosters a structured and standardized methodology for managing data effectively, promoting consistency, quality, and security across various data-related processes. Additionally, data governance prioritizes the needs of data users, their modes of data usage, access authority, and other compliance requirements (Abraham et al., 2019). This structured approach ensures that data is managed effectively and efficiently across the organization or RPO, enhancing the overall reliability and robustness of its data management practices.

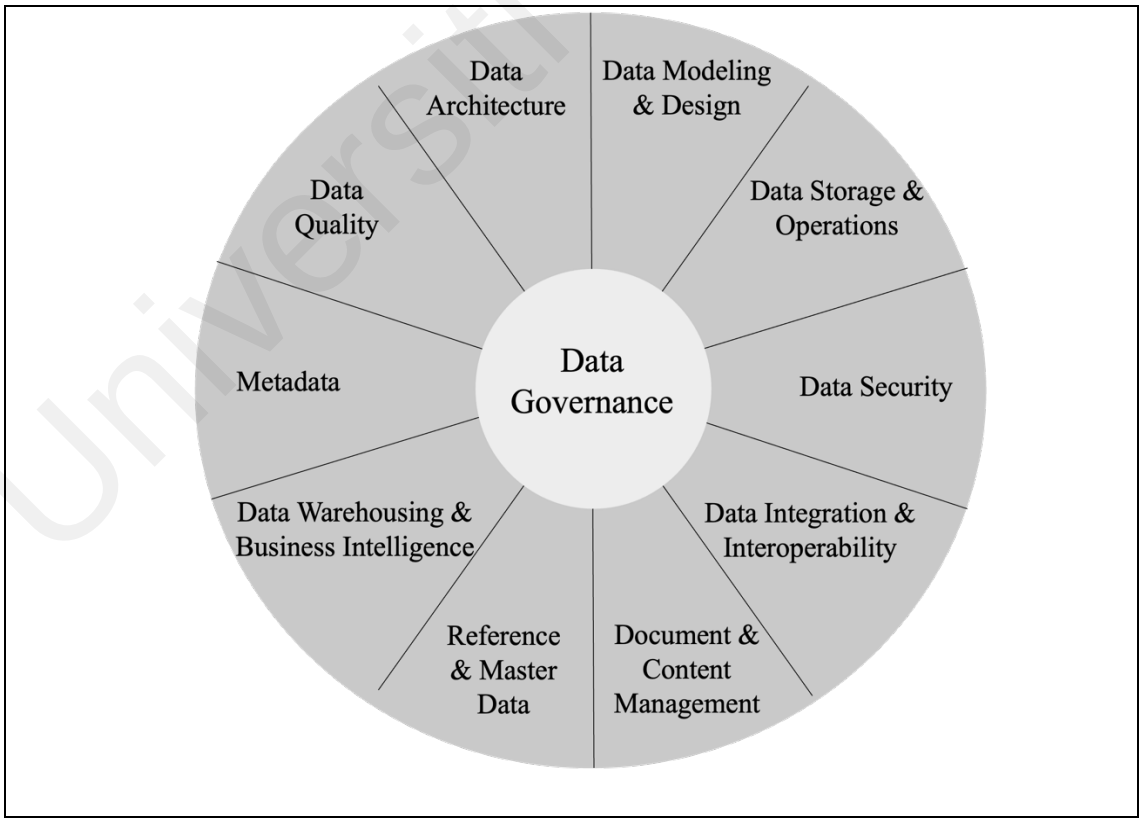


Figure 1.1: The DAMA-DMBOK2 Data Management Framework (DAMA International, 2017)

While data governance and data management are distinct concepts, they are closely connected, as shown in Figure 1.2. Data governance ensures the proper management of data (oversight), whereas data management involves handling data to meet specific objectives (execution) (DAMA International, 2017). The topic of data governance has recently received increased attention in the scientific community and among practitioners because data quality is increasingly a fundamental requirement for organizations to achieve various strategic business requirements (Otto, 2013). Appropriate data management practices enhance high-quality research, resulting in high-quality data (Lefebvre & Spruit, 2021; Marlina & Purwandari, 2019; Wong et al., 2020) and improved reproducibility and knowledge dissemination (Leonelli, 2018). To ensure data quality, apart from facilitating access and use to the community, the value of data quality improvement, loss minimization, improved resource utilization, re-examination and reuse, and increased citations must also be demonstrated (Austin et al., 2021). When high-quality data is widely accessible, it creates greater value and fosters transparency, collaboration, and innovation. Data transparency is a quality research principle that must be respected (Belcher et al., 2016).

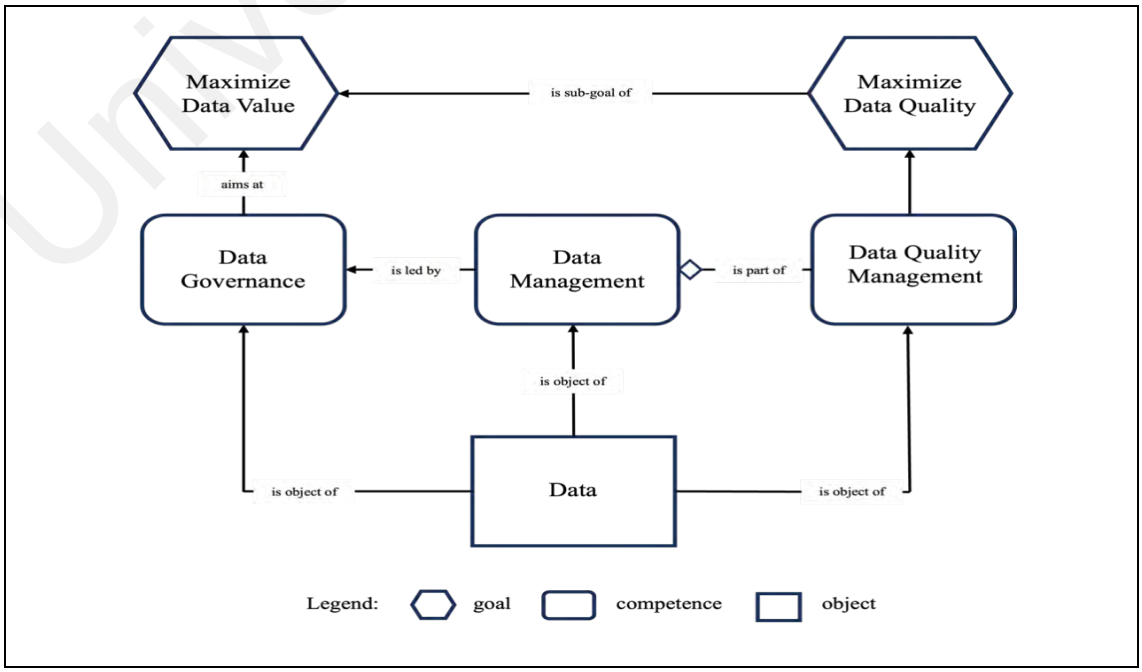


Figure 1.2: Data governance and data management (Otto, 2013)

Research data underpins the research activities undertaken within RPOs. RPOs, entities whose main activity is to carry out research and experimental development (Cruz-Castro et al., 2020), which include universities, research institutes, and similar organizations committed to the pursuit of research activities, producing multidisciplinary research outputs from scientific research across multiple disciplines (Jamiu et al., 2020; Makhlouf-Shabou, 2017). The mission of these organizations is to conduct research and disseminate its findings via teaching, publication, or technology transfer. In the rapidly evolving landscape of research, the adoption of open science principles has become increasingly essential. Open science policies and practices that promote the open availability of research data play a crucial role in accelerating the expansion of knowledge and its application (Angelopoulos & Pollalis, 2020; Janssen et al., 2017) to societal problems (Fecher et al., 2015; Rosenbaum, 2020), thereby enhancing the visibility and reputation of the organization and its researchers and scholars (Zuiderwijk et al., 2020). Ensuring broader public access to research data is a fundamental goal that can advance research and education. This aligns well with the mission of the RPOs, which is to generate and disseminate new knowledge.

The effective management and governance of research data are imperative for RPOs to uphold the integrity, credibility, and impact of their research efforts. RPOs should manage their data for a long-term competitive advantage (Singh et al., 2021). RDG implementation is vital for guaranteeing research data quality, accuracy, and usability and ensuring ethical use, privacy, and confidentiality (Hendey et al., 2018). RDG must be reasonable to ensure the whole management of the security of the data utilized in an RPO (Lefebvre et al., 2018). RDG also guarantees that research data is easy to access and creates trust in the data provided by ensuring it is valuable for decision-making (Jamiu et al., 2020). Plus, research data, a critical class of digital content, must be correct, complete, and accurate (Omar & Almaghthawi, 2020). Stakeholders could profit from the

availability of meaningful data if RDG procedures are implemented correctly (Singh et al., 2021). RPOs can also move faster, make better choices, and provide better insights into user behaviour, all of which help to boost efficiency (Omar & Almaghthawi, 2020). Thus, it is important to make sure the data is well managed, following the policies and standards of the organization (DAMA International, 2017).

Without the implementation of data management and governance, there will be a negative impact on RPOs and a decrease in the value of the data. If research data is not properly captured, it will remain fragmented, siloed, and retained by individual researchers. The benefits of research data are diminished as well if they are not curated, captured, and made available to a larger community, especially in managing and using scientific data collected through government funding by fulfilling its responsibilities to taxpaying citizens (Pisani & AbouZahr, 2010).

However, getting an organization's data governance initiatives off the ground must have been difficult (Alhassan et al., 2018). In order to have an effective RDG, members of the RPOs need to change behaviour to ensure data governance processes and principles are adhered to (Benfeldt et al., 2020). Otherwise, ensuring the data flows across departments with the defined processes will be difficult. RPOs must develop data policies, procedures, and guidelines to overcome these obstacles, utilize resources and processes, and create and expand data services and infrastructure. A well-defined policy and accessibility standards will ensure that only the appropriate person receives the correct data at the appropriate time (Brous et al., 2016; McDonald, 2018). Obtaining buy-in from stakeholders, particularly the institution's management, appears to facilitate these initiatives (Mansfield-Devine, 2017).

1.3 Context of the Study: Advancing Research Data Governance in Malaysia

Managing data is similarly to managing other assets such as human resources, which necessitates mandated controls, measurable outcomes and necessary reporting (Alhassan et al., 2018; Khatri & Brown, 2010). As a result, every RPO must implement effective data governance and stewardship. This is especially true for Malaysia's RPOs, which include at least 20 public universities and 390 private higher education institutions (MOHE, 2024), alongside numerous research institutions. The government emphasizes the management and sharing of scientific research data, viewing it as a national asset (MOSP, 2020). In contemporary times, the publication of research articles is not the only option for scholarly communication. Research data is now recognized as a valuable national asset that requires proper protection and management, entailing the adoption of strategies that promote openness and sharing.

The Government of Malaysia has actively promoted cultivating a knowledge-based economy by emphasizing data-driven research and development. The management of research data has been effectively and securely executed through the implementation of multiple policies and standards. Data management and sharing were included in the Malaysian Code of Responsible Conduct in Research in 2016 (National Science Council, 2020). The Code mandates that researchers assure data accuracy, completeness, dependability, safe storage, and data sharing upon request. The Ministry of Science, Technology, and Innovation's (MOSTI) National Policy on Science, Technology & Innovation (NPSTI) 2013-2020 (MOSTI, 2013) and Malaysia Education Blueprint (2015-2025) also emphasize research data sharing and collaboration (Ministry of Education Malaysia, 2015). Plus, NPSTI 2021– 2030 mentions open data as one of its initiatives, emphasizing research data (MOSTI, 2021).

Recognizing the importance of research data in harnessing the potential impact of open science, the Malaysia Open Science Platform (MOSP) pilot initiative was established on November 7, 2019. This initiative, led by theAcademy of Sciences Malaysia (ASM), aimed to centralize research data storage, management, and sharing (MOSP, 2020), thereby laying the foundation for enhanced collaboration and innovation in scientific endeavours. ASM launched the Malaysia Open Science Alliance (MOSA) to help MOSP become a strategic, transformative effort to boost Malaysia's STI Collaborative Ecosystem. The first phase of this initiative (2020-2021) was led by the MOSA and administered by ASM, connecting research data, researchers, and publications between five research universities and 15 Ministry of Energy, Science, Technology, Environment, and Climate Change (MESTECC) research institutes and agencies (Abd Rahman, 2019; MOSP, 2020). The MOSP was officially launched on May 16, 2023. Additionally, the ASM and MOSA Working Group on Capacity Building and Awareness initiated the Training of Trainers (ToT) Program on Data Stewardship for Open Science on September 21, 2020, with the primary objective of ensuring that future data stewards possess the necessary competencies and expertise for proficient open science data stewardship (MOSP, 2020).

As a result, in recent years, open science initiatives have been evolving in Malaysia, and researchers started to become aware of having these initiatives (Ahmed & Othman, 2021; Hodonu-Wusu et al., 2020). For instance, the University of Science, Malaysia (USM) has encouraged its researchers to develop Data Management Plans (DMPs), store their research data, and share it with others. They can either upload their data to their institutional or other free data repositories, e.g., Figshare (https://figshare.com/), Mendeley Data (https://data.mendeley.com/), and Zenodo (https://zenodo.org/) (Che Jaafar, 2021). However, according to Che Jaafar (2021), there is no mandate to make it compulsory for researchers to share data collected throughout the research.

Malaysia has embarked on a progressive journey towards advocating open science since 2019, recognizing the transformative potential of transparent and collaborative research practices. Embracing the principles of open science entails not only the dissemination of research findings but also extends to the sharing of underlying research data. As Malaysia transitions into a more open science-oriented landscape, there arises a compelling need to explore and strengthen the foundations of RDG.

Therefore, the study concurs on the importance of implementing RDG among Malaysia RPOs for better management and protection of research data (Jim & Chang, 2018; Lefebvre et al., 2018; Omar & Almaghthawi, 2020) to advance local government initiatives in supporting open science (MOSP, 2020) to foster a greater research community who keen to openly share their research data with wider users for maximizing societal value. A dedicated data governance framework focuses on research data is required for them to successfully implement RDG within their organization.

1.4 Statement of the Problem

Despite the crucial nature of data governance, several studies have uncovered a noticeable lack of knowledge and understanding in this domain (Abraham et al., 2019; Eke et al., 2022; Kuzio et al., 2022; Schöpfel et al., 2017). As demonstrated by Al-Ruithe et al. (2019), areas related to data governance that have gained increased attention and significance in recent years still remain inadequately explored. Abraham et al. (2019) assert that a holistic picture of data governance is currently lacking, which may guide practitioners and researchers alike. Likewise, when compared to other RDM bodies of knowledge, there is a notable lack of attention paid to RDG in the Library and Information Science (LIS) literature (Borda et al., 2020; Donaldson & Koepke, 2022; Tang & Hu, 2019; Wolski et al., 2017), indicating a scarcity of research on the topic (Al-Ruithe et al., 2019). This attention deficiency is particularly significant given the importance of RDG

within LIS. Far more attention is being paid to other RDM bodies of knowledge, such as data quality (Azeroual, 2020; Wagner et al., 2021), data sharing (Elsayed & Saleh, 2018; Saeed & Ali, 2019), and data literacy (Steinerová & Ondrišová, 2019; Vilar & Zabukovec, 2019).

Since RDG is regarded as a critical service, more research using mixed and qualitative methods is necessary to understand the subject in depth (Ashiq et al., 2020). Further research should be conducted on the implementation and monitoring of data governance activities rather than on the definition of concepts (Lefebvre et al., 2018; Alhassan et al., 2018; Brous et al., 2016; Alhassan et al., 2016), by defining the activities' priorities (Omar & Almaghthawi, 2020). These activities address who is accountable for governing data and their duties and responsibilities in ensuring that data governance policies and standards are adhered to, as specified in a data governance policy framework (Alhassan et al., 2016).

In addition, the current frameworks do not meet the necessity for governing research data in particular, as their scope varies across other domains (Abraham et al., 2019; Ahmadi et al., 2022; Alhassan et al., 2018; Al-Ruithe et al., 2016; Brous et al., 2016; Korhonen et al., 2013; Lee et al., 2017; Peng et al., 2016). A comprehensive data governance framework is crucial to assist scholars, organizations, and the government in making decisions (Thorat & Van Brakel, 2019). This includes having well-defined guidelines on best practices for RDM (Chigwada et al., 2017). As highlighted by Lefebvre et al. (2018), further research should collect more evidence from other research institutions worldwide following the policy screening and exploratory case study approach.

Meanwhile, other literature stresses that the activities should encompass a broader range of issues (Lefebvre et al., 2018), including research data ownership (Van Zeeland & Ringersma, 2017), active data management, data selection for long-term preservation, and data access via catalogues and repositories (Liu et al., 2020). Additionally, expanding the scope to include funders, data centres, and institutions (Naughton & Kernohan, 2016), addressing the purpose and missions of research, social and cultural norms, conflict resolution methods, and dilemmas in data sharing and use (Kouper et al., 2020), and integrating research data support with the vast diversity of research projects (Lefebvre et al., 2018) are all emphasized.

The existing literature highlights several critical issues concerning this subject matter. Previous scholars claim that RDG is better observed in developed countries such as United States (Jim & Chang, 2018), United Kingdom, Australia (Liu et al., 2020), Canada, Germany, Ireland, the Netherlands, New Zealand (Cox et al., 2017) than in developing countries such as India (Bunkar & Bhatt, 2020), Kenya (Masinde et al., 2021), Iraq (Mohammed & Ibrahim, 2019), Saudi Arabia (Elsayed & Saleh, 2018), Zimbabwe (Chigwada et al., 2017; Machimbidza et al., 2022), Kenya (Anduvare & Mutula, 2019), Indonesia (Cahyaningtyas & Priyanto, 2021; Manik et al., 2022) and Pakistan (Piracha & Ameen, 2019), since most developing country institutions lack RDM policies. However, although RDM is maturing in developed countries, some argue that it is still in its early phases (Ashiq et al., 2020). This perception is due to the numerous challenges associated with RDM methods and services (Cox et al., 2019a), which necessitate active collaboration among stakeholders and university services departments to address these challenges and issues (Ashiq et al., 2020).

Moreover, data governance arrangements differ significantly amongst universities; some established new divisions to initiate the projects (Jim & Chang, 2018), while others started with their institutional research (Briner & Rome, 2018) or information technology (IT) department (Cervone, 2017) or information governance, providing IT support (Jim & Chang, 2018). Meanwhile, allocating specific data tasks to specific personnel is ambiguous, and those responsible for research data decision-making across RPOs remain uncertain (Andrikopoulou et al., 2022; Kouper et al., 2020; Perrier et al., 2018), including RPOs in Malaysia. Existing literature (Norbib & Abu Bakar, 2021) has proposed a data governance model for the Ministry of Education, Malaysia, by comparing and analyzing several models and frameworks to uncover trends and patterns. However, there is no solid evidence and empirical studies on who governs research data and what entities are involved in RDG practices within RPOs.

To sum up, RDM supports open science, and the rapidly evolving landscape of RDM has generated interest in leveraging it to enhance RDG practices. It is evident that RDG offers a solution to ensure effective management of research data (Van Zeeland & Ringersma, 2017). This involves adopting appropriate technical standards, methods, and architecture to facilitate the management, sharing, and reuse of research data (Ng'eno & Mutula, 2018) while also determining future development directions (Zhou, 2018). As outlined by the European Commission (2019), a readily accessible governance framework should encompass distinct institutional, executive, and advisory responsibilities, thereby empowering the scientific community.

However, a gap exists in the implementation of RDG and the need for further research to understand and improve data governance practices, particularly in developing countries and within RPOs. Therefore, this study aims to address several key issues pertaining to RDG within Malaysia RPOs. Primarily, it seeks to identify RDG activities and assess

their implementation level and perceived significance among data practitioners. Furthermore, it explores the examination of roles and organizational structures linked with these activities, ultimately aiming to formulate an RDG framework. The absence of such a framework not only hampers effective decision-making among data practitioners and decision-makers but also contributes to uncertainty regarding the governance of research data and the entities involved in RDG practices.

1.5 Purpose, Research Objectives and Research Questions

The study aimed to explore how RPOs manage their research data by identifying key data governance activities and the roles associated with them. The study began with secondary research to enrich the knowledge base by examining the governance of research data within top-leading RPOs. Building on this foundation, the main objectives were to establish structural and functional dimensions based on expert opinions gathered through multiple Delphi rounds, ultimately formulating an RDG framework tailored for RPOs in Malaysia. The framework serves as a guide for these organizations, especially as they embark on implementing RDM in their institutions or refining existing RDG practices.

Therefore, to achieve the aforementioned research aim, the following research objectives were addressed throughout the study:

RO1: To investigate research data governance activities in relation to both actual implementation and perceived importance among data practitioners within research performing organizations.

RO2: To identify the stakeholders of research data governance roles within research performing organizations, including their governance structure.

RO3: To develop a research data governance framework for research performing organizations.

In this context, the following research questions were required to guide the achievement of the research objectives. By answering these research questions, the research aim was ultimately fulfilled:

RQ1: How are research data governance activities currently being implemented by data practitioners in research performing organizations?

RQ2: What research data governance activities do data practitioners consider significant for research performing organizations?

RQ3: Which stakeholders do data practitioners in research performing organizations consider responsible for research data governance roles?

RQ4: How do data practitioners identify the structural positions of research data governance roles within research performing organizations?

1.6 Research Methodology

This study employed a modified Delphi method as the primary data collection mechanism, utilizing a qualitative approach to derive quantitative results. Given the scarcity of information on RDG in Malaysia, a suitable approach is required to achieve consensus among data practitioners on RDG practices relevant to RPOs. Rowe and Wright (2011) advocate for Delphi approach as it effectively gathers expert opinions and achieves consensus. This method allows anonymity, encouraging professionals to freely express themselves without fear of criticism or bias (Lim et al., 2020; Markmann et al., 2021; Olsen et al., 2021). The study was modified in several ways: a) the initial round started with a set of carefully selected items based on prior desk research, instead of the traditional open-ended questionnaire; b) in Round I, panelists were asked about the actual implementation of RDG tasks at their institutions, without the need for consensus or agreement; and c) in Round III, the instruments were designed to gather information from a different perspective, incorporating newly introduced dimensions. Recruiting

appropriate experts is critical to ensuring valid results, and a purposive sampling and snowballing methods were used to form a panel of data practitioners with research data experience throughout its lifecycle. The panel consists of researchers, librarians, policymakers, IT and research officers.

The current study commenced with desk research to improve understanding of the RDG elements before implementing the modified Delphi study as field research. Furthermore, it served as a foundation for developing instruments to be utilized in the Delphi study. The process of conducting desk research entails summarizing and synthesizing previously published RDG/RDM policies from leading RPOs worldwide (Goundar, 2012; Kabir, 2016; Woolley, 1992). This method was frequently recommended as a starting point for the modified Delphi study because it provides a standard against which primary data findings can be compared (Kabir, 2016).

The detailed methodology is outlined in Chapter 3, covering all aspects of the study design, data collection, and analysis procedures.

1.7 Scope of the Study

RDG stands as a critical pillar in contemporary research endeavours, influencing the systematic management and decision-making processes surrounding research data. This study embarked on an exploration of RDG activities within Malaysia RPOs, delineating its scope through specific objectives:

a) Focus on exploration: The study concentrated on delving into RDG activities within Malaysia RPOs, aiming to uncover how these organizations manage and govern their research data.

b) Identification of key elements: The primary objective was to identify key activities and entities responsible for decision-making in the realm of research data, offering insights into the organizational structures that influence RDG.

c) Development of RDG framework: A key attempt was to develop an RDG framework that is accepted and approved by a panel of data practitioners in Malaysia. This was achieved through the utilization of the Delphi study methodology.

d) Inclusive participant selection: The study ensured a broad perspective by including participants from diverse backgrounds, each possessing varying levels of knowledge and experience with RDG processes and practices. The purposive selection of participants in the modified Delphi method aimed to ensure a depth of knowledge on the topic.

e) Crucial inclusion of management levels: Recognizing the importance of different perspectives, the study included individuals from various levels of RDM—strategic, tactical, operational, and researcher. This approach was deemed crucial for ensuring the validity and reliability of responses.

The findings of the study aid in developing a framework that considers necessary implementations and ultimately contributes to the advancement of RDG practices.

1.8 Significance of the Study

This study was conducted to explore the activities and entities involved in governing research data for RPOs in Malaysia, with the anticipation of contributing to the development of an RDG framework. The findings obtained from this study play a crucial role in enhancing the RDG body of knowledge and offering a collection of best practices tailored for data practitioners.

By gaining a deeper understanding of the RDG discourse through this study, data practitioners in RPOs are better equipped to implement effective RDG practices. Simultaneously, other organizational stakeholders, including librarians, can leverage this knowledge to develop educational modules that equip researchers with the necessary skills and knowledge for RDG and RDM.

Furthermore, the anticipated benefits of this study extend to aiding organizations in complying with regulatory requirements and adhering to established data protection and consideration standards. Effective RDG practices, as identified in this study, can lead to faster data access, enhanced data protection, and adherence to the principles of FAIR data access, ensuring that only authorized individuals can access data.

Specifically targeting Malaysia RPOs, the study proposes that these organizations can utilize the developed framework resulted from this study to either establish new or revise existing RDG practices. This, in turn, is expected to expand access to research data to a broader community, ensuring that the data is of high quality and trustworthy. Ultimately, researchers, taxpayers, and society can all benefit from research data that is readily accessible, usable, and reusable. Beyond data protection, the study suggests that effective RDG can positively impact social change by reducing data errors and ensuring more reliable data transfer between researchers and users. This has the potential to benefit society at large.

1.9 Operational Definitions

The following terms are operationally defined for the purpose of this study:

Data practitioner: Any stakeholder who interacts with research data and supports it at any stage of its life cycle, from generation to management to consumption.

Research data governance: An organizational framework that formalizes a set of data policies and processes for research data management throughout its entire life cycle and emphasizes the exercise of authority and control over research data-related decisions.

Research data governance activities: The specific actions or tasks carried out to define, implement, and monitor effective RDG within an (research) organization, encompassing various governance areas and decision domains.

1.10 Definition of Terms

The following is a glossary of important terms that may be unfamiliar to readers. Throughout the study, such terminologies were often used.

Data governance: The exercise of authority and control over the management of research data to enhance its value while minimizing associated costs and risks (Abraham et al., 2019).

Data stakeholders: Individuals, groups, or organizations impacted by data governance practices (Micheli et al., 2020)

Open science: A realistic means of making data open to the wider public needs to ensure that the data that are most relevant to the public are accessible, intelligible, assessable and usable for the likely purposes of non-specialists (British Royal Society, 2012).

Research data: Any material collected, observed, or created for analysis in order to validate original research results (Drysdale, 2019; Kim, 2020; Shitole et al., 2019; Singh et al., 2018), regardless of the form of media on which they may be recorded (Vanderbilt University, 2021).

Research data management: The process through which a researcher progresses from developing a management strategy for the research data towards an implemented, sustainable business plan for working with active data, as well as storing, archiving and cataloguing data upon completion of a research project (Willaert et al., 2019).

Research performing organizations: Entities whose main activity is to carry out research and experimental development (Cruz-Castro et al., 2020), which include universities, research institutes, and similar organizations committed to the pursuit of research activities, producing multidisciplinary research outputs from scientific research across multiple disciplines (Jamiu et al., 2020; Makhlouf-Shabou, 2017).

1.11 Organization of the Thesis

The organization of the thesis is presented below to provide an overview of the content discussed in each and to stimulate interest in reading the entire chapters. This thesis consists of eight chapters, each commencing with an introduction and concluding with a chapter summary. The chapters are as follows: Introduction, Literature Review, Methodology, Desk Research Analysis Exploring Research Data Governance Activities, Research Data Governance Activities in Malaysia Research Performing Organizations, Research Data Governance Roles and Structural Positions in Malaysia Research Performing Organizations, Research Data Governance Framework, and Discussion and Conclusions. Detailed descriptions for each chapter are outlined in the sections that follow.

In Chapter One, a brief introduction and background information are presented regarding RDG. The comprehensive exploration covers various aspects of the study, including the problem statement, research objectives, research questions, research methodology, scope and limitations, research rationale, justification, and the significance of the study. Additionally, selected terms are defined, and a summary is provided.

Chapter Two presents a thorough examination of previous literature related to data governance in general and, more specifically, research data governance, was conducted to identify gaps that the present study aims to address. The critical review encompassed earlier works to ensure a comprehensive understanding of the subject of research data governance. The chapter concludes with a discussion on the identified gaps in the literature, providing clarity on the objectives of the study.

Chapter Three provides a comprehensive overview of the research methodology employed in this study. It explores various facets, encompassing the research paradigm, research design, the main two phases of data collection; comprising desk research and a modified Delphi study, participant recruitment, the research instrument, data collection methods, approaches to data analysis, the rationale behind the chosen research methodology, considerations of reliability, validity and trustworthiness, ethical considerations, and concludes with a summary.

Within Chapter Four, a detailed exploration is undertaken on data collection of the desk research to examine policy documents on RDG/RDM from leading RPOs worldwide. The analysis method involves a content analysis approach, and the entire process of data analysis is exhaustively explained. The overarching goals are to gather extensive background information on RDG implementation and practices, utilizing the discovered insights for the development of Delphi instruments. The presentation of findings is conducted in a thorough and detailed manner.

Chapter Five extensively examines the findings derived from the Delphi study, with a primary emphasis on RO1. The central objective is to explore the functional dimension of RDG activities, covering tasks, areas, and decision domains. This exploration encompasses both the implementation and importance of RDG activities within RPOs in Malaysia. The chapter examines discussions surrounding RDG implementation among Malaysia RPOs, the importance of RDG activities with insights from the panelists, the progression towards research-derived descriptions for RDG areas and decision domains, and a comprehensive gap analysis comparing the perceived importance to the actual implementation of these activities.

Chapter Six focuses on the discussion of RO2, which explores structural dimension of RDG roles such as identifying RDG stakeholders and identifying structural positions for each role, advancing towards research-derived descriptions for RDG roles, and establishing a nomenclature for RDG roles. This comprehensive analysis aims to offer a thorough understanding of the roles and organizational structures associated with RDG.

Chapter Seven emphasizes the ultimate aim of the study to address RO3, that is to develop an RDG framework tailored for RPOs in Malaysia. It presents the framework, providing insights into each of its components. The framework elaborates on both structural and functional dimensions and addresses the drivers for implementation. The structural elements include people and governance bodies. Within this section, the emphasis is placed on stakeholders ideally holding each RDG role and its corresponding governance structure. Meanwhile, the functional dimension encompasses data governance action, as well as processes (and technologies). Under processes, the discussion extends to cover RDG areas and decision domains.

Chapter Eight conducts a thorough discussion of the data, enabling the research study to validate the initially established objectives. The findings are analyzed to demonstrate the consistency and relevance of the phenomenon with what was previously reported in the literature. All research questions and objectives are effectively addressed, utilizing extracts and references drawn from the data presented in Chapters Four to Six. The chapter also formulates conclusions and provides recommendations for implementation and future research. Additionally, the chapter examines the research's contribution, limitations, and implications.

CHAPTER 2: LITERATURE REVIEW

"And do not pursue that of which you have no knowledge. Indeed, the hearing, the sight, and the heart—about all those [one] will be questioned."
(Surah Al-Isra, 17:36)

2.1 Introduction

The purpose of this literature review is to conduct an examination of previous studies concerning data governance practices. By synthesizing and analyzing relevant literature, this review aims to highlight key themes, concepts, and best practices, particularly within the context of RDG. In addition, the exploration aims to support broader objectives such as promoting research integrity, collaboration, and innovation within the academic community. Furthermore, this review establishes a groundwork for subsequent chapters by offering insights into the current state of RDG practices within RPOs. It also identifies areas for further research and enhancement, specifically for Malaysia RPOs. This includes examining the roles, stakeholders, structural positioning of each role, as well as the tasks, areas, and decision domains implemented and emphasized in the literature.

2.2 Definition of Research Data Governance and related Concepts

There is no globally agreed definition of data governance, as many people describe it differently. According to Abraham et al. (2019), data governance is defined as the exercise of authority and control over the management of research data to enhance its value while minimizing associated costs and risks. This aligns closely with DAMA International's (2009, p. 19) definition of data governance as "the exercise of authority and control over the management of data". Kouper et al. (2020) emphasize that RDG involves making critical decisions about data management, which are crucial for ensuring the overall credibility of research and engaging a diverse range of stakeholders. In the context of research, data governance also encompasses the necessary decisions for effectively managing and utilizing IT assets, including defining the responsible parties to ensure optimal use and management (Khatri & Brown, 2010). Additionally, Koltay

(2016) argues that data governance operates as a service characterized by standardized, repeatable processes. These processes are designed to promote transparency, cost reduction, and compliance with established rules, policies, standards, decision rights, and accountability. This systematic framework integrates effective data management, improves data quality, and maximizes the application of data to enhance organizational performance and competitiveness (Qi, 2021).

This study aligns with the concept of data governance as defined by Abraham et al. (2019). This concept can be further elaborated with the definition by Korhonen et al. (2013, p.11), which describes data governance as "an organizational approach to data management that formalizes a set of data policies and procedures to encompass the full life cycle of data, from acquisition to use and disposal". While at the point of RDM, Lefebvre et al. (2018, p. 5) assert that RDG involves "developing policies for data management planning and assigning tasks to researchers and data management services". Therefore, in the context of this study, RDG is defined as "an organizational framework that formalizes a set of data policies and processes for RDM throughout its entire life cycle and emphasizes the exercise of authority and control over research data-related decisions".

Meanwhile, RDG activities are the specific actions or tasks carried out to define, implement, and monitor effective RDG within an (research) organization, encompassing various governance areas and decision domains. Some tasks involve defining procedures for good RDM, others necessitate active implementation, while monitoring tasks ensure the proper execution of measures, identify areas for enhancement, and verify that the implementation process aligns seamlessly with the stipulated policies and laws. RDG is concerned with how individuals and processes interact with data and deal with it (Lefebvre et al., 2018). For instance, to protect privacy, data providers and users must

ensure that data is stored securely and provided without reidentification (Hendey et al., 2018).

In contrast, RDG is different from RDM as the latter is concerned with the day-to-day activities and practices of researchers in terms of archiving, organizing, and describing the data and related research materials (Borghi & Van Gulick, 2022; Tripathi et al., 2017a), as well as promoting long-term data preservation and usage (Adika & Kwanya, 2020). In the context of this study, as Willaert et al. (2019, p. 3) summarized, RDM is defined as "the process through which a researcher progresses from developing a management strategy for the research data towards an implemented, sustainable business plan for working with active data, as well as storing, archiving and cataloguing data upon completion of a research project".

In the realm of RDM, the role of a data practitioner holds significant importance. The Education Development Center (2022) defines the term data practitioner as someone who contributes to the data life cycle by collecting, transforming, analyzing data, and communicating results to inform decision-making. Additionally, two other definitions are considered: one from Stanford Professionals (2022), which encompasses anyone involved in data-related activities such as architecture, analysis, and consumption, and another from Mozilla (2020), which refers to someone who examines data, identifies trends and creates visual representations. For this study, a data practitioner is defined as any stakeholder who engages with research data and supports it throughout its life cycle, from generation and management to consumption. Within the scope of this study, data practitioners include policymakers, librarians, researchers, research officers, and IT managers. While data stakeholders are defined as individuals, groups, or organizations impacted by data governance practices (Micheli et al., 2020). These stakeholders play

crucial roles in supporting and utilizing research data, including engaging in RDG activities.

Meanwhile, the literature contains numerous definitions of research data. In this study, the term research data refers to any material collected, observed, or created for analysis in order to validate original research results (Drysdale, 2019; Kim, 2020; Shitole et al., 2019; Singh et al., 2018), regardless of the form of media on which they may be recorded (Vanderbilt University, 2021). The research data could encompass various forms such as textual, quantitative, qualitative, images, recordings, musical compositions, verbal communication, experimental readings, simulations, or codes (Tripathi et al., 2017b). A study by Patterton et al. (2018) found that both experienced and emerging researchers commonly used data formats such as spreadsheets, image files, and text. According to Denny et al. (2015), researchers view data as the lifeblood of their work, as it is closely tied to research outputs, publications, and future funding. The increasing prevalence of digital tools in academic settings contributes to the growing volume and diversity of research data being created, shared, and stored, particularly in online environments (Adika & Kwanya, 2020).

The terms ‘data’ and ‘information’ are often used interchangeably (DAMA International, 2017), though they represent different stages in the process of converting raw facts into meaningful insights. As Pomykalski (2020) explains, at some point in the data life cycle, data becomes information, which makes the privacy of both data and derived information essentially synonymous. However, in this study, they are treated as distinct. In line with the perspective of Tilly et al. (2017), data is objective and represents a phenomenon independent of an information system, while information is subjective and provides context to data through an information system, making it easier for users to understand. Data transforms into information when it is viewed in context or analyzed to

provide insights (Al Kez et al., 2022). In the context of this study, the singular term 'data' may be used interchangeably with the longer phrase 'research data' throughout this paper.

2.3 The Roles of Research Data Governance

As open science evolves, the roles of RDG become increasingly essential, ensuring that data is not only widely accessible but also effectively managed to maintain quality and integrity, ultimately benefiting the broader community. The following discussion highlights on the roles of RDG in relation to open science and RDM.

2.3.1 Enhancing Open Science and Open Data Initiatives

RDG is at the forefront of advancing open science and open data initiatives through policies, collaborative research projects, and transparent data sharing practices. Open science is not a new concept; the term Science 2.0 was previously used to describe the evolution of scientific practices (Szkuta & Osimo, 2016). According to the British Royal Society (2012), open science is "a realistic means of making data open to the wider public needs to ensure that the data that are most relevant to the public are accessible, intelligible, assessable and usable for the likely purposes of non-specialists". Meanwhile, the U.S. National Academy of Sciences (2018, p. 23) stated that open science "aims to ensure the free availability and usability of scholarly publications, the data that result from scholarly research, and the methodologies, including code or algorithms, that were used to generate those data". Due to the lack of a universally accepted definition of open science, Vicente-Saez & Martinez-Fuentes (2018, p. 428) has defined open science through their systematic literature review as "transparent and accessible knowledge that is shared and developed through collaborative networks". These definitions suggest that open science prioritizes the importance of offering open and free access to knowledge, including publications, data, and methodologies, to the broader community, encouraging transparency, accessibility, and reproducibility. To make open science's transformative

vision possible, the knowledge must be universally accessible, emphasizing communities frequently underserved by scientific products (Bahlai et al., 2019).

The areas under the umbrella of open science include, but not limited to:

a) Open data, which aims to ensure the sharing of research materials to others to facilitate replication studies and increase data reuse (Open Knowledge Foundation, 2023);
b) Open access to publications such as efforts to increase the systematic publication of scientific results (Chalmers et al., 2013), to preregister papers on preprint websites (Bourne et al., 2017), and to promote open-access publishing (Else, 2018);
c) Open source software, which encourages the sharing of computer code (Nosek et al., 2015);
d) Open peer review, which aims to increase the transparency of the peer review (Lee & Moher, 2017);
e) Open educational resources, including making study data publicly available (National Academies of Sciences, Engineering, and Medicine, 2018), sharing study materials (Nosek et al., 2015), and providing real-time updates on the studies progress (Foster & Deardorff, 2017);
f) Open collaboration that facilitates effective communication between experts and decision-makers to make effective use of scientific information (National Academies of Sciences, Engineering, and Medicine, 2018); and
g) Citizen science that involves the public in scientific research, in collaboration with professional scientists (Bowser et al., 2020).

Collectively, these initiatives can affect sociocultural and technological change through openness and connectivity and how research is designed, conducted, captured, and assessed (Vicente-Saez & Martinez-Fuentes, 2018).

Open data, as a fundamental area of open science, was firstly developed and defined as data that can be freely used, shared, and built on by anyone, anywhere, for any purpose. The word 'open' was introduced by the Open Knowledge Foundation in 2005 (James, 2013). From the word itself, it is understood that open data can be freely accessed, reused, remixed, and redistributed for academic research and teaching purposes and beyond. There are no restrictions on reuse or redistribution, and they are appropriately licensed as such (Murray-Rust, 2008). Examples of open research data practices include implementing open data policies, sharing research data, curating open data, accessing open data from data centers, and providing open data services (Tu & Shen, 2023).

In the realm of research data, open research data constitutes an integral aspect of contemporary research data ecosystems and practices (see Figure 2.1). Open research data is a central tenet of open science and can positively increase scientific research effectiveness, transparency, and reproducibility (Tu & Shen, 2023). It demands comprehensive documentation, proper curation, reproducibility, FAIR data, and publication impact, as highlighted by Fürholz & Jaekel (2021). Open research data can be categorized as raw or primary data, derived from primary data for subsequent analysis or interpretation, or derived from existing sources held by others (The Concordat Working Group, 2016).

Open research data should, in principle, adhere to the well-known FAIR data principles, which require that data be Findable, Accessible, Interoperable, and Reusable (Fürholz & Jaekel, 2021). The FAIR data principles were first established in 2016 (Wilkinson et al., 2016). FAIR data principles serve as a cornerstone for open science,

promoting data discoverability, accessibility, interoperability, reusability, and machine-actionability (Borghi & Van Gulick, 2022). Table 2.1 summarizes the actions associated with implementing FAIR data (Austin et al., 2021).

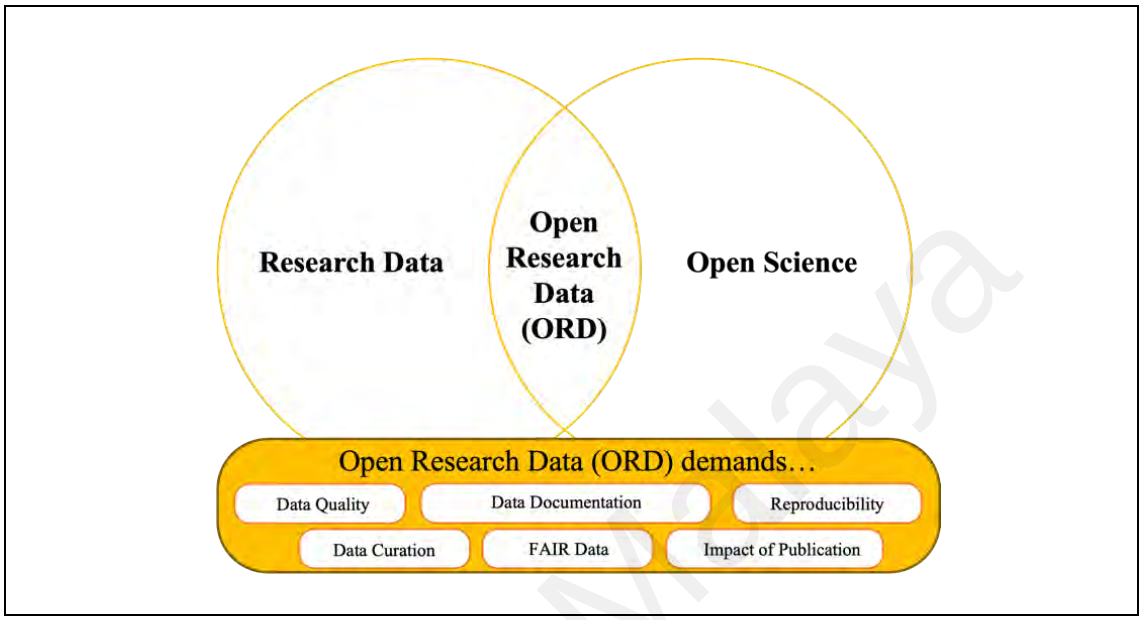


Figure 2.1: Open research data (Fürholz & Jaekel, 2021)

While data openness may not always be feasible or appropriate, adherence to FAIR principles is important. In cases involving sensitive data, only metadata describing the sensitive information is made available and accessible to the public. Sensitive data may be subject to special or limited access restrictions in exceptional circumstances, e.g., protecting individuals' identities or maintaining private and confidential information (Beyan et al., 2020; Kilbertus et al., 2018). Other than that, access to data must be strictly controlled in areas where the risk of a breach and its consequences is considered high (Faiz et al., 2020). The shift to big data also brings security and privacy challenges to the surface (Feki & Boughzala, 2016). Therefore, the FAIR principle is always accompanied by its tagline: data should be as open as possible and closed as necessary (European Commission, 2016).

Table 2.1: Implementation of FAIR data practices

Findable	Accessible	Interoperable	Re-usable
Data and metadata are assigned a globally unique and eternally persistent identifier	Data and metadata are retrievable by their identifier using a standardized communications protocol	Data and metadata are retrievable by their identifier using a standardized communications protocol	(Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation
Data are described with rich metadata	The protocol is open, free, and universally implementable	The protocol is open, free, and universally implementable	(Meta)data use vocabularies that follow FAIR principles
Data and metadata are registered or indexed in a searchable resource	The protocol allows for an authentication and authorization procedure, where necessary	The protocol allows for an authentication and authorization procedure, where necessary	(Meta)data include qualified references to other (meta)data
Metadata specify the data identifier	Metadata are accessible, even when the data are no longer available	Metadata are accessible, even when the data are no longer available	(Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation

Furthermore, it is crucial that the FAIR principles be supplemented with guidelines focusing on the entities involved in the data processes, as well as the motivation behind the collection and use of data. The CARE principles stand for Collective Benefit, Authority to Control, Responsibility, and Ethics, which were devised to support Indigenous peoples through moving consultation toward the core of relationships based on ethical values (Carroll et al., 2021). The CARE principles complement the FAIR principles by weaving in a focus on people and purpose, rather than just the data itself (Taitingfong et al., 2024). In other words, while FAIR principles focus on the technical aspects of data and information, CARE principles emphasize how data is collected, used, and shared in ways that promote the well-being of people (Carroll et al., 2021). Since its introduction, the CARE principles have been adopted and referenced across different fields and sectors globally (Taitingfong et al., 2024), inspiring many groups to rethink how to center people and the purpose of data use (Wylie et al., 2021).

Numerous researchers have emphasized the benefits of practicing and utilizing open (research) data. Sharif et al. (2018) concur that openly share research data can foster innovation. Apart from that, researchers would seek to openly share their research data in order to advance their careers, for example, by publishing their findings in journals of international renown (Zuiderwijk et al., 2020) or suitable data repository (Fürholz & Jaekel, 2021). A user service case study of national scientific data centers in China demonstrated that open research data was utilized to support government decision-making, educational activities, news reports, science popularization, and various other societal activities (Tu & Yang, 2020). Similarly, Thursby et al. (2018) concur giving access to research data exposes it to inspection, establishes the foundation for research verification and reproducibility. It may also help universities improve their research and education while expanding universities' impact beyond their institution (McKiernan, 2017). Overall, openly sharing research data forms the basis for research verification and reproducibility. It also opens a pathway to wider innovation and opportunities, leading to greater social and economic values. The actual value of open research data is recognized in its ability to be leveraged, shared, and combined with other data, enriching research and advancing analytical capacity to inform decision-making.

The adoption of open science principles promotes innovation and fosters transparency by efficiently utilizing research data (Downs, 2021; Huston et al., 2019; Leonelli et al., 2015). Therefore, to promote smooth research data sharing and collaboration, RDG mechanisms are imperative amidst the widespread adoption of open science initiatives (Borghi & Van Gulick, 2022; Fürholz & Jaekel, 2021; Napis et al., 2019; National Academies of Sciences, Engineering, and Medicine, 2018). Effective governance and management of research data ensures compliance with open research data standards, thereby enhancing the open research data initiative and advocating broadly for open

science. This approach encourages greater collaboration, innovation, access, use, and reuse of research findings.

2.3.2 Promoting Effective Research Data Management

The exponential growth of scientific data has compelled scientific communities to collect, preserve, and distribute research data globally. Buhomoli and Muneja (2022) propose that research data should be treated with care to ensure the validity of study findings, prevent duplication of effort, and save resources. The management of research data is a developing area that is gaining traction among universities, funding agencies, and academic publishers (Fürholz & Jaekel, 2021; Lefebvre et al., 2018).

Integrating RDM and open science practices can benefit RPOs. Researchers and institutions can increase the visibility of their research outputs and foster new collaborations or funding opportunities. Fürholz and Jaekel (2021) emphasize the significance of data curation, advocating for the publication of research data in discipline-specific repositories and linking journal articles to open research data via Digital Object Identifiers (DOIs) to enhance visibility. As new users explore and make discoveries, the previously collected data can yield additional societal benefits (Downs, 2021). Proper recognition of data sharing and submitting datasets to relevant repositories allow original data producers to derive new value from their initial data collection efforts (Krzton, 2018). Therefore, by adopting effective RDM practices, researchers can ensure their data is organized, documented, preserved, and readily available for sharing and reuse.

Directing data planning and organization are among the critical importance of RDM throughout the study cycle to ensure the quality, completeness, authenticity, and reliability of research data (Marlina et al., 2022). Wang et al. (2021) argue that RDM must facilitate the comprehensive documentation of research activities and provide an accurate representation of the effectiveness of researchers. The increasing acknowledgment of the

scientific crisis regarding reproducibility has amplified the need for research methodologies that are open and documented (Fürholz & Jaekel, 2021; Leonelli, 2018; Turkyilmaz-van der Velden et al., 2020).

In addition, effective management of the data lifecycle, as emphasized by Bertagnolli et al. (2017), is critical for researchers to examine and reanalyze data, verify findings, replicate studies, train future researchers, and generate new ideas and breakthroughs. Most researchers, as reported by Kabanda et al. (2023), primarily work with research and academic data. Ensuring robust RDM practices is essential for successful scientific research, dissemination, and reuse, safeguarding the quality and integrity of research outputs while fostering collaboration and innovation (Borghi & Van Gulick, 2022).

Therefore, the importance of RDM has been highlighted in the context of open science, given its crucial role in effectively preserving, safeguarding, and sharing research data (Lefebvre & Spruit, 2021), as well as fostering the development of open infrastructure (Higman et al., 2019). Shmagun et al. (2023) emphasize the significance of RDM as an exemplar of open science practices. Ultimately, increasing awareness and prioritization of RDM is crucial, as suggested by Cruz et al. (2019), who propose integrating it into hiring and performance evaluation criteria. This approach not only underscores the importance of data management but also incentivizes researchers to allocate additional time and effort to this critical aspect of their work.

To have proper RDM, effective RDG is essential, as RDG has gained significance in modern scientific practices due to the growing demands for transparency, reproducibility, and ethical standards. Establishing data policies and governance frameworks encompassing comprehensive guidelines has emerged as a crucial priority for RDM in open science (Schöpfel & Azeroual, 2021). It is imperative to note that distinct data governance is essential for research data, differentiating them from institutional or

business data (Gupta & Cannon, 2020a). Such measures can enhance the effectiveness, transparency, and reproducibility of scientific research (Fürholz & Jaekel, 2021), resulting in better decision making. This lack of implementation may have negative consequences for the research data life cycle, as insufficient guidance may result in issues with data collection, handling, storage, and re-use (Buhomoli & Muneja, 2022).

Thus, efficient RDG facilitates the validation and verification of research findings through the promotion of data transparency. Furthermore, the integration of sound RDG principles enhances the efficiency of RDM procedures, increases the standard of research data, and enables the discovery and reuse of data (Beyan et al., 2020; Kilbertus et al., 2018; Ng'eno & Mutula, 2018). Efficient RDG frameworks enable researchers to focus more on analysis and interpretation, accelerating the pace of scientific discovery.

2.4 The Importance of Research Data Governance in Research Performing Organizations

The growing volume and complexity of research data and the need for effective data management and sharing to promote scientific progress have underscored the increasing importance of RDG within RPOs. Proper implementation of data management is impossible without good data governance (Brous et al., 2016). Effective RDG provides the guiding principles and policies that inform RDM practices (Jim & Chang, 2018; Lefebvre et al., 2018; Omar & Almaghthawi, 2020). It sets the boundaries and expectations for responsible data handling, ensuring compliance with ethical guidelines, legal requirements, and institutional policies. A robust approach to policy development in RPOs requires a balance between top-down and bottom-up initiatives. A study by Napis et al. (2019) found that 70.4 percent of respondents agreed with the university's service on research data policies. Cruz et al. (2019) argue that top-down policy measures should be combined with bottom-up community-driven efforts to ensure widespread

acceptance of regulations and policies. Additionally, they emphasize that policies must align with community practices to remain relevant.

Data governance can increase data quality by increasing the degree of numerous quality factors (Koltay, 2020). For instance, proper data governance permits the management of risks that may develop due to non-compliance with information policies or a lack of control by establishing risk-mitigating policies (Abraham et al., 2019). Comprehensive coverage of RDG activities is crucial for ensuring better data quality and broader access, guarding against data loss from intentional theft, accidental leakage, and improper disclosure (Solomonides, 2019). Moreover, it addresses concerns such as a lack of trust among data consumers (Abduldayan et al., 2021) and potential data misuse (Tiffin et al., 2019). An RPO that invests in data governance may also mitigate a data breach or regulatory oversight (DalleMule & Davenport, 2017).

RDG is essential for ensuring the protection and effective use of data assets within RPOs. As a result of the enormous volume of data stored within RPOs, RDG should be implemented to safeguard data assets and lead data activities toward organizational goals (Jim & Chang, 2018; Nielsen, 2017). RPOs can move faster, make better decisions, and deliver better insights into consumer behaviour with solid data governance practices (Omar & Almaghthawi, 2020). A research data policy can assist researchers in safely exchanging research data. Similarly, they permit data interchange with other scholars if the findings are deemed valuable to society. According to Napis et al. (2019), RDG will facilitate collaboration and engagement among researchers from various disciplines within a trusted ecosystem, resulting in increased research outputs. As seen in their study findings, respondents are willing to cooperate and contribute data to other researchers if a proper policy document exists (Napis et al., 2019).

The implementation of efficient RDG including ethical considerations, which encompass the protection of research participant confidentiality and the prevention of data misuse (Al-Ruithe et al., 2016; Jamiu et al., 2020; Liu et al., 2020). Data governance frameworks that emphasize ethical principles can cultivate confidence among stakeholders and advocate for the ethical management of data.

In conclusion, the implementation of RDG within RPOs is crucial for safeguarding data assets and enhancing organizational performance. Disregarding RDG in RPOs can negatively impact data handling, preservation, ownership, access, sharing, and security, ultimately diminishing the return on investment for research resources. Arguably, organizations lacking a governance strategy will confront substantial risks, but those operating inside a governance framework perform significantly better (Omar & Almaghthawi, 2020).

2.5 Stakeholders Accountability in Research Data Governance

The concept of data governance is becoming more widely accepted. A complex research product or network with a wide range of stakeholders' access methods and legal settings poses questions regarding who should manage and control such complexity, as well as who owns and cares for data throughout its lifecycle (Kouper et al., 2020). Their study reveals that most respondents recognize RDG as a collective responsibility within the data ecosystem. Effective data governance is crucial for ensuring clarity and accountability in research practices.

Individual researchers should be primarily responsible for data decisions (Kouper et al., 2020). This underscores the importance of clearly delineating researchers' roles and responsibilities within RDG policies, as they are primarily engaged in various stages of the research lifecycle, including the production and preservation of research data. However, Lefebvre et al. (2018) highlight the expansive nature of contemporary

responsibilities, concepts, and interpretations within professional data management. They assert that the existing framework, exemplified by the DAMA-DMBOK's 32 defined roles, fails to adequately incorporate the distinctive roles pertinent to RDM. Despite RDM's restructuring of traditional academic roles such as researchers, principal investigators, and deans, it neglects to formalize essential positions like data managers or data stewards. Even though these roles actively collaborate with researchers and oversee research data across its entire lifecycle, they remain unregulated within the current schema. Consequently, Lefebvre et al. (2018) argue that researchers often find themselves burdened with data management tasks for which they lack sufficient expertise.

The implementation of RDM services demands comprehensive and systematic approaches from various research support service providers (Chiware & Mathe, 2015; Fürholz & Jaekel, 2021), especially given the complexities of today's research landscape. Cruz et al. (2019) affirm this perspective, noting that data governance policies often clarify the responsibilities of academic and support staff, covering research procedures and support services such as information technology, legal, and ethical committees. Researchers, librarians, administrators, and information technology professionals all play pivotal roles in ensuring that research data and associated information remain accessible, visible, understandable, and usable over the long term (Buhomoli & Muneja, 2022; Fürholz & Jaekel, 2021).

Besides that, in many universities, the library and research office play essential roles as stakeholders supporting research activities, although they may not have a longstanding tradition of closely collaborating (Cox et al., 2017). Librarians and other technical support staff play a vital role in acquiring research data (Singh et al., 2018) and giving research help to the researchers (Gunjal & Gaitanou, 2017). In addition to offering library services, library staff would help researchers by providing training on cutting-edge research

tools and techniques and specialized resources pertinent to their field of study (Gunjal & Gaitanou, 2017; Hamad et al., 2019; Ismail et al., 2022; Zain et al., 2023).

Meanwhile, Omar & Almaghthawi (2020) recommend that universities establish an effective functional team for data governance tasks, conduct an internal audit of data governance, monitor regulatory compliance procedures, define the priorities of data governance activities, provide frequent data governance training to employees and faculty members, establish enforcement and follow-up standards, and conduct frequent data governance audits as depicted in Figure 2.2. Thus, the community within an institution must recognize the presence of regulations and policies. It is essential for everyone to collaborate across all levels of governance to carry out their tasks effectively. Furthermore, alignment with best practices is imperative, along with establishing a competent governance team to define and oversee activities related to data governance. However, the study is limited to Saudi universities and does not specifically identify the data governance team or activities. Further research is needed to provide higher education institutions with practical and reliable solutions for the effective management of big data and the assurance of data quality, ultimately improving operational efficiency.

Similarly, according to Jim and Chang (2018), most data governance activities in universities are supported by a specified group of members in the form of a committee, council, or working group. The second and third largest groups of workers supporting and supervising data governance processes are those in the IT and institutional research departments, respectively (Jim & Chang, 2018). While some organizations maintain specialized units for data management and analytics, others adopt more collaborative approaches through shared offices or units dedicated to supporting data governance (Jim & Chang, 2018).

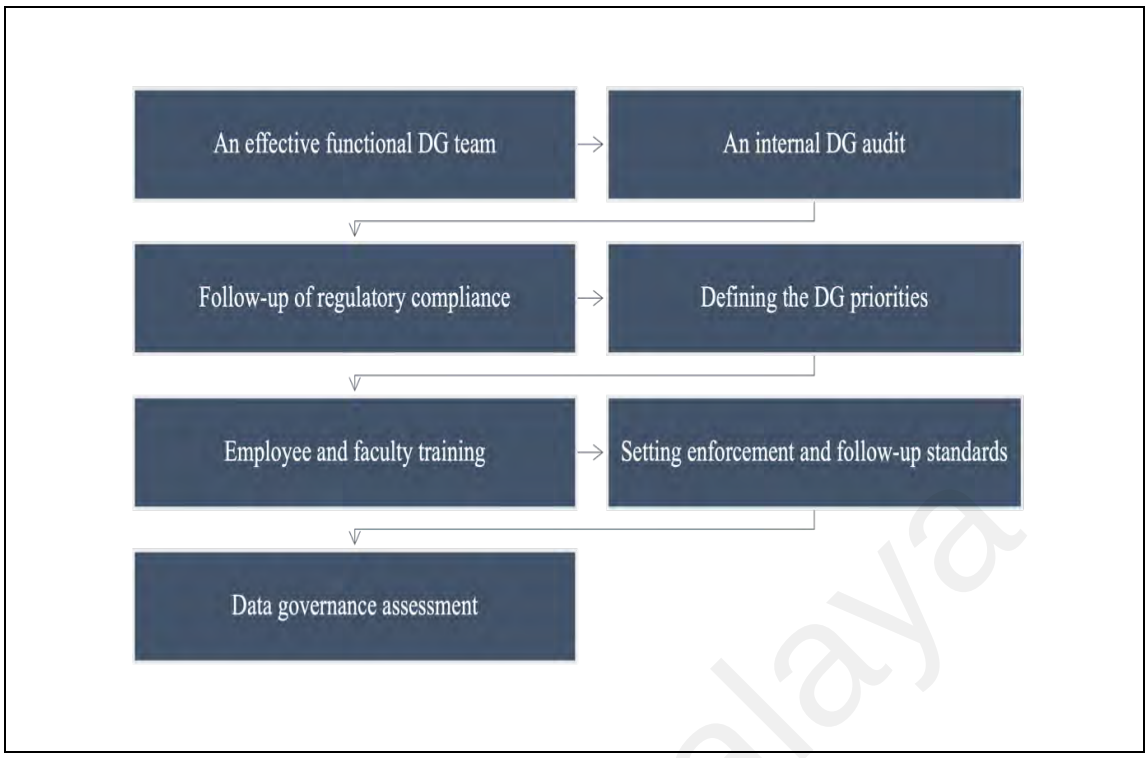


Figure 2.2: A proposed model for good data governance practices (Omar & Almaghthawi, 2020)

Furthermore, comprehensive guidance from journal publishers regarding data deposition play a pivotal role in shaping RDM practices. Journal publishers must advise and provide clear instructions on data deposition for authors (Aleixandre-Benavent et al., 2019), including specific data types shared through endorsed repositories significantly influences research data policies (Rousi & Laakso, 2020). These factors contribute to the development of RDG and the unprecedented growth of research data repositories worldwide. Furthermore, mandates from funders regarding research data sharing serve as facilitators for promoting data sharing practices (Anger et al., 2024; Cox et al., 2019b; Gaba et al., 2020).

In conclusion, ongoing engagement and coordination among researchers, institutions, and communities are vital for effective RDG, alongside the utilization of frameworks, guidelines, and policies that support responsible data management and usage. Collaboration among various stakeholders is essential to ensuring the integrity of research data and enabling future research. However, who should be responsible for research data

in Malaysia RPOs is still questionable. Therefore, the current study aims to identify the stakeholders involved in RDG and the specific key task areas unique to RPOs in Malaysia.

2.6 Understanding the Barriers for Research Data Governance

While demand for RDG has increased in recent years to improve the quality of research data, there is still a lack of a defined regulatory framework to guide the process (Chigwada et al., 2017; Liu et al., 2020), that cover RDM drivers, including preservation, storage, security, quality, compliance, sharing, and jurisdiction (Ng'eno & Mutula, 2018). In addition, embracing data governance has proven difficult for some organizations (Benfeldt et al., 2020; Chigwada et al., 2017; Manik et al., 2022).

Several studies have highlighted persistent barriers, including limited awareness among researchers about the significance of data governance practices (Singeh et al., 2013b). Effective communication and training are essential for raising awareness and understanding on RDG within RPO communities (Palsdottir, 2021). Researchers must first understand the critical role of RDM in sustaining research through disseminated regulations and policies (Cahyaningtyas & Priyanto, 2021). Studies show that universities in the UK emphasize offering RDM training to all members (Liu et al., 2020), while 77 percent of Malaysian universities agree on the necessity of proper training and information services for successful RDM implementation (Napis et al., 2019). Without adequate education and training, researchers may overlook data management principles, leading to inconsistencies and inefficiencies in data handling (Buhomoli & Muneja, 2022). According to Wiley and Kerby's (2018) study, there are indications that researchers need assistance with data management. This applies to both experienced researchers and those who are just starting their careers.

Other than that, the lack of standardized approaches to data management across different fields and disciplines poses a significant challenge. Implementing effective data governance in research organizations faces barriers such as the need for a deeper understanding of professional data management complexity and the failure to adequately connect research data support with diverse projects (Lefebvre et al., 2018). To address these challenges, various frameworks and guidelines, including the widely accepted FAIR principles, have been developed to prioritize the findability, accessibility, interoperability, and reusability of data (European Commission, 2016; Wilkinson et al., 2016).

RPOs share similar principles in managing their research, with data management plans (DMPs) being a key requirement to guide researchers in managing their data throughout the research process (Burgi et al., 2017; Nanyang Technological University, 2019). The National Science Foundation, for example, mandates that all scientists receiving federal grants develop a DMP as part of their research approach (Luesebrink et al., 2014). The DMP typically includes an overview of the data created and managed, as well as a detailed description of the methods used for data collection, processing, and analysis, and compliance with the funding agency, copyright, and ethical requirements (Chigwada et al., 2017; Marlina et al., 2022). Liu et al. (2020) found that 77 percent of UK RPOs require DMPs for all research, while all new proposals must include strategies for data acquisition, management, integrity, confidentiality, retention, sharing, and publication. In Australia, 48 percent of universities require DMPs for all research, and 44 percent require them only for funded research.

However, despite the recognized importance of RDM, the adoption of DMP remains inconsistent across different regions and disciplines. The findings from a study by Buhomoli and Muneja (2022) suggested that only 34 percent of researchers at selected RPOs in Tanzania utilize DMPs in their research activities, while the majority (66 percent) do not. According to the same study, most researchers refrain from using DMPs because they consider them irrelevant to their research topics or unrelated to their work. From another perspective, according to a study by Palsdottir (2021), researchers in sciences demonstrated a lower likelihood of being unaware of a DMP and standard file naming system compared to researchers in social science and humanities.

Without proper DMPs for strategically handling research data, data management efforts can become disorganized, and preserving existing research data remains a challenge, as data often resides with individual researchers (Buhomoli & Muneja, 2022; Chigwada et al., 2017). In their study, Buhomoli and Muneja (2022) show no standardization of research data preservation procedures among researchers. Most of them keep their research data after the research, while others discard it after the research, and the rest give it to project supervisors or share it with friends.

Having different timeframes for archiving research data between researchers may also indicate a lack of guidelines and coordination over the actual timeframe for archiving data (Buhomoli & Muneja, 2022; Chigwada et al., 2017). In some cases, institutions do not even have a policy for archiving research data, which implies that they are deleted after being analyzed (Chigwada et al., 2017). In other contexts, when papers are published in scholarly journals, only a portion of the data is made public (Chan et al., 2014). Plus, as a result of ICT's presence, data is stored in different media, making retrieval difficult at the time of need (Jamiu et al., 2020).

Since researchers retain complete control over research data, the study's outcomes (Chigwada et al., 2017) indicate that the researchers would determine access to research data. They may allow anybody to access the research data or restrict it based on the data privacy classification. Surprisingly, the study by Cahyaningtyas & Priyanto (2021) reveals that some researchers refuse to disclose data because they either do not want to contribute, do not have legitimate data, or fear embarrassment if others learn their data are invalid. Additionally, some researchers desired to disclose their data but had misplaced or forgotten where it was stored (Cahyaningtyas & Priyanto, 2021).

Ensuring data security and privacy are also crucial, with the General Data Protection Regulation (GDPR) serving as key legislation governing data protection in the European Union (Peukert et al., 2022; Starkbaum & Felt, 2019). Researchers face hurdles in complying with data privacy standards like GDPR, which are constantly changing (Labadie & Legner, 2023). Adhering to various and sometimes intersecting regulations complicates data sharing and imposes additional administrative burdens. Additionally, ensuring robust data security measures has become increasingly critical in higher education, as evidenced by the rising frequency of data breaches (Chapman, 2019) and the vulnerability of sensitive research data to malicious threats (Nasir et al., 2023). Other studies have concluded that a lack of security awareness is directly related to how faculty value the information system assets of their universities (Nyblom et al., 2020). Verizon (2023) noted that the 2023 Data Breach Investigations Report (DBIR) examined 16,312 incidents, of which 5,199 were confirmed data breaches.

Besides that, while open (research) data is widely recognized, actual data sharing is somewhat limited (Alsheikh-Ali et al., 2011). Vasilevsky et al. (2017) found that only a minority of journals (12 percent) require data sharing as a pre-condition for publication. Additionally, researchers are frequently unprepared or unwilling to communicate their

findings to the public or decision-makers as they believe that simply publishing their findings in peer-reviewed journals will result in data use (Joo & Peters, 2020; Tripathi et al., 2017a; Vilar & Zabukovec, 2019). Among others barriers still surround data sharing practices include (Beno et al., 2017; Fecher et al., 2015; Ma & Lam, 2019; Tenopir et al., 2020; Tu & Shen, 2023):

a) Cost of sharing, e.g., time and effort to clean up data, create documentation and metadata, and check the integrity and consistency of data;
b) No or little rewarding of professional credit, e.g., data as a publication, social recognition, promotion, etc.;
c) Concerns about the misuse of data, e.g., the risk of data being misinterpreted, combined inappropriately, or incorrectly represented;
d) Ethical issues, e.g., concerns about confidentiality as certain data may contain sensitive personal information; and
e) Researchers' reluctance is due to their keen sense of ownership of their data.

Implementing RDG systems also faces numerous challenges that extend beyond organizational culture and technical barriers. Organizational cultures that prioritize individual autonomy over collective data stewardship may impede efforts to implement centralized data governance structures, as resistance to change and concerns about data ownership can hinder collaboration and data sharing initiatives (Bietz et al., 2016). Besides that, insufficient resources hinder the development of data management plans, infrastructure, and training programs necessary for effective governance. According to Lacagnina et al. (2022), resource limitations, including funding, time, and expertise, pose significant barriers to implementing data governance practices. Insufficient infrastructure and the absence of established protocols impede data integration and sharing across platforms and fields (Machimbidza et al., 2022). Additionally, technical barriers such as

interoperability challenges, data format compatibility, and data security risks can also hinder efficient RDG (Alvarez-Romero et al., 2023; Filgueiras & Raymond, 2023). Wang et al. (2021) highlight the difficulties caused by the rapid evolution of information technologies, including big data, the Internet of Things, mobile Internet, and cloud computing, which have resulted in ongoing changes to the scientific application of research outputs and dissemination techniques.

In conclusion, addressing challenges related to awareness, resources, regulations, and technology is crucial. Data inaccessibility due to incompleteness, inaccuracies, data silos, and unpublished studies may limit the scope of knowledge expansion and innovation activities. Thus, representing a complete lack of return on investment for research resources and study participants (Chan et al., 2014). Leveraging momentum from reproducibility concerns, efficiency gains, ethical obligations, funder requirements, and collaboration can help. This approach can create a more robust and transparent research data ecosystem.

Taking into account the aforementioned information, RPOs may struggle to manage their research data effectively without the existence of RDG (Kouper et al., 2020; Omar & Almaghthawi, 2020). Therefore, it is worthwhile and necessary to promote the development, preparation, and implementation of RDM (Gupta et al., 2021; Lau et al., 2021; Nie et al., 2021), including RDG (Kouper et al., 2020; Lefebvre et al., 2018; Makhlouf-Shabou, 2017; Omar & Almaghthawi, 2020) and the process should involve a variety of stakeholders (Abraham et al., 2019; Borkakoti, 2021; Cox et al., 2017; Kariotis et al., 2020; Parmiggiani & Grisot, 2020; Rosenbaum, 2020), including library staff (Ismail et al., 2022), research scholars, information technology experts, and administration (Piracha & Ameen, 2019). The characteristics of each data type must be understood from the start to protect the data assets strategically. Additionally, RPOs must

understand where the data are located, how they are used, and where and when they are integrated to determine a data value (Alhassan et al., 2019b; Gupta & Cannon, 2020b).

2.7 The Significance of Research Data Governance for Malaysia Research Performing Organizations

The Malaysian government aims to turn the nation into a hub for knowledge and innovation through research and development (R&D), as outlined in the National Higher Education Plan: Beyond 2020 and the Malaysia Education Blueprint 2015–2025 (Ministry of Education Malaysia, 2015). The government has restructured higher education to enhance research outputs for nation-building and social welfare. For instance, in 2018, the Ministry of Higher Education allocated RM13.89 billion to universities, which was a 13.5 percent rise from RM12.28 billion in 2017 (Jusoh, 2018). Malaysian universities utilized RM5.58 billion in federal research grants to generate RM7.17 billion between 2007 and 2015, resulting in a 28.5 percent return on investment (Chik et al., 2018). The global ranking of Malaysian institutions has improved as a result of their achievements in research outputs (QS World University Rankings, 2022). Additionally, Malaysia showed its dedication to an open research data policy by introducing the Malaysia Open Science Platform (MOSP) in 2020 (Ibrahim & Wei, 2023).

Despite the increasing research production and data availability, Malaysian universities have only partially embraced open science (Ahmed & Othman, 2021). For instance, the study by Singeh et al. (2013a) conducted a quantitative study that utilized a web-based survey approach in five Malaysian research-intensive institutions to determine the readiness of the authors to self-archive in open access repositories. The study employed the Unified Theory of Acceptance and Use of Technology (UTAUT) paradigm to evaluate the authors' behavioural intention to self-archive in institutional repositories.

The results indicate that despite academics' recognition of the effectiveness of institutional repositories for disseminating knowledge, most of them have yet to fully embrace self-archiving.

In a similar study, Singeh et al. (2013b) examined academics' awareness of self-archiving, perceptions of self-archiving scientific information, perceptions of obligations to self-archive, and possible barriers to contributing to institutional repositories at five Malaysian research-intensive universities. The researchers utilized a web-based survey to select 72 academics. The study found that most academics needed more understanding of institutional repositories and were unaware of self-archiving options. Researchers believe that sharing data can be time-consuming and may lead to plagiarism, which could be a barrier to contributing to institutional repositories.

In more recent studies like Mohamad Hashim (2019), the author believes implementing model guidelines to address legal barriers to open access to publicly funded research data in Malaysia is necessary. The author compared civil society's conceptions, policies, processes, government organizations, research funding agencies, and research institutions in Australia, Canada, the EU, the United Kingdom, and the United States on open access to research data. The model guideline is appropriate for implementation by public research funding bodies and RPOs in Malaysia. Hence, the model rules can become a standard in giving open access to publicly funded research data in Malaysia. It is also ideal for policymakers to refer mainly to intellectual property protection, ownership, copyright and licensing, author's moral right, data confidentiality, data privacy, data security, and legal responsibility to assure data quality.

While, Napis and colleagues (2019) explored the development of a data repository policy framework at a Malaysian public institution, emphasizing raising awareness among researchers of the importance of sharing data with others. A total of 164 respondents completed questionnaires focusing on the benefits of data sharing, the lack of a requirement for data sharing, and the types of data that should be included in managing data repository systems (Napis et al., 2019). The findings indicate that if an appropriate policy document is in place, respondents are keen to participate and share data with other researchers (Napis et al., 2019). Aside from that, this paper discusses recommended principles for data repository policies, as well as strategies for encouraging data sharing between researchers (Napis et al., 2019).

A similar study by Hodonu-Wusu et al. (2020) explored the awareness of Malaysian researchers regarding open research data sharing and usage, revealing that, although they are aware of open data, they are not currently integrating it into their research practices. The authors suggested that researchers should be rewarded and exposed to tools and guidelines that allow data sharing and reuse to better understand open research data and be aware of the advantages associated with these open data initiatives (Hodonu-Wusu et al., 2020).

Another study (Olesen et al., 2018) found that misbehaviour such as, manipulating research data, misrepresentation of research outcomes, plagiarism, authorship disputes, breaching of research protocols and unethical research management, were witnessed by participants among junior and senior researchers. Although Malaysian's institution of higher education have taken steps to monitor research misconduct, it still occurs in the research community (Olesen et al., 2018).

Besides that, the study by Zain et al. (2023) on Malaysian information science researchers revealed a reliance on Google Drive for data storage, collaboration, and backup, posing risks like data loss. Moreover, the use of WhatsApp as a means of communication and the lack of standardization in data management procedures, such as file naming conventions and a hybrid approach to filing documents, may result in ineffective data retrieval, security concerns, and challenges in data sharing and collaboration.

Meanwhile, in the context of Malaysian academic librarianship, involvement of librarians in RDM services is important to ensure research data is findable, accessible, interoperable, and reproducible (Amanullah & Abrizah, 2023). However, findings from the study conducted by Amanullah & Abrizah (2023) underscore the absence of clear information and established RDM plans or policies in Malaysian academic libraries. Consequently, libraries may misinterpret RDM services, leading to basic extensions of traditional services. The study advocates for the adoption of formal RDM plans by universities or libraries to address these gaps effectively, preventing libraries from merely providing basic RDM services.

Similarly, another study by Ismail et al. (2022) aims to investigate researchers' behaviours and practices in RDM and proposes ways for the library to integrate RDM into the research services provided. Unfortunately, the findings show that most researchers managed their research data based on individual perspectives and practices, without adhering to proper guidelines and standards. They used personal solutions for research data storage and preservation, and they also refrained from sharing their research data due to privacy and confidentiality concerns.

In summary, while the Malaysian government strongly supports open research data, researchers show limited awareness and willingness to share their research data due to various perceived reasons. Furthermore, the absence of established data policies fails to incentivize the research community to participate in research data sharing, leading to potential instances of research misconduct. Therefore, despite efforts by higher education institutions to enhance research outputs for nation-building and social welfare and to monitor research misconduct, there is a need for RDG within Malaysia RPOs to mitigate data-related risks. The establishment of clear information and well-defined RDM policies is crucial to ensure that researchers adhere to proper guidelines, thus maximizing the potential benefits associated with research data.

2.8 Theoretical Frameworks and Core Concepts

This section provides a discussion of previous studies, encompassing both theoretical and empirical literature. It aims to identify theoretical frameworks to underpin the study, key concepts, and gaps within the existing literature, focusing on studies related to data governance, key roles, and specific areas of interest.

2.8.1 Studies on Data Governance

Existing literature address various perspectives of data governance, such as the conceptual framework on data governance (Abraham et al., 2019), data governance activities (Alhassan et al., 2018), factors and activities of data governance (Ahmadi et al., 2022), data principles (Brous et al., 2016), roles and accountabilities (Korhonen et al., 2013), scientific data stewardship (Peng et al., 2018), data governance for cloud computing (Al-Ruithe et al., 2016), platform ecosystems (Lee et al., 2017), and third-generation platforms (Yebenes & Zorrilla, 2019).

Abraham et al. (2019) in their study state that data discrepancies must be identified and resolved prior to drawing conclusions. However, a holistic approach is lacking due to publications focusing on specific choice domains or including smaller reviews. Therefore, their study tries to formalize data governance concepts, synthesize existing literature, and propose a research agenda. A comprehensive literature review was undertaken on 145 research publications and practitioner articles between 2001 and 2019. They define and decompose the major building pieces of data governance along six dimensions, namely governance mechanisms, organizational scope, domain scope, data scope, antecedents, and consequences. The study has certain limitations, which are as follows: firstly, it places less emphasis on the broader concept of data management, focusing more on specific aspects; secondly, the terms 'information governance' and 'data governance' are used interchangeably, potentially leading to confusion or lack of clarity in their respective definitions and scopes. A conceptual framework for data governance is provided in Figure 2.3.

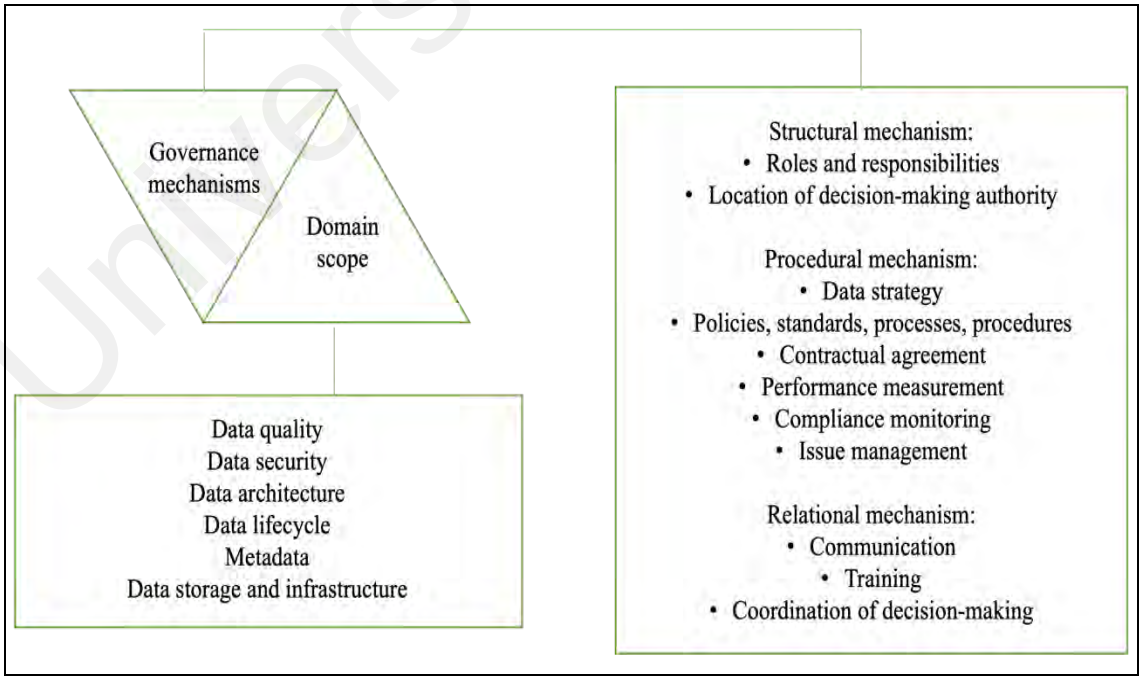


Figure 2.3: The Conceptual Framework for Data Governance by Abraham et al. (2019)

Using Abraham et al.'s (2019) theoretical framework to guide both data collection and analysis and employing thematic analysis by Braun and Clark for data analysis, Al-Wahshi et al. (2022) explored data quality issues in the Omani banking sector and assessed how various data governance mechanisms enhance data quality. The study employed a qualitative case study methodology, gathering data through semi-structured interviews and document reviews. The findings reveal that governance mechanisms like performance measurement, compliance monitoring, and training significantly improve data quality in Omani banks. The study provides empirical evidence on the effectiveness of these mechanisms and their application in the Omani banking context. It offers practical insights for practitioners aiming to implement systematic data quality management approaches. However, the findings are specific to the Omani banking sector and may not generalize to other industries or regions without considering their unique contexts. Further research should validate these findings across diverse organizational settings, explore factors influencing the adoption of data governance mechanisms, and investigate how enhanced data quality impacts organization's performance.

binti Azizatun Nafi'ah (2021) conducted a study using the concept of data governance introduced by Abraham et al.'s (2019); focusing on structural, procedural, and relational mechanisms. This study focused on how the national COVID-19 task force in Indonesia manages data to ensure it is valid and real-time. Researchers used qualitative analysis, collecting data through interviews with three representatives from the Ministry of Communication and Informatics. The study found that data management relied on structural, procedural, and relational mechanisms. The structural mechanism is well-established from national to regional levels. The procedural mechanism has improved and now focuses on data integration. The relational mechanism shows that coordination and communication among task force members are effective and quick.

Similarly, Rachmat et al. (2023) conducted a study based on Abraham et al.'s (2019) data governance framework. The study explores the challenges in creating effective institutional structures, implementing detailed procedures from central to regional levels, and maintaining coordinated relationships. Miscommunication with data input officers and the need for fostering empathy in them were identified as key issues. Using a qualitative approach with assessments, questionnaires, and interviews, the study evaluates data governance in the One Data Indonesia portal at Sukabumi City's Communication and Information Agency to ensure quality data from Regional Apparatus Organizations. The findings reveal gaps in data governance maturity, highlighting risks related to people, processes, and technology. The study offers recommendations for improving data governance and master data management at the organization.

While, another authors (Alhassan et al., 2018) suggest that since data has become a key organizational asset, data governance is necessary and critical. The absence of a data governance program may cause failure in running an organization. They conducted a study to explore the current literature on data governance in scientific and practice-oriented publications and provide a comparative analysis of the activities reported for data governance. According to Alhassan et al. (2018, p. 304), data governance activities are "the conditions or things that need to be performed in order to be considered as doing data governance". The authors conducted a comprehensive literature assessment on 61 articles that specifically discuss data governance initiatives. Alhassan et al. (2018) introduce a model of data governance activities, as depicted in Figure 2.4. The model comprises three data governance activity constructs: action, governance area, and decision domain. The decision domains are based on the framework proposed by Khatri and Brown (2010). They claim that this model provides a thorough understanding of the required data governance activities, as the activities are examined from academic and practitioner perspectives. Their analysis reveals that practice-oriented publications report a greater

volume of data governance activities associated with the "implement" and "monitor" actions of the governance areas across decision domains than scientific publications, despite the authors' observation that scientific publications focus more on defining activities. The study has some limitations, including the absence of a complete description of each of the 120 identified activities, and the model of data governance activities developed has yet to be empirically tested and validated.

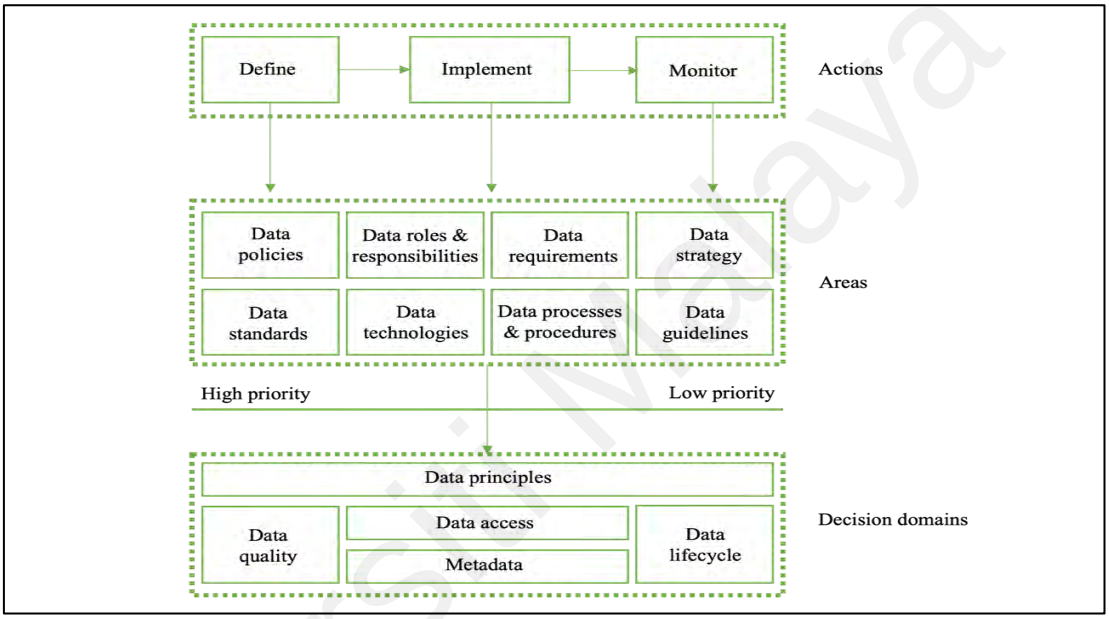


Figure 2.4: Data Governance Activities Model by Alhassan et al. (2018)

In a later study, Alhassan et al. (2019a) compared findings from this research with the data governance activities model, highlighting significant 'implement' actions identified through interview coding in the case study. This underscores the importance of focusing on implementation activities for each identified critical success factor (CSF) to achieve successful data governance across different governance areas. This study sets out to uncover what makes data governance successful and suggests how it can be effectively implemented in a major telecommunications provider in Saudi Arabia. Through semi structured interviews focused on CSFs, the research gathered insights and analyzed them using various coding techniques. It identified nine key CSFs, ranking them in terms of importance and providing actionable recommendations for each. The study contributes

by revealing both commonalities and differences in data governance approaches and emphasizes the importance of business-driven data integration strategies as crucial for success. Despite its single-case focus, the findings provide practical insights and encourage further research into effective data governance strategies for similar organizations.

Data management is vital for improving business operations and gaining a competitive edge by transforming organizations into data-driven entities where decisions are based on data insights. However, organizations face various challenges in establishing effective data management systems. Therefore, Kaewkamol (2022) aimed to propose a framework and guideline for implementing data governance in higher educational organizations. The framework was developed by adapting the data governance activities model by Alhassan et al. (2018) and key elements in data governance from Mahanti (2021), and by conducting semi-structured interviews with five staff members representing data stewards. These interviews explored the current state and challenges of data governance. The resulting framework is designed to kickstart a data governance scheme, promoting the shift to data-driven organizations. It provides a practical guideline for higher education institutions to manage data effectively. While the study offers valuable insights, it is based on a single case study with five participants, which might not cover all potential challenges and solutions.

While high-quality data is critical for corporate value creation, data quality issues are also frequently addressed in businesses, necessitating appropriate data governance. Inadequate access to trustworthy information increases risks and expenses and bad managerial decisions. Although data quality has been linked to data governance, the research has not extensively examined the accountability aspect of data governance. Therefore, using design science methodology, Korhonen et al. (2013) conducted a study

to analyze data governance roles and responsibilities from an organizational design perspective. The study aimed to identify typical data management roles and explore how the Agile Governance Model (AGM) could serve as a foundation for designing a suitable governance structure for the organizational arrangement of data governance accountabilities. The data governance team are divided into five tiers: strategic steering, strategic implementation, tactical, operational, and day-to-day, as seen in Figure 2.5. The limitations are that the analysis is based solely on secondary sources, and additional features can be identified when applying the Agile Governance Model for governance design.

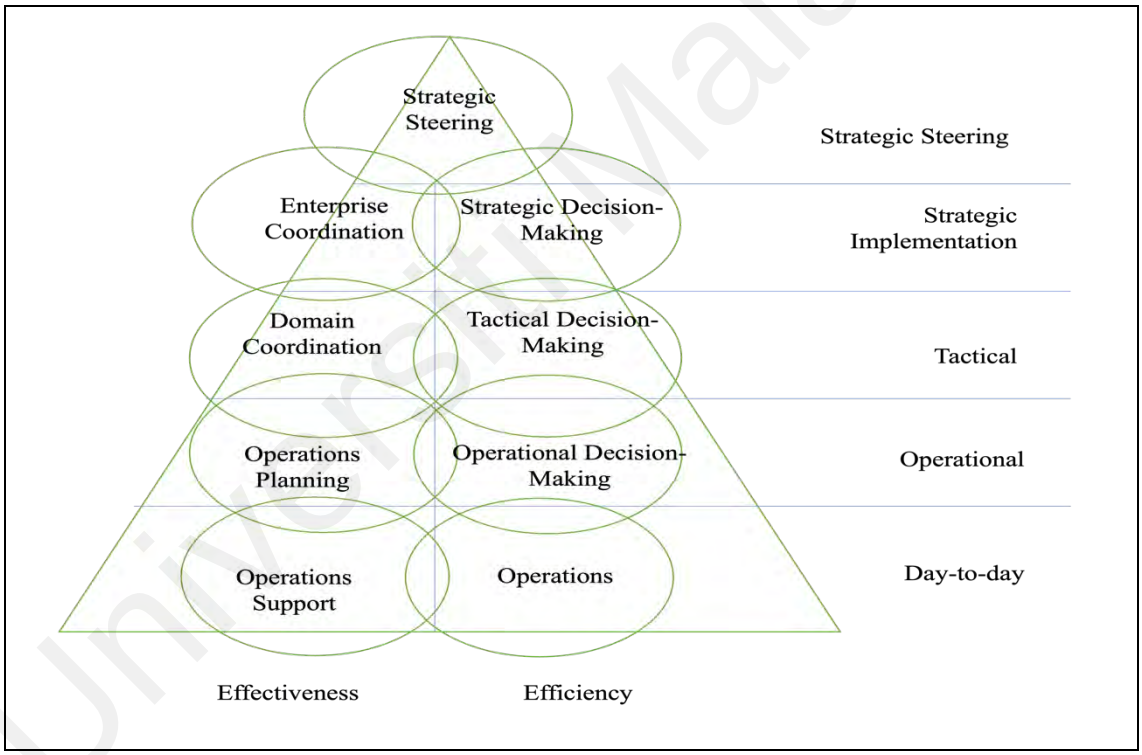


Figure 2.5: Agile Governance Model by Korhonen et al. (2013)

Data governance enables the proper management of an organization's data and information. As more data becomes available, the demand for data governance increases. Brous et al. (2016) believe prior research has concentrated on data governance structures, with less consideration for the fundamental ideas. They use a systematic literature review to analyze the guidelines and policies for data governance. They have derived four data

governance principles. For future research, it is encouraged to determine which principles apply to various units of analysis, such as organizational, group, or person.

In a study by Ahmadi et al. (2022), analyze the crucial elements necessary for effective data governance. The study aimed to propose an approach for managing relevant factors and activities for implementing data governance in an organization. By employing fuzzy logic, they offer a nuanced methodology that specifically targets organizations looking to enhance their data governance practices. The research identifies key factors and activities that are essential for implementing data governance successfully. This contribution is significant as it provides organizations with a clear framework to follow, helping them navigate the complexities of managing and securing their data more efficiently.

In their study, Norbib & Abu Bakar (2021) set out to develop a comprehensive data governance model tailored specifically for the Ministry of Education in Malaysia. By comparing and analyzing existing data governance models and frameworks, they were able to identify seven key contributions that could enhance data governance within the Ministry. Building on these insights, they proposed a conceptual model designed to address the unique needs and challenges faced by the Ministry of Education. Their work offers a structured approach to managing educational data, ensuring better data quality, security, and accessibility, ultimately supporting the Ministry in its mission to improve the educational landscape in Malaysia.

Meanwhile, Benfeldt et al. (2020) examine the complexities surrounding data governance within local government organizations, particularly focusing on Danish municipalities. Through interviews and analysis, they introduce a theoretical framework known as the problem triangle. This framework highlights and explains the intricate relationships between six critical challenges: value, collaboration, capabilities, overview, practices, and politics. By theorizing these challenges, the study sheds light on why

governing data proves challenging in local government contexts, providing valuable insights for policymakers and practitioners seeking to improve data management strategies and governance frameworks within municipal settings. Their research underscores the importance of addressing these multifaceted issues to enhance the effectiveness and transparency of data governance practices in local governments.

Scientific data stewardship is crucial in ensuring the long-term preservation, trustworthiness, and effective use/reuse of digital research data. However, many data centres lack a comprehensive, integrated framework for managing scientific data stewardship operations. In a study, Peng et al. (2018) aimed to present a comprehensive, data-centric view of managing scientific data stewardship operations through a Plan-Do-Check-Act (PDCA) cycle flow diagram. The study intends to guide improvements at all levels of data stewardship and proposes a data-centric conceptual enterprise paradigm for managing stewardship activities. However, one constraint of the framework is the lack of explicit reference to certain critical areas of scientific data stewardship.

Meanwhile, the vast amount of data generated by the environment necessitates the adoption of regulations, directives, and standards to ensure its proper governance. The complexity of today's data pipelines is compounded by current industrial supply networks. Yebenes & Zorrilla (2019) believed that there are no frameworks that have been adapted to the cloud computing architecture. Thus, they conducted a study to establish a preliminary schema for a third-generation platform's effective data governance program. The study employs a systematic review of the literature. They suggested that models, templates, and tools be developed to assist businesses in applying the framework in real-world case studies for future work.

Similarly, another authors (Al-Ruithe et al., 2016) identified a research gap regarding the lack of studies on data governance in cloud computing. To address this gap, they researched to explore the importance of data governance from both the perspectives of cloud users and providers. They employed an analytical theory approach to develop a data governance design framework for cloud computing environment. The proposed framework focused on five key processes: data governance structure, data governance assessment, data governance function, negotiation, and data governance level agreement. However, the framework has not been empirically validated, which is acknowledged as a study limitation. The authors intend to develop a more comprehensive and holistic framework for cloud data governance in future research.

In addition to the cloud computing environment, platform ecosystems such as Facebook, YouTube, and Twitter also experience significant growth through user-generated content sharing and demand. However, concerns arise regarding data misuse, privacy, revenue sharing, and the inadequacy of traditional data governance approaches in platform ecosystems. To address this gap, Lee et al. (2017) conducted a pioneering study on data governance in platform ecosystems, analyzing and surveying four platforms: Facebook, YouTube, eBay, and Uber. They compared nineteen governance models against data governance elements specific to platform ecosystems. However, the study has limitations, including the need for analysis from different perspectives (e.g., platform owners vs. users), potential validity issues in case studies, and the possibility of missing relevant literature in the literature search.

The discussion of studies on theoretical frameworks for data governance covers a wide range of domains, contexts, and settings, offering valuable insights for ensuring data quality and effective governance. Existing literature covers a broad spectrum of data governance frameworks and objectives, including formalizing concepts in various sectors

such as banking, implementing open data portals, exploring scientific publications, and comparing data governance activities. Studies also investigate data governance in telecommunications, higher education institutions, Agile Governance Model, local government organizations, scientific data stewardship, third-generation platforms, cloud computing, and platform ecosystems. However, there remains a notable gap in research on data governance specifically tailored for RPOs in Malaysia, particularly in defining the necessary activities and roles for governing research data. Addressing this gap in the current study could greatly contribute to improving RDM practices in Malaysia.

2.8.2 Key Roles

The term data governance refers to the decisions made about who should be in charge of what types of data and how they should be handled (Alhassan et al., 2016). According to Alhassan et al. (2018), establishing data roles and responsibilities is the first stage in executing a data governance program. Everyone who has a relationship with data should do everything they can to keep it safe. However, they must understand the data strategy and the rationale for each step (Gupta & Cannon, 2020b). As a result, RPOs need to form an efficient functional team to carry out the data governance tasks (Alhassan et al., 2019a; Omar & Almaghthawi, 2020) as a structural dimension. The designated tasks and duties will influence how the data governance program's other activities are carried out (Alhassan et al., 2018).

Abraham et al., (2019) split the structural dimensions into two components: the location of decision-making authority and roles and responsibilities. The function of structural dimensions and the delegation of decision-making authority are intended to define reporting structures, governance bodies, and accountability. The allocation of decision-making authority determines which organizational body has the mandate for data governance action (Abraham et al., 2019; Khatri & Brown, 2010). Making clear who

has decision-making authority, what decision-making authority he holds, and whether decision-making authority is centralized, decentralized, or both (Abraham et al., 2019). The main roles and governance bodies comprise the executive sponsor, data governance leader, data owner, data steward, data governance council, data governance office, data producer, and data consumer (Abraham et al., 2019).

a) The executive sponsor should preferably be the highest level executive (Abraham et al., 2019; Korhonen et al., 2013). The executive sponsor delivers strategic direction, business prioritization, and funding for data management (Abraham et al., 2019; Korhonen et al., 2013; Yebenes & Zorrilla, 2019), clarifies and defines the scope of the data governance program, and aids in setting completion milestones and targets, as well as assuring compliance with data laws and regulations (Yebenes & Zorrilla, 2019). At the same time, the executive sponsor must have enough authority to engage in long-term strategic planning and decision-making for the entire organization (Korhonen et al., 2013).
b) Data governance leader is responsible for the day-to-day management of the data governance program (Abraham et al., 2019) to carry out the board's directives and ensures compliance with established standards (Wende & Otto, 2007). In addition, the data governance leader gives guidance concerning data design, delivery, and maintenance and ensures compliance with data policies (Abraham et al., 2019). Furthermore, the data governance leader coordinates tasks for data stewards and delivers periodic updates on data governance performance (Abraham et al., 2019). Chief data steward is another term used to refer to data governance leaders.

c) Data owners are typically executives in charge of data assets to communicate general data requirements and hazards (Abraham et al., 2019). They make final decisions regarding the data domain, its maintenance, and development (Vilminko-Heikkinen & Pekkola, 2019), and have approval authority for data-related actions (DAMA International, 2017).

d) Data stewards are classified into business data stewards, technical data stewards, and chief stewards (Korhonen et al., 2013; Wende & Otto, 2007). Business data stewards must understand the value of data to the organization and translate business strategy into data techniques that achieve business goals, which involves coordination between business and technical teams (Abraham et al., 2019; Korhonen et al., 2013). Technical data stewards are familiar with various fields, including software development, database management, web service application development, and system integration (Abraham et al., 2019; Peng et al., 2016). Meanwhile, the duties of a chief steward have already been described under the data governance leader. Data stewards are responsible for providing data management support to other data stewards, the team from documentation and metadata, and other key stakeholders (Peng et al., 2016). They also assist in identifying data management needs and maintaining compliance with data management standards, including community standards on data quality metadata and rules (Peng et al., 2016).

e) The data governance council consists of senior officials who will determine the program's objectives and broad principles and practices (Yebenes & Zorrilla, 2019) by concentrating on enterprise-level strategy execution through design, planning, and support (Korhonen et al., 2013) and aligning it with organizational goals (Abraham et al., 2019).

f) The data governance office is a body that assists the data stewardship teams and the data governance council with governance and decision-making (Abraham et al., 2019). The data governance office is responsible for developing communication channels, scheduling meetings, coordinating issue resolution, and educating stakeholders (Yebenes & Zorrilla, 2019). This group also ensures that operational teams design and implement proper data management policies and procedures (Yebenes & Zorrilla, 2019).

g) Data producer creates the data or integrates and maintains data given by others (Abraham et al., 2019) but may not have total control over it, particularly in the case of public-funded organizations like universities (Dijkers, 2019), since the data owners normally belong to the organizations or funding agencies.

h) A data consumer is the one who consumes the information, establishes demands and reports issues with data (Abraham et al., 2019).

2.8.3 Key Areas

Understanding the key areas of RDG is essential to ensure effective management and ethical use of research data, enabling its functional dimension to maximize data value and minimize any related risks (Austin et al., 2021; Downs, 2021; Matthewson, 2019; Redkina, 2019). According to Wong et al. (2020, p. 57), functional dimension "describes the right things to do in implementing data governance in all activities within all areas". Governance areas encompass aspects or functions that require an action, as highlighted by Alhassan et al. (2018, p. 306). Meanwhile, the decision domains are the areas of RDG that require control to attain organizational objectives (Sung et al., 2019). The following section discusses governance areas categorized into procedural and relational mechanisms, as conceptualized by Abraham et al. (2019), alongside decision domains.

2.8.3.1 Procedural mechanisms

Procedural mechanisms ensure that data is accurately recorded, securely stored, effectively used, and appropriately shared. These mechanisms cover a wide range of governance areas, including data policy, data requirements, data strategy, data standards, processes and procedures, guidelines, compliance monitoring, and more.

a) Data policy - A well-defined, easy-to-follow data policy is essential to prevent data leakage, monitor data access activities, and guarantee that data is available at the appropriate time and in the correct format (Alhassan et al., 2019a). Data governance policies must align with the organization's strategic goal and provide high-level definitions of expected behaviour and outcomes in the domain areas where the business aspires (Yebenes & Zorrilla, 2019). On the other hand, few data policies explicitly indicated infrastructure support, training, and the establishment of future rules (Lefebvre et al., 2018). Bottom-up systems should be avoided since they create inconsistency and complexity; hence, governance should be a corporate-wide duty to eliminate policy misunderstandings (Tallon et al., 2013). The policies and procedures apply to decision-making at the strategic, tactical, and operational levels (Korhonen et al., 2013).

b) Data requirements - Data requirements are critical to the success of any data governance initiative. Alhassan and colleagues (2019a) believe that data requirements must be controlled and maintained through communication among all parties involved to ensure that data is implemented appropriately. Furthermore, the demands and needs of data consumers may vary over time; consequently, successful user engagement programs must connect interactively with present and future data consumers (Peng et al., 2018).

c) Data strategy - A high-level plan of action based on strategic business goals is known as a data strategy (DAMA International, 2009). Data strategy is considered the most essential of the identified governance areas because it affects the IT departments and the entire business (Omar & Almaghthawi, 2020). Organizations need a complete and integrated data governance strategy to maximize the value of their digital information assets, such as research data, while mitigating risks (Omar & Almaghthawi, 2020). A robust and complete data strategy includes all the necessary responsibilities associated with data as an asset and the preparedness for data governance (Alhassan et al., 2019a). The data strategy and business plan are the first stages in establishing a successful and long-term data governance program (Gupta & Cannon, 2020b).

d) Data standards - Data governance is the process of deciding on data quality standards and data quality management (Khatri & Brown, 2010). According to DAMA International (2009), data standards guarantee that data representation and data-related procedures are uniform and consistent. They ensure interoperability inside and across organizations and their continuing usefulness (DAMA International, 2009). Internally, data stewards and data architects produce data standards, while standardization groups such as the ISO define data standards externally (DAMA International, 2009). It is about establishing how data will be managed, including metadata management, data quality, and data security (Yebenes & Zorrilla, 2019).

e) Organization should have documented processes and procedures for each stage of the data life cycle when dealing with a significant volume of data from multiple sources (Alhassan et al., 2019a). For this specification to be utilized and monitored appropriately, the data owner of every data collection

should establish data processes and procedures (Alhassan et al., 2019a). Transparent data processes and procedures enable data quality trust (Alhassan et al., 2019a). However, to guarantee that data governance processes and principles are followed, people in the organization must first be willing to adjust and change (Benfeldt et al., 2020).

f) Data governance mandates established guidelines and rules for managing data quality (Otto, 2011). It is critical to consider establishing organizational structures, employing governance mechanisms to improve data quality, managing resources across a single company, and establishing intra-organizational data management guidelines (Al-Ruithe et al., 2019).

g) Compliance monitoring, for instance, aims to track and enforce adherence to legal requirements, business rules, standards, and processes (Al-Ruithe et al., 2019). This includes professional data supervision, data management initiatives, and service control (DAMA International, 2009).

Besides that, risk management, disaster recovery planning, data security, data privacy, data integrity, incident response, data access management, and accountability are all data governance areas of concern for digital data stored in the cloud (Al-Ruithe et al., 2016). According to some authors (Feki & Boughzala, 2016), risk management includes identifying risks and vulnerabilities and designing and implementing processes to mitigate their effects on the organization. Other data governance procedural mechanism areas of concern include strategic alignment, master data management, data profiling, data cleansing (Ahmadi et al., 2022), data stewardship (Peng et al., 2018), and data ownership (Brous et al., 2016).

2.8.3.2 Relational mechanisms

While, relational mechanisms comprise communication, training, and decision-making coordination to facilitate collaboration between stakeholders (Abraham et al., 2019). A communication strategy may help by defining stakeholders, communication channels, supporting tools, and metrics to retain commitment (Al-Ruithe et al., 2019). Open and ongoing communication among key product players and stakeholders and a clear approach to prioritizing and addressing problems presented by data consumers are vital (Peng et al., 2016). In addition, an organization should offer staff and faculty members periodic data governance training, create enforcement and follow-up requirements, and frequent data governance plans and policies assessments (Omar & Almaghthawi, 2020) as the capacity to analyze and apply data effectively to guide decisions (Mandinach & Gummer, 2013).

2.8.3.3 Decision domains

The identification of decision domain areas is based on frameworks by (Abraham et al., 2019; Alhassan et al., 2016; Khatri & Brown, 2010), which includes data principles, data quality, metadata, data access, data life cycle, data security, data architecture, data storage, and infrastructure. It aligns with the importance of having a consistent and agreed-upon data domain structure to simplify efforts toward data management, data governance, and data quality improvement initiatives (Allen & Cervo, 2015). Each decision domain is elaborated upon in the subsequent sections to enhance comprehension.

a) Data principles - From the perspective of data principles, Brous et al. (2016) have defined four principles based on their literature review; they are organizational dimension, alignment, compliance, and common understanding.

b) Data quality - The potential to boost organization performance is by improving data reliability (DAMA International, 2017). Data governance should be seen as a tool to correct poor data quality emerging from data quality management (Alhassan et al., 2016, 2018; Al-Ruithe et al., 2019; Benfeldt et al., 2020). Data quality management is vital for organizations to respond to strategic and operational concerns that need high-quality data (Weber et al., 2009). The planning, distribution, organization, consumption, and disposal of high-quality data are all covered under data quality management (Weber et al., 2009).

c) Metadata - Metadata specifies an organization's data, representation, classification, origin, movement, evolution, who can use it, and whether it is of high quality (DAMA International, 2017). Metadata management typically serves as a starting point for better data management (DAMA International, 2017). It integrates, controls, and provides high-quality metadata (Wang et al., 2018).

d) Data access – It is the process of storing, accessing or transferring data in a database that requires authorization (Chambers et al., 2019). Data access privileges and data availability at the proper time and in the right format are all part of accountable data access and availability (Alhassan et al., 2019b). Therefore, it is critical to review and tighten data access rules to ensure data safety (Al-Ruithe et al., 2018), security and prevent data leaking (Alhassan et al., 2019b).

e) Data lifecycle - Data governance is based on many processes associated with the data lifecycle (Omar & Almaghthawi, 2020). The process includes creating, processing, analyzing, preserving, giving access, and reusing research data (Ghent University, 2016; UK Data Archive, 2018). It is critical

to identify each piece of data, monitor its evolution, connect it to strategic business objectives (Yebenes & Zorrilla, 2019), and continuously examine data quality throughout its lifecycle.

f) Data security - Data security relates to preserving data accessibility, authenticity, availability, confidentiality, integrity, privacy, and dependability security standards (Singh et al., 2021). Much effort is required when dealing with big data in data security and privacy. As a result, policies must address all data issues, including data security and risk, metadata management, and data quality. However, the researchers' primary concern is not data management, preservation, or sharing but data security and the security of the devices used for storage and analysis (Schöpfel, 2019).

g) Data architecture - Data governance includes determining an organization's data requirements, emphasizing data architecture (Abraham et al., 2019; DAMA International, 2009). Data architecture is a key requirement of a successful data management organization (DAMA International, 2017). Data architecture defines the data asset management strategy to meet the needs of the organization's data (Wang et al., 2018). Data strategy and data architecture are critical for coordinating 'doing things correctly' and 'doing the right things' while managing data assets (DAMA International, 2017).

h) Data storage and infrastructure - They focus on IT artefacts that enable effective data management across the organization (Abraham et al., 2019). In recent years, the application demand for massive data storage has supported the development of high-performance storage technology (Al-Ruithe et al., 2018). Compared to past data storage responses, researchers used their own devices to keep their data due to a lack of institutional policies and concerns about data sharing (Palsdottir, 2021; Vilar & Zabukovec, 2019). However,

since researchers are fully aware of their institution's data storage requirements and its support for open data sharing, they must be capable of implementing comprehensive research data storage methods (Adika & Kwanya, 2020).

Overall, the review of related frameworks is useful for constructing a conceptual framework for understanding RDG. By incorporating these frameworks, models, and introduced concepts, general understanding of data governance activities and the key roles in promoting effective data governance were obtained.

2.9 Research Gap

This review examines various dimensions and domains within the existing literature. RDG is essential for ensuring quality, access, and ethical use of research data. Numerous research gaps exist in RDG, particularly in research methodology, stakeholder involvement, and frameworks for data management.

One notable gap in RDG research methodologies is the reliance on traditional methods like literature reviews, interviews, and questionnaires to gather data and answer research questions (Abraham et al., 2019; Alhassan et al., 2018; Al-Ruithe et al., 2016; Brous et al., 2016; Lee et al., 2017; Norbib & Abu Bakar, 2021; Yebenes & Zorrilla, 2019). While these methods can provide valuable insights, they often have limitations in addressing the complexities of RDG practices. To fill this gap, research on RDG could benefit from more innovative or mixed method approaches that combine qualitative and quantitative data collection methods. These methods would allow researchers to observe RDG practices in real time and engage with a wider range of stakeholders. Such methodologies would provide richer, more actionable insights that could guide the development of more effective, context-sensitive RDG frameworks.

In addition, there is a gap in the literature regarding the specific activities and roles involved in RDG and how they should be implemented within RPOs. Many studies describe RDG in other aspects (Benfeldt et al., 2020; Kaewkamol, 2022; Norbib & Abu Bakar, 2021; Omar & Almaghthawi, 2020; Peng et al., 2018; Rachmat et al., 2023) but lack actionable insights into the day-to-day governance activities that ensure effective data management, data sharing, and security. In Malaysia RPOs, data quality, storage, compliance, and sharing remain critical challenges due to the lack of standardized guidelines and insufficient data management practices (Amanullah & Abrizah, 2023; Hodonu-Wusu et al., 2020; Ismail et al. 2022; Napis et al., 2019; Olesen et al., 2018; Singeh et al., 2013b; Zain et al. (2023). These issues are compounded by unclear governance activities and roles within the organizations. By providing a clear roadmap of governance activities, it could ensure that research data is consistently managed, securely stored, and appropriately shared across departments and disciplines.

RDG in RPOs is a multifaceted effort involving a diverse array of stakeholders, entities, individual roles, and governing activities. Currently, there is no overarching framework that interrelates the involved parties, roles, and activities. These organizations generate a diverse range of research data, and it is vital to identify roles and its governance structure (structural dimension) and key activities; tasks, areas, and decision domains (functional dimension) for effective RDG and stewardship, improving data quality and ensuring equitable access to research data. Thus, a more comprehensive framework specifically designed for RPOs to govern research data is needed.

Therefore, the current study holds significant importance in addressing the research questions and fulfilling the research objectives. The modified Delphi method, chosen as the field research approach, is suitable for data collection due to its ability to gather insights from a panel of experts comprising data practitioners from various research

institutions. These experts hold diverse levels of management and/or governance roles in handling research data and possess knowledge and expertise in managing research and/or publication data. The study aims to explore the RDG activities implemented in these institutions with the key stakeholders who govern research data, to develop best practices to be incorporated into an RDG framework. This framework will serve as a guide for Malaysia RPOs looking to initiate RDM or improve their existing practices.

2.10 Summary of Chapter Two

The evolving body of literature on data governance sheds light on RDG in a broader sense. A notable research gap lies in exploring the governance roles and the activities necessary for establishing effective RDG. The literature review provides evidence that an RDG model or framework has yet to be proposed for RPOs in Malaysia. Thus, the study aims to develop a framework on RDG to promote a data-driven research community and well-managed research data specifically for RPOs. Two primary methods of data collection utilized to address the research questions of the study, including desk research and a modified Delphi study. The subsequent Chapter Three discusses the research design and rationale. Furthermore, the chapter examines the implementation of the Delphi method as a form of action research, elucidating its specific application in this study to gather insights from a panel of experts regarding the topic under examination.

CHAPTER 3: METHODOLOGY

"If a disobedient person brings you information,
investigate it, lest you harm a people in ignorance…"
(Surah Al-Hujurat, 49:6)

3.1 Introduction

This chapter describes the research methods, outlining the steps from initial assumptions to the collection and analysis of research results. It outlines the procedures employed in this study and provides the rationale for selecting a qualitative approach for both desk research and the application of the modified Delphi technique.

The study aimed to investigate the actual implementation of RDG tasks within RPOs in Malaysia and identify essential RDG activities for their implementation (functional dimension). Additionally, the objectives included identifying stakeholders for each role and establishing the governance structure (structural dimension), based on expert opinions gathered through multiple rounds of questionnaires. Ultimately, the study sought to formulate an RDG framework tailored for RPOs in Malaysia. To achieve these objectives, the following research questions were formulated to guide the study:

RQ1: How are research data governance activities currently being implemented by data practitioners in research performing organizations?

RQ2: What research data governance activities do data practitioners consider significant for research performing organizations?

RQ3: Which stakeholders do data practitioners in research performing organizations consider responsible for research data governance roles?

RQ4: How do data practitioners identify the structural positions of research data governance roles within research performing organizations?

This chapter first provides an overview of the research paradigm, design, and sampling strategy, followed by a description of the data collection and analysis procedures, validity reliability, and trustworthiness of data, as well as ethical considerations. It then concludes with the chapter summary.

3.2 Research Paradigm

The methodological approach adopted in research is significantly influenced by the research paradigms and philosophical considerations. In this instance, the study was guided by a post-positivist philosophical paradigm, which acknowledges that scientific knowledge is not fixed but depends on specific circumstances and can be further refined based on new evidence and observations (Creswell & Creswell, 2018). This perspective is particularly relevant to the study of data governance, as it recognizes that a one-size-fits-all approach is not suitable (Weber et al., 2009) since each entity possesses unique requirements and preferences, necessitating a flexible and context-sensitive approach. Post-positivism also accepts the presence of an objective reality independent of human perception (Ryan, 2006), however, the understanding of this reality is mediated by the interpretation and analysis of study participants (Beeftink, 2005).

This dual acknowledgment of an objective reality and the interpretive process is a key connection to the critical realist perspective, which posits that while there are objective realities in the world, the understanding of these realities is always shaped by human cognition, social context, and the limits of perception (Bhaskar, 1975). Similarly, constructivism emphasizes the subjective nature of knowledge, suggesting that individuals construct their understanding through lived experiences and social interactions (Alanazi, 2016). This makes post-positivism, in combination with elements of critical realism and constructivism, a suitable research paradigm for studying the

complex and evolving field of data governance, where diverse perspectives and contexts play a crucial role.

The modified Delphi technique used in this study aligns well with the post-positivist research paradigm, while also drawing on principles of critical realism and constructivism. The modified Delphi method provides a flexible, iterative, and qualitative approach that recognizes the subjectivity of knowledge and values diverse perspectives (Cuhls, 2023). It enables the exploration of complex issues through comprehensive desk research, the consideration of multiple viewpoints presented by a panel of expert practitioners, and the convergence towards shared understandings over multiple rounds to arrive at consensus. This process reflects key principles of post-positivism, particularly its emphasis on fallible knowledge, subjectivity, and interpretation (Tauber, 2022). Moreover, it resonates with critical realism by acknowledging that the objective realities of data governance and related phenomena exist but are always understood through the subjective lenses of expert participants. The iterative nature of the modified Delphi also complements constructivism by highlighting that knowledge is constructed through the collective insights of individuals, with meaning emerging through ongoing dialogue and the sharing of perspectives. The modified Delphi method, thus, becomes an appropriate tool for uncovering and refining expert knowledge in the context of a post-positivist and constructivist approach, where knowledge is seen as evolving, socially constructed, and contingent on the context. Several key connections between the post-positivist paradigm, critical realism, constructivism, and the modified Delphi technique are highlighted in Table 3.1.

Table 3.1: Connections between Post-Positivism, Critical Realism, Constructivism, and Delphi technique in this Study

Post-Positivism	Critical Realism	Constructivism	Modified Delphi Technique
Emphasizes the empirical evidence and testing of theories using existing data (Panhwar et al., 2017).	Acknowledges that an objective reality exists but that our understanding is mediated through subjective perception (Bhaskar, 1975).	Knowledge is constructed through interaction and interpretation, emphasizing the role of social context and individual experience (Alanazi, 2016).	Information is collected via instruments based on measures completed by participants (Creswell & Creswell, 2018).
Recognizes that scientific knowledge is constructed through interpretation and analysis (Tauber, 2022).	Explores the mechanisms and structures that exist independently of human perception, but always understood through subjective experiences.	Knowledge emerges through social processes and collaborative interaction among participants.	The instruments are designed based on desk research, acknowledging that knowledge construction is influenced by interpretation.
Acknowledges that knowledge is influenced by the researcher's perspective, and multiple interpretations may exist (Panhwar et al., 2017).	Knowledge is shaped by human cognition and social context, acknowledging both objective and subjective dimensions.	Emphasizes the subjective and contextual nature of knowledge, highlighting the role of individual experience in the construction of meaning.	Incorporates multiple rounds of anonymous feedback from a panel of experts, minimizing individual biases and exploring diverse perspectives (Ryan, 2006).
Knowledge is conjectural and values diverse perspectives in understanding complex phenomena (Creswell & Creswell, 2018).	Recognizes that reality is complex and multifaceted, with mechanisms that can only be uncovered through deep investigation and interpretation.	Knowledge is not absolute but constructed through collective human interaction, meaning-making, and shared understanding.	Involves a panel of experts from various backgrounds, ensuring a wide range of opinions and insights are considered.
Recognizes the significance of subjective experiences and interpretations in the construction of knowledge (Panhwar et al., 2017).	The objective reality is independent but unknowable in its entirety; understanding is partial and contingent on human perception.	Encourages exploration of diverse perspectives to construct a richer, multifaceted understanding of the issue at hand.	Facilitates the convergence of expert opinions over multiple rounds to achieve consensus or highlight persistent disagreements.

The research paradigm and the data collection technique are closely aligned through their shared emphasis on the subjective nature of knowledge and the importance of diverse perspectives in understanding complex phenomena. Post-positivism, critical realism, and constructivism acknowledge that knowledge is always influenced by the researcher's perspective and interpretation, and that multiple interpretations can exist depending on context and experience (Panhwar et al., 2017). Similarly, the modified Delphi technique integrates multiple rounds of anonymous feedback from experts,

enabling the exploration of a range of viewpoints and minimizing the impact of individual biases. By gathering diverse perspectives and facilitating the convergence of expert opinions, the modified Delphi technique aims to achieve a deeper understanding of complex issues, potentially reaching consensus on specific topics. Therefore, the connection between the post-positivist paradigm, critical realism, constructivism, and the modified Delphi technique lies in their shared commitment to embracing subjectivity, collaborative knowledge construction, and the provisional nature of knowledge.

3.3 Research Design

Several research designs in the qualitative method include interviews, case studies, narrative, ethnography, phenomenology, and grounded theory (Creswell & Poth, 2017; Denzin, 2017). In the context of this study, the approach utilized was the sequential exploratory method (Whitehead & Day, 2016). In this method, the development of instruments was grounded in qualitative content analysis derived from the findings of the desk research. Quantitative and qualitative data were subsequently collected from the panelists of the Delphi technique, selected for their expertise and knowledge, as demonstrated by Beiderbeck et al. (2023) and Saihi et al. (2023) in their studies. These data were analyzed to develop an RDG framework based on the obtained findings. Table 3.2 provides an overview of the research questions and methodologies for studying research data governance in Malaysia RPOs.

The study commenced with a review of the relevant literature to propose a conceptual framework, followed by two phases of data collection: initially through desk research, and subsequently via a modified Delphi study. The process flow of the study across all phases is illustrated in Figure 3.1.

Table 3.2: Data collection strategy

NO	QUESTION	METHOD	SUBJECT & SAMPLING	OUTCOME
RQ1	How are research data governance activities currently being implemented by data practitioners in research performing organizations?	1. Desk research 2. Modified Delphi-Round I	1. Policy documents 2. Data practitioners from Malaysia research performing organizations	1. List of RDG tasks with their assigned roles were determined (Desk research). 2. Modified Delphi Instruments were developed. 3. RDG tasks implemented by Malaysia RPOs were identified (Delphi study).
RQ2	What research data governance activities do data practitioners consider significant for research performing organizations?	Modified Delphi-Rounds I-III	Data practitioners from Malaysia research performing organizations	A consensus on RDG activities was reached.
RQ3	Which stakeholders do data practitioners in research performing organizations consider responsible for research data governance roles?	Modified Delphi-Rounds I-II	Data practitioners from Malaysia research performing organizations	A consensus on the stakeholders (key players) for RDG roles was identified.
RQ4	How do data practitioners identify the structural positions of research data governance roles within research performing organizations?	Modified Delphi-Rounds III-IV	Data practitioners from Malaysia research performing organizations	A consensus on the structural position of each RDG role was determined.

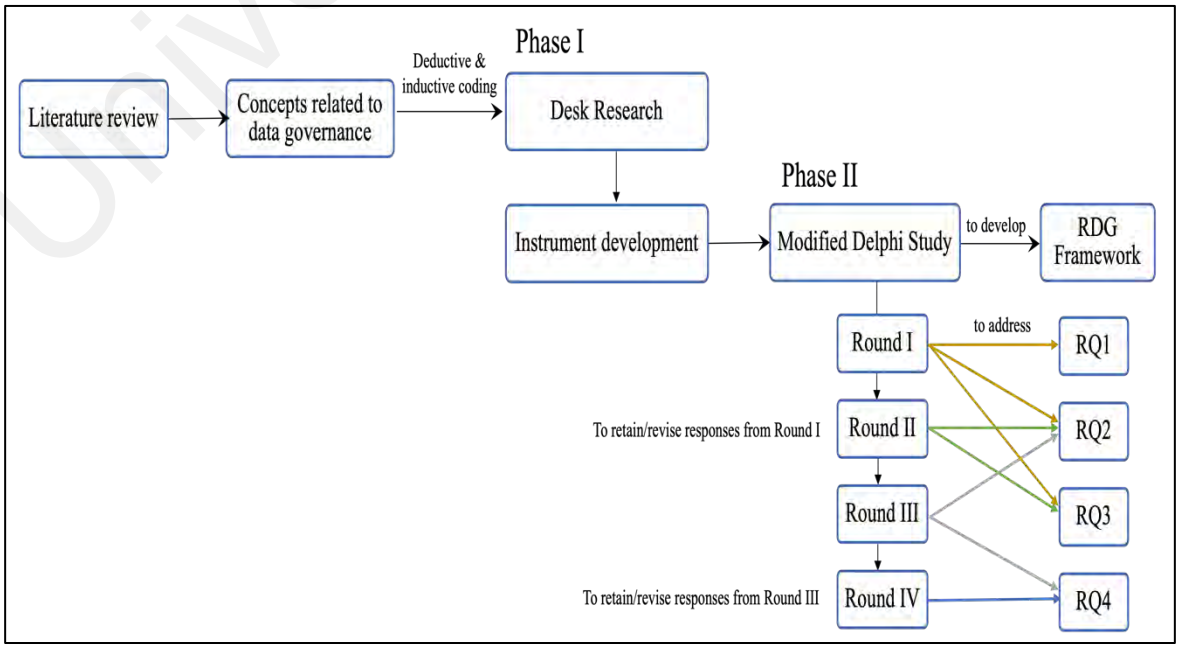


Figure 3.1: Overview of key phases and steps in the research design process

3.3.1 Conceptual framework to comprehend RDG

The main concepts and terms related to data governance practices were derived from the literature review. After reviewing the existing literature and identifying the theoretical frameworks that underpin the study, a conceptual framework for understanding RDG (see Figure 3.2) was derived by summarizing the main elements of data governance. It contains essential data governance concepts and offers supportive details for prescriptive knowledge, which are sourced from various data governance frameworks.

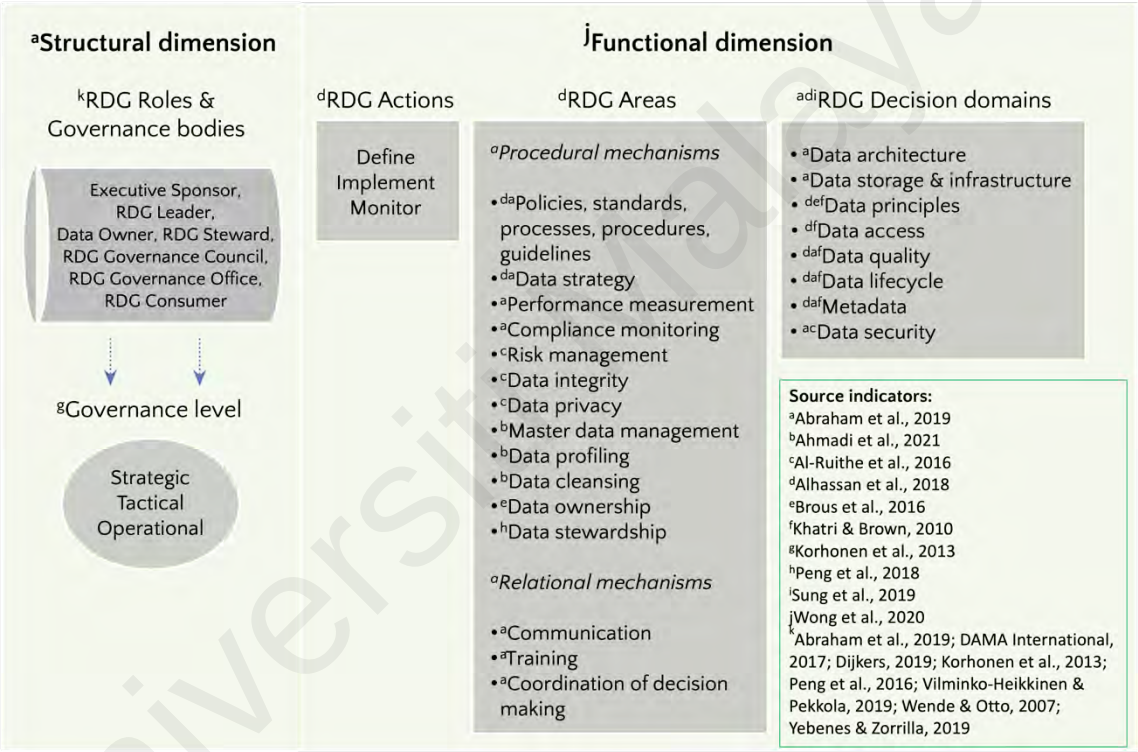


Figure 3.2: A conceptual framework for understanding RDG

Figure 3.2 illustrates the conceptual framework incorporating two main dimensions: structural and functional. The governance mechanisms proposed by Abraham et al. (2019)—structural, procedural, and relational—are integrated within this framework and have been adapted by other studies, including those by Al-Wahshi et al. (2022) and Rachmat et al. (2023). The structural mechanism corresponds to the first dimension, while the procedural and relational mechanisms are categorized as RDG areas within the

functional dimension. In this study, the focus was on data governance mechanisms where this dimension is very strategic in striving for the implementation of RDG in Malaysia.

Besides that, to explore the RDG activities implemented by RPOs, the current study relied on the Data Governance Activities Model proposed by Alhassan et al. (2018), which was also utilized in later studies (Alhassan et al., 2019a; Kaewkamol, 2022). Given that among the study's objectives was to gain a deeper understanding of how RDG is executed in practice, following the approach used by Alhassan et al. (2016; 2018) was essential.

According to Alhassan et al. (2018), data governance activities are the necessary actions and requirements needed to be performed to ensure that data is effectively governed. The literature related to the emergence of data governance helps to conceptualize the areas and decision domains related to RDG. In reviewing the literature on data governance frameworks (Abraham et al., 2019; Ahmadi et al., 2022; Al-Ruithe et al., 2016; Alhassan et al., 2018; Brous et al., 2016; Peng et al., 2018), the areas of data governance were identified and used as a preliminary guide to identify the RDG areas mentioned in the selected policy documents.

Furthermore, a combination of the frameworks proposed by Abraham et al. (2019) and Alhassan et al. (2016) were selected to present the decision domains that should be considered for RDG. Decision domains are the areas of RDG that require control to attain organizational objectives (Sung et al., 2019, p. 6380). Decision domains are crucial because the definition of data governance also pertains to who has decision rights and accountability regarding an enterprise's data assets. Therefore, identifying the decision domains is necessary to assign the appropriate responsibilities and duties (Alhassan et al., 2016). By incorporating these frameworks and models in the study, a comprehensive analysis of RDG activities and the roles in promoting effective RDG could be provided.

Additionally, the study focused on exploring the structural perspective that would be most suitable for Malaysia RPOs. Thus, the Agile Governance Model by Korhonen et al. (2013), (see Figure 2.5), acted as a guiding framework to ascertain the most suitable placement of RDG roles whether to emphasize effectiveness or efficiency aspect within the organizational structure. The identification of the main data governance roles were informed by various authors (Abraham et al., 2019; DAMA International, 2017; Dijkers, 2019; Korhonen et al., 2013; Norbib & Abu Bakar, 2021; Peng et al., 2016; Vilminko-Heikkinen & Pekkola, 2019; Wende & Otto, 2007; Yebenes & Zorrilla, 2019).

3.3.2 Phase One: Desk Research

The first phase (desk research) involved examining RDG/RDM policy documents to identify key aspects, such as major activities and responsible stakeholders, aiming to gain insights into the current RDG implementation within leading RPOs worldwide. This phase was crucial not only for understanding existing practices but also for developing instruments needed for the subsequent Delphi study. The concepts for the conceptual framework were applied to guide the content analysis as part of the desk research by initiating a deductive coding process.

In the context of this study, desk research, also known as secondary research, was initially conducted to analyze existing information before engaging in field research and collecting primary data through the Delphi study. This approach is in line with the understanding that desk research involves summarizing, organizing, and synthesizing previously published research in various formats (Goundar, 2012; Kabir, 2016; Woolley, 1992). The primary objective of the desk research in this study was to acquire background information on RDG activities to establish a foundational understanding. In addition, it served as the basis for developing instruments for the Delphi study.

In this desk research, a qualitative content analysis was employed to examine existing data, as exemplified by Jacobsson et al. (2022) & Sukmasetya et al. (2018) in their respective studies. Desk research, involving the examination of written policies on RDG/RDM, naturally aligned with content analysis to categorize data into themes or categories not immediately apparent from the raw data (Vaismoradi et al., 2013) and to identify recurring themes and concepts (White & Marsh, 2006). Although there is no universal definition of content analysis, for this study, it is defined as "a research technique for making replicable and valid inferences from texts to the contexts of their use" (Krippendorff, 2004, p. 18). It was selected for its reliability and validity as a tool for systematically analyzing large amounts of data (Kabir, 2016; Krippendorff, 2004).

As a result, this desk research provides a solid foundation for designing subsequent Delphi study in the second phase, serving as a benchmark for evaluating the findings of the collected primary data, as recommended by Kabir (2016). The complete research process of the desk research, including its procedures and findings are elaborated in Chapter Four.

3.3.3 Phase Two: The modified Delphi study

This section provides background information of a Delphi study to enhance readers' understanding. The subsequent discussion will elaborate on how the modified Delphi method was applied in this study.

3.3.3.1 Introduction of Delphi study

Developed in the mid-20th century, the Delphi technique has proven to be a valuable tool across various scientific disciplines. In the 1950s, Norman Dalkey and Olaf Helmer, who worked at The Rand Corporation, created the Delphi technique to reach a trustworthy expert agreement on the technological impact on warfare (Barrett & Heale, 2020; Custer et al., 1999). This approach was named after the ancient Greek oracle of Delphi, renowned

for its ability to predict the future (Olsen et al., 2021). The Delphi approach has been frequently used in various scientific disciplines such as health care (Azimi et al., 2021; Fanniff et al., 2023; Lim et al., 2020; Negev et al., 2023; Wardlow et al., 2023), medicine (Garner et al., 2022; Schropp et al., 2023), education (Berger et al., 2023; Tangalakis et al., 2023), commercial (Nemati et al., 2020; Tamošaitienė et al., 2021), social sciences (Markmann et al., 2021), and information science (Marlina et al., 2022; Saihi et al., 2023).

The conventional Delphi method is a research approach that involves generating a research question and posing it to select experts in a specific field. These experts participate in multiple rounds of questionnaires to generate and refine a list of items in response to the research question (Wardlow et al., 2023). The initial round typically employs open-ended questions to gather expert opinions and potential solutions (Olsen et al., 2021). Subsequent rounds involve narrowing down and ranking responses until a consensus is reached among the participants.

According to Dalkey (1969), a Delphi study is distinguished by three key characteristics: anonymity, controlled feedback, and statistical group response. Due to its anonymity, a Delphi study allows professionals to communicate their knowledge and opinions without group dynamics or peer pressure (Campbell et al., 2021; Custer et al., 1999). The controlled feedback, which involves summarizing previous stage results and presenting them to participants for consideration in subsequent rounds, enables the efficient solicitation and synthesis of expert opinion, particularly with respect to complex and uncertain topics (De Lima & Seuring, 2023; Huang et al., 2022; Jünger et al., 2017; Naranjo et al., 2023; Olsen et al., 2021). While statistical group response establishes a consensus threshold and protects individual responses by sharing data in aggregate form.

There continues to be methodological uncertainty and a lack of clear guidelines for conducting a Delphi study. More specifically, ongoing debates persist regarding the number of rounds, the size of panel members to recruit, definition of expertise, the definition of consensus, and issues with the use of different forms of statistical analysis.

The appropriate number of rounds in a Delphi study can vary depending on various factors, such as the complexity of a research topic (El Miedany et al., 2021), the level of disagreement among panelists (Olsen et al., 2021), and the available time and resources (Keeney et al., 2006). Typically, studies may consist of two to four rounds, however, this may vary. Previous research has demonstrated a variety of round counts, including studies with a single round (Marlina et al., 2022), two rounds (Aghimien et al., 2023; Colonna et al., 2022; Lim et al., 2020; Negev et al., 2023; Saihi et al., 2023), between three to five rounds (De Lima & Seuring, 2023; Delaney, 2023; Drury et al., 2023; Esfandiari et al., 2023; Kiernan et al., 2023), and six or more rounds (Glascock et al., 2023).

There is ongoing debate regarding the size of Delphi study panels (Zartha Sossa et al., 2019). According to Akins and Cole's (2005) study, using a small sample of a limited number of experts in the field of study is permissible and would produce reliable results. Based on the findings of a study (Zartha Sossa et al., 2019), the number of experts ranges from 11 to 20, 21 to 30, and 31 to 40 have the highest frequency of use in the previous Delphi studies. This is consistent with the findings of de Loë et al. (2016), who found that the ranges of experts with the highest frequency of use were 11 to 20, 21 to 30, and 41 to 50. However, for the initial round, the sample size required to get expert opinions must be large enough to allow possible dropouts (Gotay, 2020; Huang et al., 2022). According to Humphrey-Murto et al. (2017), less than six participants for a Delphi study are likely insufficient; the ideal number of panel members to facilitate the consensus-building process is 10 or more (Jünger et al., 2017; Lim et al., 2020). With the increased sample

size, the quality of consensus is more convincing and can be verified (Mustaffa & Ghani, 2021). However, rather than focusing on sample size alone, it is also crucial to assess the representativeness of the expert panel (Beiderbeck et al., 2021; Bloor et al., 2015; Gotay, 2020).

Determining consensus is one of the most controversial methodological issues in the Delphi process. Although achieving consensus among participants is a key feature in Delphi studies, what is accepted as consensus, or how it is reached, remains unclear (Barrett & Heale, 2020; Keeney et al., 2006). Within the context of Delphi research, consensus denotes the extent of agreement or the convergence of viewpoints among participants concerning a specific issue or topic (Beiderbeck et al., 2021; Mustaffa & Ghani, 2021).

In a previous study by Sforzini et al. (2022), they characterized consensus as a measure of agreement versus disagreement concerning a specific recommendation. They further delineated strong consensus when this agreement percentage reached or exceeded 95 percent, moderate consensus falling within the range of 61 percent to 94 percent, and weak consensus spanning from 51 percent to 60 percent. In contrast, another study conducted by Huang et al. (2022) employed a 5-Likert scale, utilizing the median, interquartile range (IQR), and the level of agreement as metrics for measuring consensus. Specifically, consensus was established when the median score fell within the top two response options (such as "suitable" or "very suitable"), the IQR was one or less, and the combined responses from the top two categories exceeded 70 percent. In this study, an agreement rate of 70 percent among the expert panel was defined as the consensus threshold. Another research conducted by Barber et al. (2023), where the objective was to identify which subjective and objective tests content experts utilize to inform decision-making regarding the return to sport (RTS) for athletes following an upper extremity (UE)

injury. Statements that failed to achieve a 70 percent agreement rate were modified and subsequently reintroduced in future surveys.

Conventional Delphi method can be modified for various purposes, audiences, and accessibility. Modifications may involve starting with pre-selected items from existing literature or expert interviews (Kiernan et al., 2023), which can expedite the process and reduce the number of rounds required. There is no standard Delphi procedure; rather, each one is tailored to the specific needs of the project and the questions being asked (Skulmoski et al., 2007). Researchers and practitioners can adapt the Delphi technique to their needs, as it includes a variety of methodologies such as classical Delphi (Gotay, 2020), policy Delphi (Akartuna et al., 2022; Negev et al., 2023), real-time Delphi (Garner et al., 2022), e-Delphi (Gonçalves et al., 2023), and modified Delphi (Azimi et al., 2021; Barber et al., 2023; Huang et al., 2022; Mustaffa & Ghani, 2021; Wardlow et al., 2023).

3.3.3.2 Application of the modified Delphi study

Given the limited information on RDG in Malaysia, the application of a modified Delphi technique was deemed appropriate for this study, aligning with the recommendation of Rowe and Wright (2011), Barrett and Heale (2020), De Lima and Seuring (2023), and Drury et al. (2023), who suggested using a modified Delphi when information is scarce. Due to the necessity of obtaining consensus from a group of experts, this modified Delphi approach was well-suited as it offered a systematic methodology for collecting and consolidating expert viewpoints (Barrett & Heale, 2020; Beiderbeck et al., 2021; Olsen et al., 2021).

The modified Delphi approach used involves a qualitative approach in collecting quantitative and qualitative results in all its rounds. This approach was chosen in this study because it is well-suited for gathering insights from experienced data practitioners and achieving a consensus on important elements related to RDG (Keeney et al., 2006;

Olsen et al., 2021; Rowe & Wright, 2011) for the development of an RDG framework. This approach aligns with the post-positivist paradigm described and justified earlier (see Table 3.1), which emphasizes the importance of empirical evidence and theories based on existing data (Jick, 1979, as cited by Prayag, 2023). Overall, utilizing a post-positivist paradigm and using a qualitative approach in examining quantitative and qualitative results within a post-positivistic framework is more effective, promoting in-depth exploration of phenomena across diverse perspectives and contexts (Panhwar et al., 2017).

In this study, in order to eliminate subject bias, anonymity was protected, providing an equal chance for each panelist to present and react to ideas (Negev et al., 2023). The process was iterative, involving four rounds of surveys and feedback. Each round built upon the results of the preceding round, all with the aim of reaching the a priori consensus threshold (Hasson et al., 2000; Olsen et al., 2021). The modified Delphi study was conducted to address research questions of the study, as described below:

a) The initial round focused on exploring the actual implementation and perceived importance of RDG activities among RPOs in Malaysia, while also gathering expert opinions and insights on the significant component of RDG (i.e., RDG roles' key stakeholders).
b) The second round aimed to build expert consensus on the key stakeholders and essential tasks for RPOs in Malaysia, which did not meet the a priori consensus threshold in the first round.
c) The third round of the Delphi study aimed to achieve consensus on the RDG structural positions, identifying the essential RDG areas and decision domains, determining suitable nomenclature for roles, and verifying descriptions for RDG roles, areas, and decision domains.

d) The last round offered the panelists a chance to refine their replies for the items that did not receive a consensus in the previous Round III.

In this adaptation, a common modification was implemented involving commencing the initial round with a set of selected items drawn from the previous desk research. The similar approach was undertaken by other authors (Albotoush & Shau-Hwai, 2023; Custer et al., 1999; Kiernan et al., 2023; Münch et al., 2021). It differs from the conventional Delphi method, where the initial step typically entails distributing an open-ended questionnaire to the panelists (Olsen et al., 2021). Instead, in Round I of the modified Delphi, the process was initiated with four sets of instruments, mainly comprising of closed-ended, structured questions.

Aside from that, in the first round of the study, the panelists were asked on the actual implementation of RDG tasks at their institutions, with no need for agreement or consensus. It was conducted to explore the phenomena by gaining a better understanding of the current RDG practices in Malaysia RPOs. Although the current study used instruments featuring structured, closed-ended questions, it nevertheless permitted participants to provide input through open-ended sections, allowing them to share suggestions, comments, or seek clarifications for each provided statement or question. The other modification was that, in the third round, the instruments were crafted to seek information in a different perspective to accommodate newly introduced dimensions.

This modification to the Delphi method offers two key benefits: firstly, it generally enhances and manages the quality of responses in the initial round, and secondly, it provides a versatile approach for obtaining answers to the research questions. Data from all rounds were utilized to develop an RDG framework for Malaysia RPOs. This stage entailed synthesizing expert inputs and proposing an RDG framework based on the

panelists' collective agreement. This framework will aid in the implementation of best RDG practices within RPOs.

3.4 Rationale for the modified Delphi approach

There is a strong rationale for using the modified Delphi approach in this study. This study necessitates engaging individuals who are geographically dispersed, as highlighted by the authors (Humphrey-Murto et al., 2019; Taylor et al., 2020), to tap into their expertise in assessing the importance of RDG elements, including governance roles and their ideal stakeholders, as well as RDG tasks, areas, and decision domains. The modified Delphi method presents logistical advantages ideally suited for this purpose. Specifically, the modified Delphi method was selected for this study because it collects expert opinions from geographically scattered individuals (El Miedany et al., 2021).

The ability of the modified Delphi study to provide both a reliable and creative means for exploring ideas further underpins the rationale (Lim et al., 2022). This method is widely acknowledged as an efficient communication tool among experts, systematically facilitating the exchange of information and nurturing the development of a collective group consensus (Mustaffa & Ghani, 2021; Olsen et al., 2021).

Furthermore, the approach allows for a gradual accumulation of opinions while avoiding bias towards adversarial procedures. It ensures that group members are consistently kept informed about the status of the collective opinion. It empowers the panelists to appraise and scrutinize RDG components considered important for the implementation of RPOs in Malaysia. The panelists may then discern items that were previously considered insignificant, allowing them to clarify, enhance, or revise their viewpoints.

However, it is crucial to recognize that this method has several limitations as well. It is time-consuming as it involves multiple study rounds (Barrett & Heale, 2020) and the preparation of questionnaires for each round. Furthermore, the choice of experts and the quality of the questions asked could impact the overall quality of the results (Esfandiari et al., 2023; Rowe & Wright, 2011).

3.5 Recruitment of the modified Delphi participants

The post-positivist paradigm significantly influences the expert panel selection process in this study, recognizing the importance of recruiting participants with direct knowledge and varied experiences (Barrett & Heale, 2020; Rowe & Wright, 2011). The panelists were selected based on their expertise and experience in dealing with research or publication data and their willingness to share that knowledge (Goodman, 1987; Ziglio, 1996), emphasizing purposive sampling over random selection. The study's population were data practitioners with diverse expertise individuals from various departments to get the best experience and knowledge of the organization's subject to ensure comprehensive coverage of RDG.

Purposive sampling, also referred to as judgment sampling, is a non-probability technique in which researchers intentionally select the panelists based on specific qualities or characteristics they possess (Etikan et al., 2016). According to Neuman (2011), this method is applied to choose unique cases that provide particularly informative insights. Etikan et al. (2016) further support Neuman's explanation, stating that purposive sampling involves the identification and selection of individuals or groups well-versed and knowledgeable about a phenomenon of interest. The deliberate use of purposive sampling in this study ensured that the panelists were selected from diverse institutions, departments, and research organizations, each possessing distinctive skills in handling research data. This approach offers the advantage of encountering subjects likely to have

expertise in the matter under investigation, aligning with Creswell and Creswell's (2018) assertion that purposive sampling is suitable when researchers believe they can obtain a representative sample through sound judgment.

Given the limited number of experts and challenges in recruiting panelists, snowball sampling was also utilized. This method involved selecting participants based on recommendations from current panel members, similar to approaches used by other Delphi study authors (De Lima & Seuring, 2023; Lim et al., 2020; Naranjo et al., 2023; Seuring et al., 2022). Snowball sampling is defined by Berg & Lune (2017) as a sampling approach where researchers reach out to a group of people and gathers other contacts through this group of people. This process was continued until the necessary sample size for the initial round was achieved. For the recruitment strategies, a contact book in an Excel spreadsheet file containing a list of data practitioners believed to meet the following general sampling criteria was prepared:

a) Participants acknowledge having practical expertise and knowledge encompassing diverse research data handling and support facets, including data generation, management, and consumption, AND/OR
b) Participants have previous experience or ongoing engagement in the formulation of (research) data policies for their affiliated institution(s), AND
c) Participants must be affiliated with any RPO in Malaysia. For those from government agencies, they should be associated with the MOSP, AND
d) Participants demonstrate a willingness to engage and share their valuable experiences actively.

Efforts to identify potential participants were initiated through online searches commencing from December 3, 2022. As a start, the MOSP played a crucial role in pinpointing data practitioners associated with the organization. This collective consisted of (MOSP, 2020):

a) Members of the Malaysia Open Science Alliance (MOSA) (16 members),
b) Members of the Working Group on Guidelines (10 members),
c) Members of the Working Group on Capacity Building and Awareness (7 members),
d) Members of the Working Group on Infrastructure (9 members),
e) Certified Data Stewards (12 members), and
f) Participants of the TOT on Data Stewardship for Open Science coming from diverse geographical areas and background (261 members) (MOSP, 2024).

Individuals not initially part of the TOT program could be identified through the MOSP website. However, contact with these TOT participants was established through a snowballing approach initiated by MOSP, as MOSP has the list of names and access to the contact information of these individuals.

Moreover, invitations were extended to individuals who had interacted with research or publication data from universities and research institutions. Another approach involved exploring digital platforms and repositories, such as Dimensions (https://www.dimensions.ai/), UM Dataverse (https://researchdata.um.edu.my/), and MOSP (https://mosp.gov.my/) itself, to identify researchers who had become acquainted with the principles of open data sharing. After being identified, a thorough online search was conducted for their email addresses. The potential participants were contacted via email and invited to participate in the study.

The recruitment of data practitioners occurred in a two-phase process:

a) The initial phase spanned from January 3, 2023 to March 31, 2023. Seventeen (17) individuals were willing to contribute to the study, including five from snowballing approach. The sample of email invitations for both phases are provided in Appendix A. They were then assigned to a specific set of instruments based on their responses regarding their level of involvement in governing and/or managing research/publication data, as indicated in question no. 4 of the demographic information form (referenced in Appendix B).

b) The second phase of recruitment occurred between April 4, 2023 and May 31, 2023. Potential participants were identified based on their current involvement, role, and position to determine their classification within the four main categories, as indicated in Table 3.3. The first three categories correspond to the levels of governance delineated by Korhonen et al. (2013).

Table 3.3: Categories of Potential Participants

No.	Group	Study Invitee	Description
1.	Set A (Strategic)	67	MOSA members, universities' deputy vice-chancellors, directors, deputy directors of research management offices, and chief librarians.
2.	Set B (Tactical)	27	Certified Data Stewards under MOSP.
3.	Set C (Operational)	51	Data practitioners who received TOT on Data Stewardship training, research officers, IT officers responsible for institutional repositories, liaison librarians and librarians who handle or deal with research or publication data.
4.	Set D (Researcher)	147	Researchers who have shared their datasets in various available repositories, including Dimensions, UM Dataverse, and MOSP.

3.6 Research instrument for the modified Delphi study[1]

During the first round of the modified Delphi study, the option of conducting structured interviews to extract insights from participants on RDG implementation was initially considered. However, it was decided that using thoughtfully crafted survey instruments with structured closed-ended questions would be more efficient and effective for initially gathering input from the panelists, as they allow for consistent data collection and easier analysis across a larger group (refer to Appendix C for a screenshot of the instrument on Cognito Forms). The panelists were provided with a set of questions in the same manner, this method served to reduce the possibility of bias or perspectives influencing the outcomes. Moreover, the instruments incorporated statements that had been derived from desk research involving policy documents from prominent RPOs worldwide. In the context of this study, all Delphi rounds mostly gathering quantitative data sought to reach a consensus not necessarily mean that it was a viable solution, nonetheless it showed the mutual agreement among the panelists on the RDG framework's acceptance and validation (Mustaffa et al., 2021).

All instruments were written in English and designed to be clear, concise, and focused on the research questions under investigation. A structured, closed-ended survey instrument was essential to ensuring that the obtained results were consistent and in line with the research objectives. Previous research has suggested that the number of anchors to be use in a survey should be primarily determined by the survey designer's judgement, while ensuring that respondents can adequately express their opinions (Chyung et al., 2017; Khodyakov et al., 2023). For this study, 5-point scale anchors were selected and

[1] All survey instruments employed in the modified Delphi study are available at: Hazmi, Norzelatun Rodhiah (2024). Research data governance framework for research performing organizations in Malaysia: Modified Delphi survey instruments. *Figshare*. Dataset. https://doi.org/10.6084/m9.figshare.24988116.v4.

employed as ordinal scales (Creswell & Guetterman, 2019) for participants to rank the level of implementation and importance of governance activities.

Survey instruments used for Round I were structured and customized based on the unique characteristics of each group, i.e. Set A, Set B, Set C, and Set D. The aim was to guarantee that individuals within each group could provide accurate responses that matched their expertise and knowledge. This was because, in the initial round, questions were tailored to gather information pertinent to the panelists' experience based on various levels of involvement in dealing with research data related activities. Additionally, the survey was organized into sets to restrict the number of questions, preventing an undue time commitment.

Delphi participants are usually experts with busy schedules and are unlikely to allocate an extensive amount of time for the survey. At the early stage of the initial round, three panelists answered for two sets of surveys. However, their feedback indicated that the survey took a longer time to complete, and furthermore, they were unable to pause due to the absence of a save button in the online form used. Acknowledging this time constraint among the panelists, each panelist was limited to providing responses for only one best set, even if they could answer more than one set. These sets of questions reflect a standard survey volume, i.e. the extent or amount of content contained within a survey instrument, indicating the overall size or number of questions participants are expected to answer. The survey instruments were designed with content that can be reasonably completed within an hour, ensuring comprehensive and insightful responses from the panelists. Completing the instrument for each round was estimated to require approximately 30 to 45 minutes.

The administration of the survey instruments in this study was carried out electronically, utilizing the secure online form builder Cognito Forms (https://www.cognitoforms.com). The panelists received a secure link via email, allowing them to conveniently complete the survey instrument online. Each round's questionnaire was accessible to respondents for approximately two weeks. The confidentiality of panelists' responses was maintained by summarizing their responses and not disclosing their identities to the other panel members. After all, Delphi research is distinguished by its emphasis on panel members' anonymity (De Lima & Seuring, 2023; Tangalakis et al., 2023).

3.7 Data collection for the modified Delphi study

Initially, the aim was to include over 30 participants into the survey to distinguish among the four sub-groups of data practitioners, ensuring a minimum of seven (7) participants per sub-group. Remarkably, the study exceeded expectations by recruiting a total of 47 data practitioners, including three (3) who completed two sets of questionnaires. Given the absence of a consensus on the requisite sample size for an adequate Delphi technique panel, as noted by Lim et al. (2020), the current study decided to include all 47 individuals. This marked the formal completion of the modified Delphi study: Round I. Therefore, the process of selecting the panelists aligns with the post-positivist paradigm, which underscores the importance of empirical evidence and the rigorous testing of hypotheses and theories based on the expertise and experiences of professionals in the field.

The panelists' voluntary completion of the survey served as an implicit agreement to participate in the study, as conducted by Petersen et al. (2019). Several efforts were made to retain these panelists so that they could participate in the subsequent rounds, as suggested by Hall et al. (2018). These efforts involved providing aggregated, anonymous

feedback from prior rounds to allow participants to reflect on others' responses (sample is attached as Appendix D), issuing reminders emphasizing the value of their involvement (refer to Appendix E), ensuring rounds concise and minimal intervals between rounds to sustain engagement, and personally acknowledging those who completed all rounds through tokens of appreciation (see Appendix F).

The four-round modified Delphi study spanned approximately ten months, commencing with Internet searches for potential participants on December 3, 2022. The study included the creation and administration of the Delphi instrument, as well as the subsequent processes of data collection and analysis, concluding with the final data collection on September 30, 2023. Analysis and reporting on prior modified Delphi rounds were prepared and presented to the panelists before each of the subsequent rounds (attached as Appendix G). The panelists were given an average of two weeks' time to respond to each of the four modified Delphi rounds with reminder emails sent to the panelists who did not respond. The panelists were asked to respond to all instrument items in each round. Participation in the survey was voluntary, and no effort was made to provide information to survey participants prior to the instructions in the email and the instrument. Other than that, the instrument was supplemented with a glossary of terms in the introductory Round I and Round III. This was done to ensure consistency among the panelists in using the same definitions for crucial terms. Interpretations of loosely defined terms will be shaped by perceptions of the participants, who have different educational, professional and cultural experiences (Van Mil & Henman, 2016).

Cognito Forms and Microsoft Excel software were opted to design, collect, and analyze the survey data. Additionally, Yet Another Mail Merge (YAMM) (https://yamm.com) was employed to send personalized emails to panelists. When compared to the blind carbon copy (BCC) feature available in Gmail, YAMM proved invaluable by sending default emails that include customization of recipient names, personalized links to different sets of survey instruments, and settings for various due dates for survey completion (refer to Appendix H). In addition, YAMM facilitated real-time tracking of engagement metrics directly from Google sheets, allowing for monitoring of email activity, including reading emails and clicking links. The use of YAMM significantly eased and expedited the email communication process by sending an email to many participants, a task that would have otherwise been challenging and time-consuming.

To ensure a cohesive and continuous engagement throughout the study, participation in the initial round was a prerequisite for the panelists to progress to Rounds II and III. Similarly, participation in Round III was a prerequisite for the panelists to advance to Round IV. Delphi Round I was carried out to tackle RQ1. In contrast, three Delphi rounds—Rounds I, II, and III—were utilized to address RQ2. RQ3 was answered in Rounds I-II, while RQ4 was addressed in Rounds III-IV. As a result, insights from all four rounds (Rounds I-IV) hold significant importance in guiding the development of an RDG framework tailored to meet the specific requirements of RPOs in Malaysia. This offers recommendations for improved RDG implementation and adherence to best practices. Detailed discussions on each modified Delphi round are presented in the subsequent sections. Figure 3.3 depicts the entire modified Delphi process.

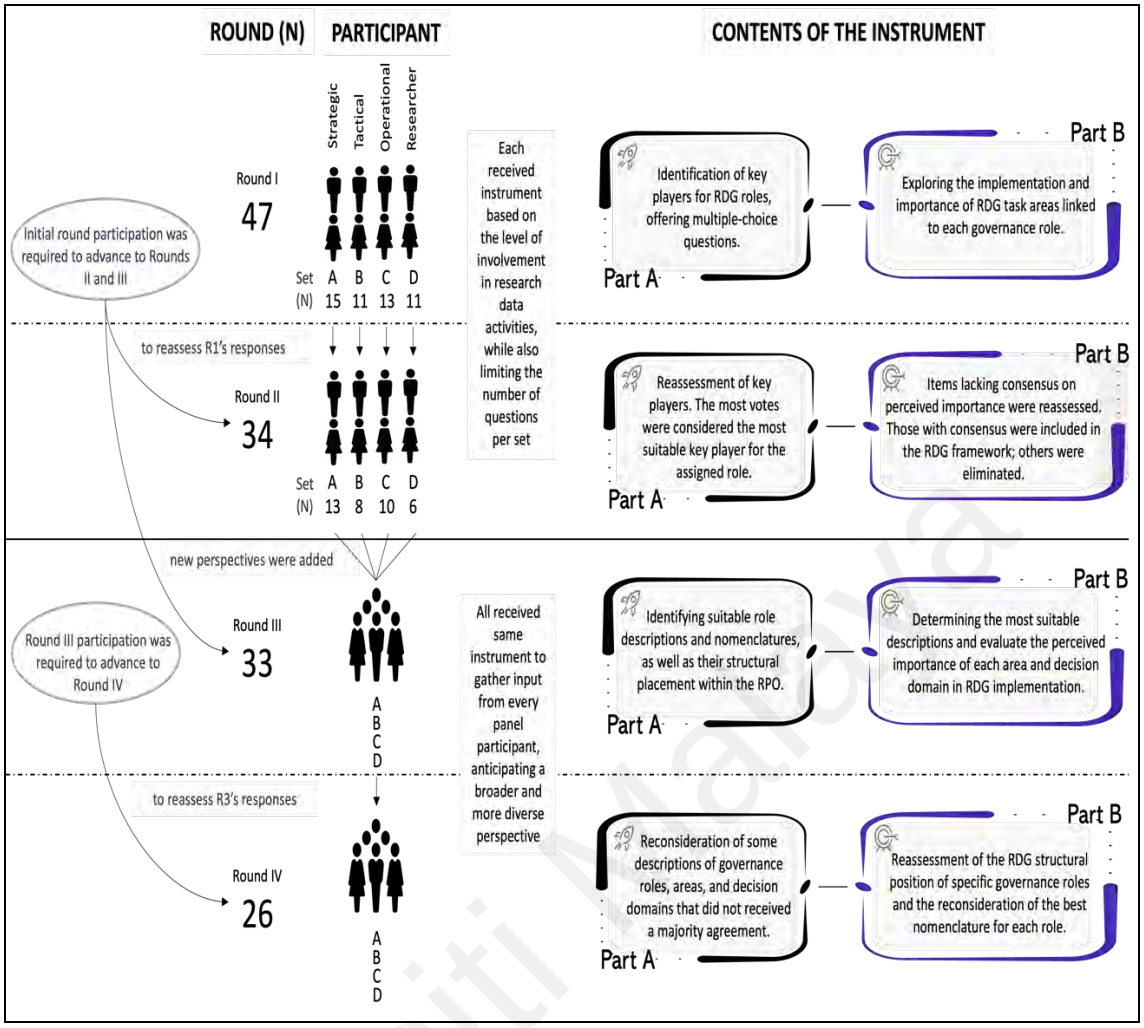


Figure 3.3: A comprehensive overview of the modified Delphi process — tackling research questions and guiding RDG framework development

3.7.1 Modified Delphi: Round I

The study utilized four survey instruments by aligning them with the panelists' assigned group. The survey instrument was emailed to 292 data practitioners (please see Appendix I for a sample email). In Round I, the instrument comprised two primary constructs: Part A and Part B.

a) Part A encompassed questions concerning the identification of key players for RDG roles, offering multiple-choice questions to allow the panelists to select one or more applicable answers. In total, Part A included eight (8 items). All items were automatically presented again in Round II to gather a majority consensus from the group responses. This aspect addressed RQ3.

b) Part B consisted of more specific statements designed to investigate the actual implementation and perceived importance of RDG task areas linked to each governance role. This aspect was pertinent to addressing RQ1 and RQ2, respectively. In Part B, 5-point scale questions were employed for the panelists to choose the most suitable response. The response options for the actual implementation level ranged from 1 = Not Implemented to 5 = Very Highly Implemented. This scale was also employed by Eisma (2015) for students to assess the degree of implementation of Student Services Programs within both the College and the University in the College of Arts and Sciences. Simultaneously, the response options for the perceived importance level ranged from 1 = Not Important to 5 = Very Highly Important. Notably, there was no neutral midpoint to force the panelists to take a stance of either between not implemented/not important to very highly implemented/very highly important. The panelists were required to indicate the extent to which each task was implemented within their organization. Furthermore, they were instructed to express the degree of importance of each task to the implementation of RDG in Malaysia RPOs. Each item was provided with free-text comments that they could explain their response or express disagreement with the statement's relevance. In total, Part B was divided into eleven subsections corresponding to various governance roles, such as the Organization (18 items), Executive Sponsor (2 items), Data Governance Leader (7 items), Research Data Governance Committee (7 items), Office of Research Data Governance (4 items), Research Data Governor (13 items), Research Data Steward (17 items), Administrative Offices (17 items), Research Data Consumer (3 items), Researchers (24 items), and External Bodies (7 items), resulting in a comprehensive total of 127 item statements.

While different sets (A, B, C, and D) included various governance roles for Part B, the questions for Part A on identifying RDG key players and Part B concerning the tasks associated to the organization were presented in all sets. This brought the total to 46 items for Set A, 56 items for Set B, 53 items for Set C, and 57 items for Set D. The rationale for categorizing the panelists into distinct groups, each responding to a customized set of instruments according to their level of governance and profession, was to improve the accuracy and reliability of the responses. This categorization also aimed to minimize the number of questions in each set, ensuring that the estimated time required to answer all questions was 45 minutes.

3.7.2 Modified Delphi: Round II

In Round II, questions were categorized into two parts:

a) Part A, consisting of all eight items from the first round. The questions were presented again, and the findings from the previous round were shared with all the panelists from all sets. This was done to identify key players occupying various governance roles. The stakeholder who received the majority of votes after this round was considered the most suitable key player for the assigned role, thereby addressing RQ3.

b) Part B was to build expert consensus on the perceived importance of governance tasks for RPOs in Malaysia, which did not meet the consensus criteria in the first round. Part B was particularly conducted to further answer RQ2 and RQ3 as these research questions required achieving a certain a priori consensus threshold. Each panelist in Round II received an email invitation with a unique link (see Appendix J), consisting a personalized questionnaire that included their individual rating as well as the overall group median rating for each item; this feature allowed the panelist to see the group collective

opinion (Hasson et al., 2000). A section of free-text responses was included to represent the diversity of opinions. In Part B, using the 5-point scale questions, the same statements as in the previous round (items that did not reach the identified a priori consensus threshold) were presented. The panelists were instructed to specify the level of perceived importance of each listed statement. Therefore, Round II served as a qualifying round, determining whether items that had not reached a consensus level in Round I would reach a consensus this time or remain below the a priori consensus threshold. Items falling into the former category were considered significant tasks and would be included in the RDG framework. However, for RDG tasks that fell into the latter category were deleted and eliminated.

The number of items and statements depended on their designated set. The questions on Data Governance Leader (6 items), Research Data Governance Committee (7 items), Office of Research Data Governance (4 items), Research Data Governor (5 items), Research Data Steward (4 items), Administrative Offices (7 items), Research Data Consumer (1 items), Researchers (3 items), and External Bodies (4 items). This brought the total to 25 items for Set A, 17 items for Set B, 20 items for Set C, and 15 items for Set D, resulting in a comprehensive total of 41 item statements. Table 3.4 provides a breakdown of the questions across each set in Rounds I and II.

3.7.3 Modified Delphi: Round III

Rounds III and IV were conducted to address RQ2, and RQ4. Every panelist received an email invitation to participate and was granted access to the same instrument for everyone, specifically crafted for this round (refer to Appendix K). The approach shifted away from grouping the panelists into specific sets. The reasoning behind this adjustment was the belief that having all participants respond to the same survey instrument would

be more beneficial. The goal was to gather input from every panelist, anticipating a broader and more diverse perspective. This round aimed to establish a consensus regarding another RDG perspectives, i.e. governance areas and decision domains (RQ2), and structural positions (RQ4). This information was essential as the foundational basis for crafting an RDG framework tailored to RPOs in Malaysia. This framework is intended to facilitate the effective initiation and execution of RDM practices within RPOs.

Table 3.4: Breakdown of questions for each set in Rounds I and II

Construct	Round I				Round II			
	Set A (#)	**Set B (#)**	**Set C (#)**	**Set D (#)**	**Set A (#)**	**Set B (#)**	**Set C (#)**	**Set D (#)**
Part A: Key Players for RDG Roles	8	8	8	8	8	8	8	8
Part B: The Organization	18	18	18	18	-	-	-	-
Part B: Executive Sponsor	2	-	-	-	-	-	-	-
Part B: Data Governance Leader	7	-	-	-	6	-	-	-
Part B: Research Data Governance Committee	7	-	-	-	7	-	-	-
Part B: Office of Research Data Governance	4	-	-	-	4	-	-	-
Part B: Research Data Governor	-	13	-	-	-	5	-	-
Part B: Research Data Steward	-	17	-	-	-	4	-	-
Part B: Administrative Offices	-	-	17	-	-	-	7	-
Part B: Research Data Consumer	-	-	3	-	-	-	1	-
Part B: Researcher	-	-	-	24	-	-	-	3
Part B: External Bodies	-	-	7	7	-	-	4	4
Total	**46**	**56**	**53**	**57**	**25**	**17**	**20**	**15**

In Round III, Part A featured 31 items that addressed twelve distinct governance roles, including new roles considered essential for the RDG framework. These roles were the Research Data Owner, proposed by Abraham et al. (2019), and the Research Data Custodian, recommended by Norbib & Abu Bakar (2021). Recognizing the significance of standardized definitions for governance roles, the panelists were required to determine whether the provided description for each RDG role was acceptable or required revision. Another Delphi study that incorporated definitions into their investigation was conducted by Aronson et al., (2012), aiming to establish a consensus definition of professional

engagement. Subsequently, the panelists were instructed to select the most appropriate option from the provided choices to denote the level of governance associated with each role and the best suited nomenclature for those roles. The primary goal was to ascertain the most appropriate descriptions and nomenclatures for these roles, as well as to gain a deeper understanding of their structural placement within RPOs.

Meanwhile, Part B consisted of 38 items encompass nineteen (19) unique RDG areas and 16 items distributed across eight (8) unique RDG decision domains. The objective of Part B was to determine the most suitable descriptions and evaluate the level of importance of each item within the context of these RDG aspects. Response options for the level of perceived importance of each item ranged from 1 = Not At All Important to 5 = Extremely Important. The panelists were required to specify the degree of importance for each area and decision domain in RDG implementation. Furthermore, each item was accompanied by a section for participants to provide free-text comments. This feature allowed them to expand on their responses, offer explanations, or express any disagreements related to the statement's relevance.

3.7.4 Modified Delphi: Round IV

Initially, the plan was to carry out a three-round Delphi study. Nevertheless, it became evident that this approach did not provide ample opportunities for the panelists to achieve a consensus, especially given the introduction of new information in Round III. The inclusion of a fourth round of responses, in which the panelists could compare their ratings to the aggregated responses from their fellow panel members, might have allowed each panelist to contemplate and potentially revise their responses in the previous round (Keeney et al., 2006).

Email invitations for participation in this final round were exclusively sent to panelists who had taken part in Round III (refer to Appendix L). Each panelist received a personalized link, incorporating both their individual original responses alongside the group responses. They also had the opportunity to offer additional comments on the statements if they want to. Round IV covered statements regarding descriptions of some RDG elements related to the roles, areas and decision domains, as well as structural position (RQ4), and nomenclature for each governance role for which a majority acceptance had not been reached in the previous round. This final round provided the panelists with an opportunity to refine their responses.

The instrument was divided into two parts. Part A featured 5 items, in which the panelists had the option to either retain the original description of governance roles, areas, and decision domains or use a revised description of each item based on suggestions provided by other members in Round III. On the other hand, Part B comprised 20 items related to the reassessment of the RDG structural position of specific governance roles within RPO and the reconsideration of the best nomenclature for each role. For a detailed breakdown of questions in Rounds III and IV, please refer to Table 3.5.

Table 3.5: Breakdown of questions for Rounds III and IV

Construct	**Round III (#)**
Part A: RDG Structural Dimension	31
Part B: RDG Functional Dimension	54
Total	**85**
Construct	**Round IV (#)**
Part A: Description of RDG Key Elements	5
Part B: RDG Structural Dimension	20
Total	**25**

The final round was aimed at achieving a majority consensus on each item listed and served as a qualifying round. The process, akin to Round II, was repeated in Round IV, affording the panelists one more opportunity to re-evaluate their responses from Round III and decide whether to retain or revise their answers.

3.8 Data Analysis of the modified Delphi study

This modified Delphi study primarily utilized descriptive and inferential statistics in its analysis. The analysis includes percentages, interquartile range, mean and median scores, as well as mean differences, t-values, and significance levels (Sig.) obtained from paired-sample t-tests. Microsoft Excel and RStudio (R Foundation, 2023) were employed for dataset examination. Excel spreadsheet software has an advantage over other statistical software due to its wide availability and relatively user-friendly interface (Gasigwa et al., 2022). Screenshots of Excel sheets utilized for data analysis are included in Appendix M. On the other hand, RStudio offers flexibility and is capable of generating high-quality analyses (Rajoo et al., 2020), which utilized for obtaining inferential statistics during data analysis (please consult Appendix N for a sample view of the source and console pages in RStudio).

In this modified Delphi study, the identification of consensus occurs in three situations:

a) Multiple-choice questions were used to achieve a majority consensus from group responses, aiding in the identification of stakeholders for specific roles and their structural positions, along with the formulation of appropriate nomenclatures. The process also involved delineating descriptions of role, governance areas, and decision domains.

b) A 5-point scale was employed to assess the level of implementation of each RDG task by Malaysia RPOs. The interpretation of the scale's results was derived from mean scores, following the methodology established by Alfaras (2023), as outlined in Table 3.6.

Table 3.6: Interpretation of Scale Results based on Mean Scores

Mean score	Interpretation
1.0-1.49	Not implemented
1.50-2.49	Slightly implemented
2.50-3.49	Moderately implemented
3.50-4.49	Highly Implemented
4.50-5.0	Very highly implemented

c) A 5-point scale was used to assess the level of perceived importance of the RDG activities; tasks, areas, and decision domains, for Malaysia RPOs. A priori consensus thresholds were determined by an interquartile deviation (IQD) of ≤ 0.5, a consensus level (CL) of $\geq 85\%$, and a median of ≥ 4. Items that did not meet these criteria for group consensus were excluded. This stringent consensus requirement was intended to develop an evidence based RDG framework that would achieve widespread acceptance among data practitioners and RPOs in Malaysia.

Besides that, to examine the comparison between the actual implementation of RDG tasks and the perceived importance of each item, a gap analysis was conducted, with a paired-sample t-test p-value ≤ 0.05 indicating significance.

3.9 Reliability, Validity, and Trustworthiness

Several strategies were employed to ensure the content validity of the study, the instruments and the findings. Initially, the research area or topic was comprehensively defined. The introduction and literature review provided a foundation by evaluating pertinent studies within the topic area. This process facilitated a thorough understanding of the study area, aiding in identifying gaps and concepts crucial for data collection aligned with the research questions.

Subsequently, the instruments of the Delphi study were crafted based on secondary data from the desk research of policy documents produced by RPOs worldwide. Previous authors have highlighted that desk research is often used to gather background information or to validate primary research findings (Ashiabi & Avea, 2019; Dijkxhoorn et al., 2019; Drossaert et al., 2018; Dwiatmoko et al., 2018; Zhang, 2021).

Additionally, to ensure the data's reliability, emails were sent to the executives, policy stewards, or individuals leading the RDM initiative at the respective institutions, requesting permission to use the document on their website and confirming that it was the latest version. The aim was to collect relevant policy documents from diverse RPOs to gain a comprehensive understanding of the contexts that informed the study, as suggested by White & Marsh (2006), that the sampling should be theoretical and purposive.

Furthermore, a rigorous analysis of the documents was conducted by coding the texts. After the coding process was completed, a collaborative review process was initiated, involving joint reviews of the codes with subject experts. Interpretations were discussed to ensure the accuracy and consistency of the findings.

Meanwhile, pilot testing was conducted in Round I for the Delphi survey instruments to enhance the reliability and validity of data collection procedures and results (refer to Appendix O). Pilot testing has been recognized as a pivotal research approach. Lancaster et al. (2004) defined pilot testing as a small-scale preliminary study conducted to evaluate an instrument and enhance its efficiency before the full-scale investigation. The instruments' design for this study underwent a testing phase involving a small group of eight (8) data practitioners selected from a pool of 47 panelists across all levels of governance. This pilot testing was conducted to identify issues related to question meanings, wordings, structures, and sequences. Several adjustments and refinements were made to the instruments in response to feedback from the panelists.

Additionally, this testing phase helped evaluate the comprehensibility of response categories and determine the average time required for completion (Check & Schutt, 2012). The insights gained from this testing phase were instrumental in refining the instruments. The refined instruments were aimed to enhance participants' engagement while ensuring the validity and accuracy of their responses (Check and Schutt, 2012). This test included the verification of the content validity through feedback from the eight participants representing both universities and research institutes. Based on the pilot test results and feedback from the expert panels, the questionnaires were modified, and then the online survey was conducted using the finalized questionnaires.

Furthermore, the panelists had the opportunity to share their perspectives on each statement presented in Round I, with new additions being reassessed in subsequent rounds. Similarly, in Round III, they could add more options for each statement, which were then re-rated by all members in Round IV. Their responses were recorded and analyzed to enhance the reliability and validity of the research findings.

Before proceeding with the pilot testing, the questionnaires underwent a thorough proofreading process to identify any grammatical errors or double-barreled statements (see Appendix P). It is worth noting that the Delphi instruments for each round were also evaluated and assessed by the supervisors to ensure they effectively measured the intended constructs before disseminating them to the panelists.

The study adopted a triangulation approach, combining data from various sources and methods to strengthen the credibility of its findings. It started with a thorough desktop research phase, where existing literature and data were reviewed to establish a solid foundation on the study's main topics. Following this initial phase, multiple rounds of Delphi surveys were conducted. These iterative rounds engaged a panel of experts who provided structured responses and feedback through questionnaires. By triangulating

insights gathered from these different phases, the study aimed to comprehensively address its research questions. This approach ensured a thorough exploration of the topic from diverse perspectives and methodological approaches.

3.10 Ethical consideration

Ethical considerations, especially with human subjects were essential (Creswell, 2012), ensuring integrity in the study and selecting the most appropriate approach (Busher & James, 2007). The study was conducted in accordance with the three fundamental ethical principles outlined in the Belmont Report (OHRP, 1979): (a) respect for persons, (b) ensuring personal safety, and (c) conducting the Delphi rounds fairly.

The researcher assumed responsibility, demonstrated integrity, and recognized the importance of protecting the confidentiality rights of all panelists (Creswell & Creswell, 2018). Participant anonymity was preserved, and their responses were treated as confidential information (Merriam & Tisdell, 2015). It is important to note that this study did not include any "off the record" comments. A neutral position was maintained throughout the study, refraining from sharing any personal experiences, comments, or judgments to ensure content validity (Goodman, 1987).

The design of the Delphi study aim to minimize potential biases or conflicts of interest among the panelists (Keeney et al., 2006). Therefore, panelists were selected in a fair and equitable manner, without any discrimination based on gender, race, ethnicity, religion, or any other personal characteristic (Hasson et al., 2000). They were informed of the study's objectives, procedure, confidentiality, and needed time to complete the surveys. In addition, they were also provided with a consent form, a demographic information form and a document outlining the experts' criteria (McMillan, 2012), as evidenced in Appendix Q. It was important for the panelists to understand the study and remain engaged throughout the study to increase its validity (Creswell & Creswell, 2018; Hall et

al., 2018). A summary of the study findings shared with the participants, offering them insights into how their contributions contributed to the overall results. This debriefing helped close the loop on the research process. Additionally, the participants received summaries of aggregated responses after each round, allowing them to reflect on the collective opinions while maintaining the anonymity of individual response.

Steps were taken to ensure research integrity and the protection of participants in the current study. A thorough application describing the current study was submitted to the Universiti Malaya Research Ethics Committee (UMREC), completing all required research submission procedures by December 21, 2022. Upon review, UMREC determined that the study protocols posed no apparent risk to participants. All participants were provided with information about the study's nature, and invitations were extended to modified Delphi participants via email. The initiation of recruiting data practitioners to participate as panelists in this modified Delphi study commenced after obtaining approval UMREC, with the assigned reference number: UM.TNC2/UMREC_2372. The UMREC approval letter pertaining to the current study is represented in Appendix R.

3.11 Summary of Chapter Three

This chapter elaborated on the research paradigm and research design, leading to the strategic selection of research methods. It outlined two key phases: desk research and the modified Delphi study, shedding light on the rationale behind choosing the modified Delphi approach. The chapter covered aspects such as the recruitment of participants, outlining the research instrument employed, the process of data collection, and the subsequent analysis for the modified Delphi study. Additionally, it discussed the measures taken to ensure the reliability and validity of the study. Ethical considerations throughout the research process were also highlighted, emphasizing the ethical framework guiding the research activities. Meanwhile, the following Chapter Four delves

into the methodology and findings of the first phase, laying the groundwork for understanding the RDG activities and the development of the Delphi instruments. It offers a detailed explanation of data collection, analysis, and results from the desk research, helping the reader comprehend the content and recognize the strong link between these two phases in addressing the study's research questions.

CHAPTER 4: DESK RESEARCH ANALYSIS EXPLORING RESEARCH DATA GOVERNANCE ACTIVITIES[2]

"Indeed, Allah commands you to render trusts to whom they are due..."
(Surah An-Nisa, 4:58)

4.1 Introduction

In this chapter, the discussion synthesizes the secondary research conducted, relying on a key data source: desk research utilizing a content analysis of RDG/RDM policy documents, which generated qualitative data. The desk research commenced in July 2022 and ended in September 2022. The findings from the desk research were employed to form an initial understanding of RDG practices implemented among leading RPOs worldwide and to gather background information on the topics under consideration. This groundwork enabled the creation of the modified Delphi instruments designed to enhance the understanding of RDG implementation within Malaysia RPOs.

4.2 Selection of policy documents

Purposive sampling was used to choose the sample for the desk research, focusing on existing RDG/RDM policies. Each policy document was sourced from RPO listed in the Scimago Institutions Rankings filtered by Research Rank (Scopus, 2022). Initially, the Leiden Ranking, developed by the Centre for Science and Technology Studies (CWTS) at Leiden University in the Netherlands, was also considered for use in this study. However, it primarily offers comprehensive information on universities' scientific performance based on bibliographic data from publications indexed in the Web of Science (CWTS, 2022), with no inclusion of research institutions (Lancho-Barrantes & Cantu-Ortiz, 2021). Therefore, the Scimago Institutions Rankings was chosen for this study

[2] Several sections of this chapter have been previously published as a preprint: Hazmi, N.R., Abrizah, A. and Yanti Idaya, A.M.K. (2023). Analyzing policy documents: A desk study exploring research data governance practices among leading research performing organizations. *Advance*. Preprint. Available at: https://doi.org/10.31124/advance.24516061.

above other well-known rankings because it measured research activity at universities and research institutes worldwide. Since there were 8084 institutions listed, the list was filtered to only the top 350 ranked institutions. The selected policy documents must be easily accessible, downloadable, and written in English. Although the sample was not restricted to any specific publication date, all selected institutions must have a publicly available document version. Meanwhile, Atlas.ti was used as an annotation, searching aid and analysis tool in the study. Table 4.1 outlines the inclusion and exclusion criteria for sampling purposes.

Table 4.1: Inclusion and exclusion criteria of the sample

No	Inclusion	Exclusion
1	The document must be produced by an RPO.	The document is produced by a non-RPO.
2	The RPO must be one of the top 350 ranked institutions in Scimago Institutions Rankings by Research Rank.	The RPO is not listed in the top 350 ranked institutions in Scimago Institutions Rankings by Research Rank.
3	The document is on RDG/RDM policy.	The document is not related to RDG/RDM policy.
4	The document must be easily accessible online and downloadable.	The document is not easily accessible online or downloadable.
5	The sample is restricted to documents in English.	The document is not in English.
6	The sample is not restricted to documents on any specific publication date.	Not applicable.

The online tracking of the policy documents was made in May 2022. The terms used for searching the web were ‘research data governance’ OR ‘data governance’ OR ‘research data management policy’ OR ‘data policy’. After considering all the inclusion criteria, 36 out of 56 identified documents (64%) were suitable for use, including 34 RDM policies and 2 RDG frameworks from 34 universities and 2 research institutes. The list of the policy documents is attached as Appendix S and a diagram on the document sampling process is depicted in Figure 4.1. The purpose of the desk research was to gather more background information about data governance elements by identifying key governance roles and the numerous RDG activities they implement at various leading educational and research institutions around the world, as determined by their administration boards. It

was done to guide the development of instruments for the modified Delphi study in the next phase. Overall, the purposive sampling method has allowed for the selection a sample that met the criteria needed for the research, and the Scimago Institutions Rankings provided a reliable source for identifying suitable RPOs.

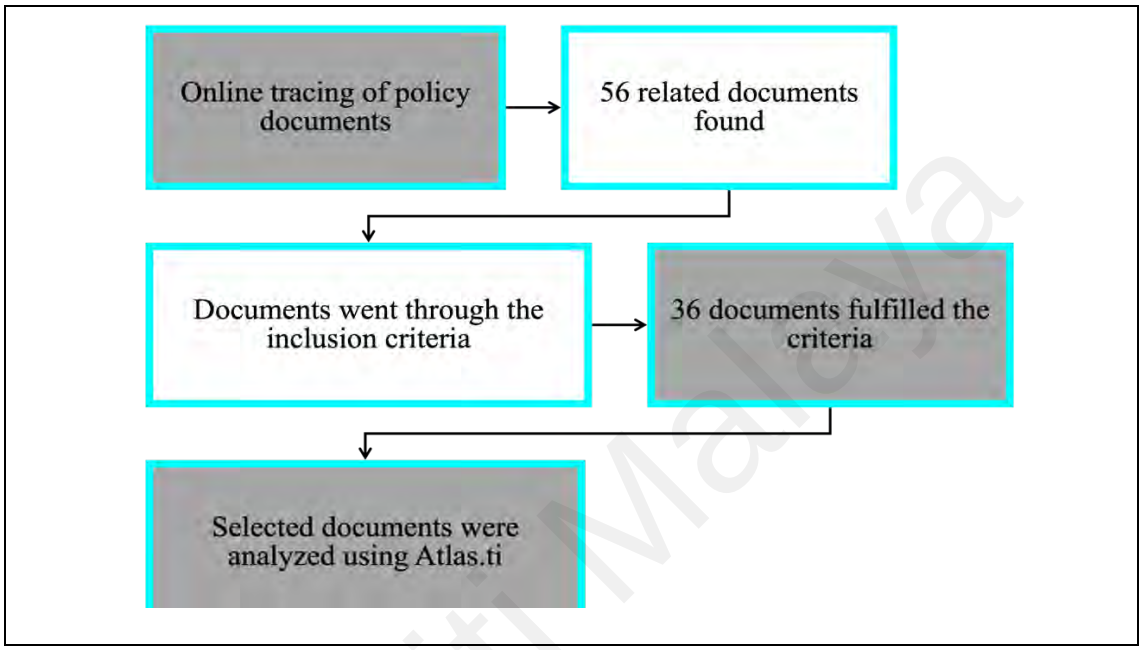


Figure 4.1: The sampling process of documents for desk research

4.3 Data Collection Procedure and Data Analysis

Conducting desk research was part of the initiative to identify the primary RDG roles and tasks, enabling an exploration of the varied governance activities implemented by RPOs across the globe. To accomplish this, the desk research used a content analysis, which involved selecting a concept for examination (White & Marsh, 2006). The analysis focused on examining the existence of chosen terms within the documents, which may be found as synonyms or expressed using different words (Carley, 1993; White & Marsh, 2006). The entire text was first read in-depth to identify related excerpts within the documents and evaluate their characteristics (White & Marsh, 2006). To systematically code the texts, the process involved identifying “structures and structured regularities in the text” (Myers, 1997, p. 13) and then applying a code to extract meaning from the text.

Following the methodology adopted by Alhassan et al. (2018), eight coding steps were employed to conduct the content analysis of the policy documents. These procedures, which involved data collection and coding, contributed to the clarity and openness of the research processes. The eight steps are described in Figure 4.2.

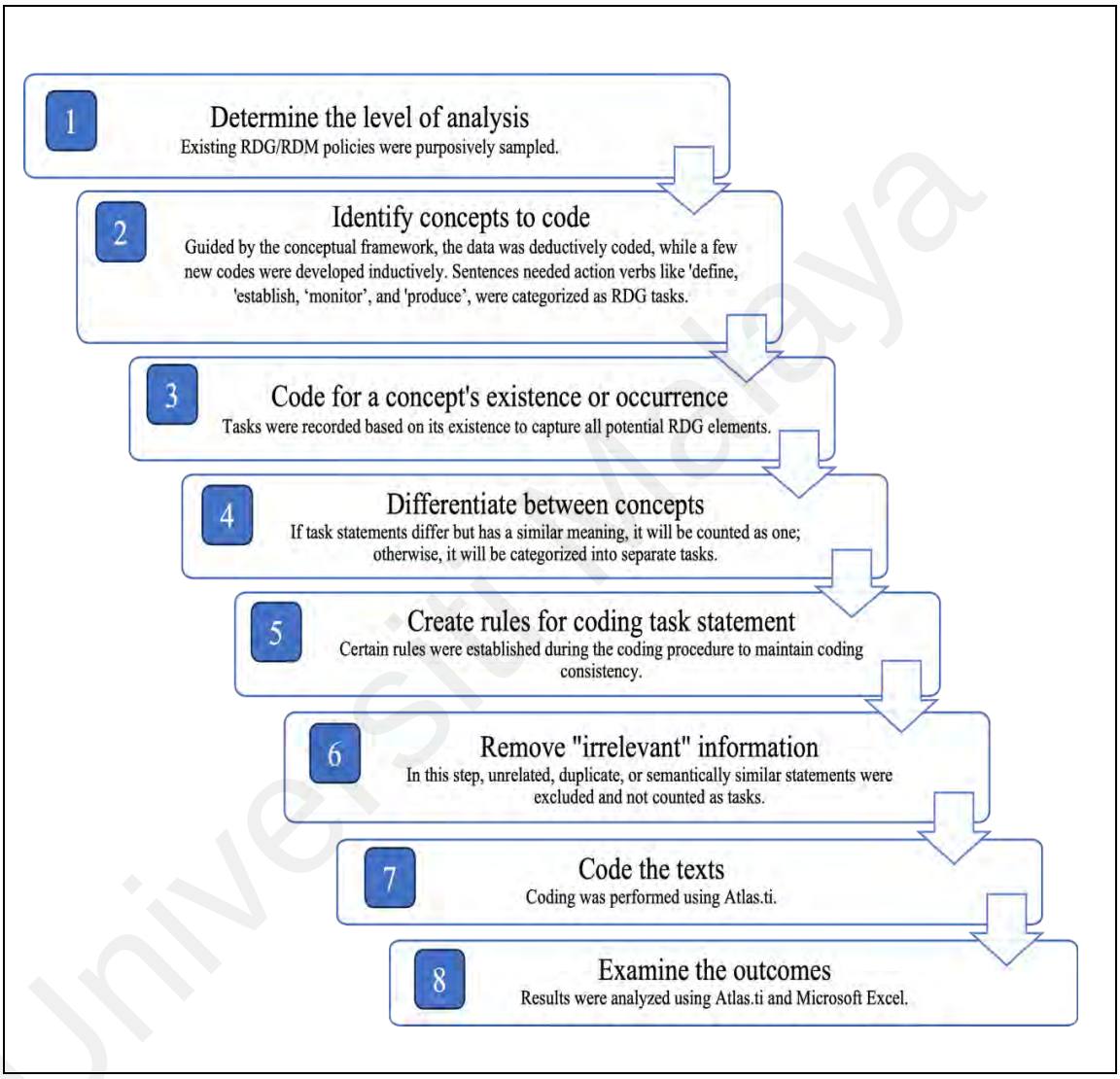


Figure 4.2: Coding steps for content analysis of selected policy documents

Step 1 – Determining the level of analysis. This step involved determining the level of analysis, which involved selecting the unit of language to be examined, such as a single word, a set of words, phrases, or an entire document (Carley, 1993). This content analysis examined 36 RDG/RDM policies from prominent RPOs that have integrated RDG into their institutions, aiming to delve into the RDG tasks and the roles associated with them.

The analysis primarily targeted statements or sentences concerning the essential tasks in RDG. Additionally, the associated role for each task was also recorded.

Step 2 – Identifying concepts to code. To ensure a structured and rigorous approach, directed content analysis techniques described by Hsieh and Shannon (2005) was employed, in which an initial coding scheme was developed based on an existing theoretical or conceptual model, as conducted in the previous research (Potter & Levine-Donnerstein, 1999; Vreugdenhil et al., 2022). This deductive approach allowed for identifying critical categories. The initial coding scheme for this content analysis was structured according to the conceptual framework illustrated in Figure 3.2 from Chapter Three. Throughout the analytical process, new codes for RDG areas, such as data citation and data custodianship, were inductively developed while reviewing the texts. All 36 policy documents underwent multiple thorough readings to identify recommended RDG key roles with their tasks comprehensively. A sentence must contain an imperative verb that indicates a specific action, such as “define”, “set”, “approve”, or “provide” to qualify for coding. These actions signify the conditions or tasks that must be implemented to satisfy the requirements for RDG.

Step 3 – Coding for a concept’s existence or occurrence. To determine the most appropriate coding strategy, researchers should consider whether to code concepts based on mere existence or frequency after enough concepts have emerged (Alhassan et al., 2018). In this content analysis, the existence of a task was prioritized over its occurrence. This approach facilitated the identification of all potential RDG tasks performed by various key roles within the selected RPOs, providing a deeper understanding of the governance practices.

Step 4 – Differentiating between concepts. During this step, researchers should decide whether to code the concepts verbatim or in a modified or compressed version (Alhassan et al., 2018). In this content analysis, task statements containing imperative verbs, as outlined in Step 2, were singled out. These statements underwent both explicit and intrinsic comparison to cluster them according to their semantic similarities. Similar-meaning statements or excerpts were assembled to form a distinct task statement, while those conveying different meanings were categorized separately. This process aimed to prevent any duplication of tasks across roles.

Step 5 – Creating rules for coding texts. The following rules were established during the coding procedure to maintain coding consistency (Alhassan et al., 2018):

a) Policy documents were initially read to identify RDG key roles with their associated tasks with imperative verbs indicating an action to be taken; and
b) Emerged task statements were compared to identify similarities and differences and grouped under relevant tasks.

Step 6 – Removing "irrelevant" information. In this content analysis, any information not directly contributing to the study's objectives, or duplicate task statements with similar meaning, were excluded following Carley's (1993) recommendation to streamline the analysis by discarding irrelevant data.

Step 7 – Coding the text. After establishing the appropriate approach for handling irrelevant information, researchers must follow the translation rules outlined in Step 5 to initiate the coding process (Alhassan et al., 2018). In this content analysis, systematic coding procedures were utilized using Atlas.ti 9.1.3, a qualitative analysis software system. The encoded data were subsequently imported into a Microsoft Excel spreadsheet for further analysis. Figure 4.3 provides a detailed illustration of the research model,

outlining the RDG tasks, roles, and activity dimensions according to the framework of the Data Governance Activities Model.

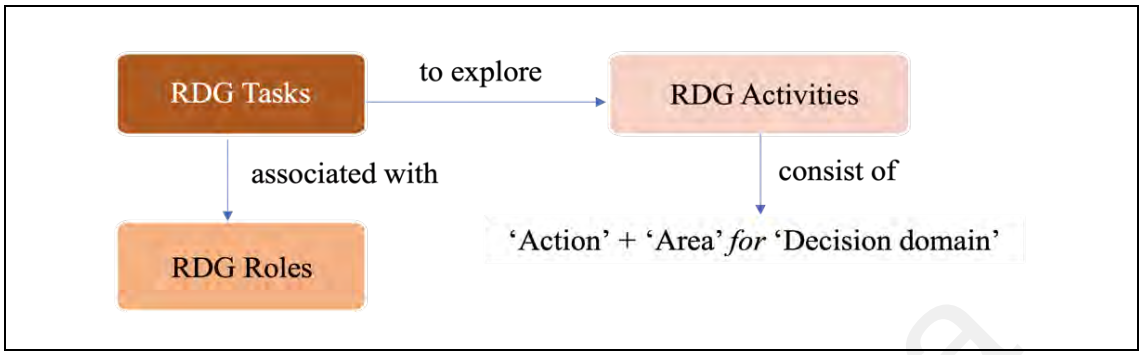


Figure 4.3: A research model of the RDG tasks, roles, and activity dimensions

These task statements linked with its associated roles were classified into broader categories, referred to as RDG activities. For this classification, it was established that each category would involve three essential components: (i) the action, (ii) the governance area, and (iii) the decision domain, following the framework of the Data Governance Activities Model. To determine whether the action was a 'define', 'implement', or 'monitor' task, it was essential to always consider the context-based connection between the task and the role. Meanwhile the area of governance was identified based on the task itself. Once this area was determined, the decision domain could be easily linked, as the relationship between the area and decision domain was predefined. It is important to highlight that the classification process was conducted in a straightforward manner, relying on subjective judgment and interpretation. To ensure accuracy and reliability, the data underwent evaluation and validation by two individuals:

a) a certified data steward specializing in RDM, ensuring a thorough understanding of data governance principles and practices relevant to the research context; and

b) a researcher who actively advocates for open science initiatives, reflecting a commitment to transparency, reproducibility, and collaboration, aligning closely with the principles of good data governance.

Step 8 – Examining the outcomes. The findings from the desk research were crucial and significant, as they provided guidance for the development of survey instruments during the second phase of the study, particularly for Rounds I and III of the modified Delphi study.

4.4 Results

After completing the coding and analysis processes, this section outlines the findings from the desk research.

4.4.1 Descriptive findings of policy documents

In analyzing the policy documents related to RDG, it is important to consider the context and scope of the study. RDG/RDM policy documents were collected from 36 RPOs, including universities and research institutions worldwide, as depicted in Figure 4.4.

Emails were sent to these RPOs requesting consent for the use of the policy documents (see Appendix T). However, many responded affirmatively, stating that permission had been granted. Additionally, some mentioned that if not requested, permission could still be obtained, as the policy documents fall under fair use. Notably, the United Kingdom had the highest number of RDG-related policy documents, accounting for 47.2 percent (n=17) of the total sample. The United States is the second highest contributor, contributing 13.9 percent (n=5) of the sample. Then followed by Australia and Netherlands, each accounting for 11.1 percent (n=4) of the sample. Other countries, including Ireland (n=2), South Africa (n=2), Belgium (n=1), and Singapore (n=1),

accounted for smaller proportions of the sample, with percentages ranging from 2.8 percent to 5.6 percent.

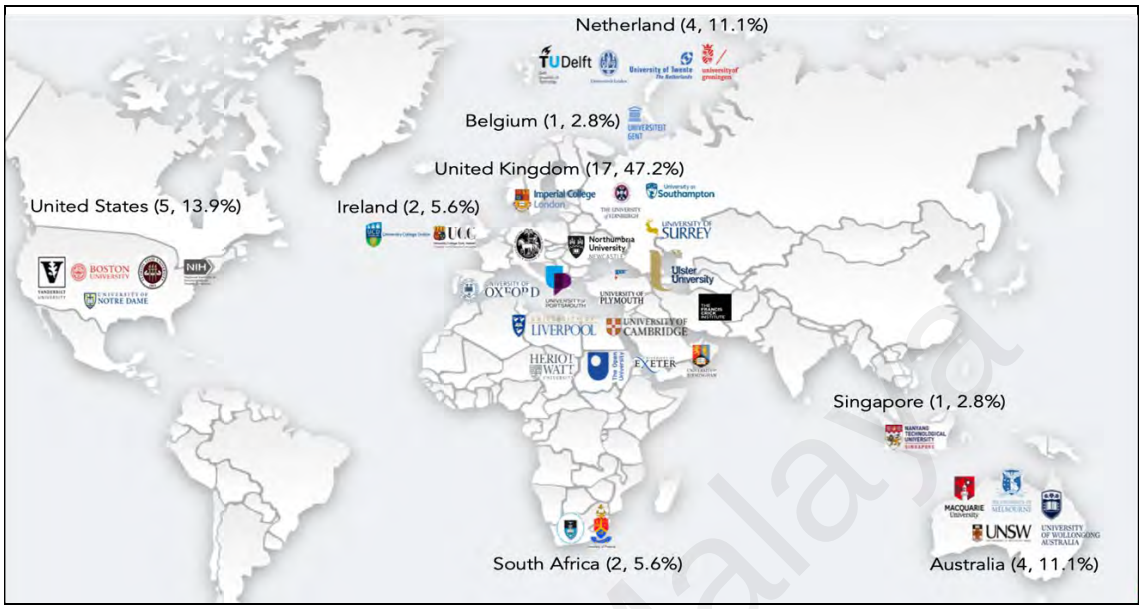


Figure 4.4: Distribution of RDG policy documents across different countries in the sample[3]

These countries have recognized the significance of research, as evidenced by their substantial investments in research and development (R&D) relative to their Gross Domestic Product (GDP) expenditure (World Bank, 2022). Belgium allocated 3.16 percent of its GDP to R&D in 2019, while Australia invested 1.83 percent, the United Kingdom invested 1.71 percent, the United States invested 3.12 percent, and the Netherlands invested 2.18 percent. Furthermore, other countries such as Singapore, Ireland, and South Africa also marked notable investments in R&D as percentages of their GDPs.

Out of the 36 policy documents analyzed, 34 originated from universities, identified by codes U1-U34. The remaining 2 documents were from research institutions and were designated as R1 and R2. The RPOs that contributed to the policy documents in the study

[3] The map and institutional logos were retrieved separately from Google between October 1-4, 2022.

are listed in Table 4.2. The institutions are recognized for their substantial publication presence in two prominent citation databases, namely Scopus and Web of Science (WoS).

Table 4.2: Research performing organizations contributing policy documents in this study

Institution Code	Research Performing Organization	Institution Code	Research Performing Organization
U1	Macquarie University	U19	University of Southampton
U2	University of Birmingham	U20	University of Surrey
U3	Boston University	U21	Delft University of Technology
U4	University of Cambridge	U22	University of Twente
U5	University of Cape Town	U23	University of Ulster
U6	University of Edinburgh	U24	University of New South Wales
U7	University of Exeter	U25	Vanderbilt University
U8	Ghent University	U26	University of Wollongong
U9	Heriot-Watt University	U27	University of Groningen
U10	Imperial College London	U28	University College Cork
U11	Leiden University	U29	University of Notre Dame
U12	University of Liverpool	U30	Florida State University
U13	London School of Hygiene & Tropical Medicine	U31	University of Oxford
U14	Northumbria University	U32	University College Dublin
U15	The Open University	U33	University of Melbourne
U16	University of Plymouth	U34	Nanyang Technological University
U17	University of Portsmouth	R1	The Francis Crick Institute
U18	University of Pretoria	R2	The National Institute of Health

Figure 4.5 illustrates the cumulative count of publications indexed in Scopus and WoS for the period spanning 2018 to 2022 across these institutions. The trend of publications indicates a greater propensity of these institutions towards publishing in WoS as opposed to Scopus. The University of Oxford has the highest number of publications, totaling 68,252 in Scopus and 86,492 in WoS, surpassing all other institutions. Remarkably, a study (Lancho-Barrantes & Cantu-Ortiz, 2021) emphasizes that the University of Oxford is recognized for its extensive collaborations in publication production, firming its position among the most collaborative universities. On the other hand, the Open University produced the least number of publications, which is 3,396 in Scopus and 5,768 in WoS.

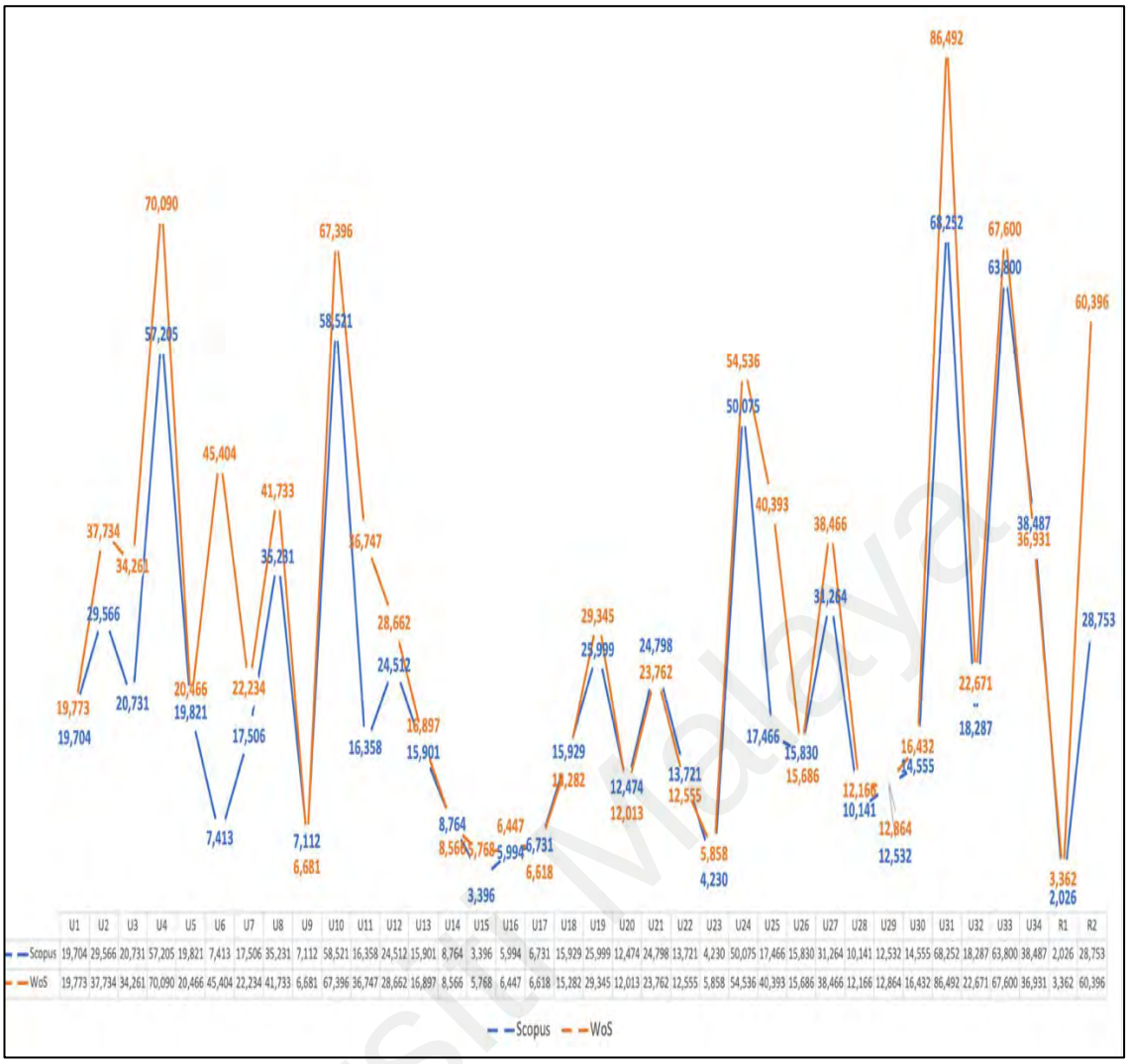


Figure 4.5: Cumulative count of publications indexed in Scopus and Web of Science (2018-2022) by the represented institutions

These institutions are highly esteemed RPOs, as indicated by their Scimago rankings. They exhibit a commitment not only to publications but also to the proper management of research data. This commitment is evident through their well-defined policy documents about RDM, underscoring the importance of RDG within these organizations. Information on the RPOs' Scimago Ranking and the implementation/revision year of their policy documents is presented in Table 4.3.

Table 4.3: Scimago ranking and implementation/revision years of RPOs' policy documents

Item	Number (#/36)	Percentage (100%)	Institution Code
Scimago Ranking			
1-50	5	14	U4, U10, U31, U33, R2
51-100	3	8	U6, U12, U34
101-150	8	22	U2, U3, U8, U11, U19, U21, U25, U27
151-200	1	2.8	U7
201-250	2	5.6	U13, R1
251-300	9	25	U1, U5, U18, U20, U22, U26, U29, U30, U32
301-350	6	17	U14, U15, U16, U17, U23, U28
351-400	1	2.8	U9
401-450	1	2.8	U24
Implementation / Revision Year			
2023	1	3	R2
2022	5	14	U6, U10, U13, U24, U33
2021	5	14	U1, U4, U17, U20, U25
2020	4	11	U15, U16, U26, U32
2019	4	11	U12, U14, U19, U29
2018	6	17	U2, U3, U5, U21, U22, U27
≤2017	8	22	U7, U8, U9, U11, U18, U28, U30, R1
No Data	3	8	U23, U31, U34

4.4.2 Research Data Governance Activities by Top Leading RPOs

The 36 policy documents were analyzed to extract excerpts or statements that contained imperative verbs related to RDG tasks and their associated governance roles. A complete overview of the desk research is illustrated in Figure 4.6. Following all the coding steps discussed earlier (see Figure 4.2), the procedure identified 118 distinct tasks from 640 excerpts of statements that contained imperative verbs related to RDG tasks and their associated governance roles. Each task could be mapped to one of the three RDG activity categories: action, area, or decision domain, based on the context of the identified governance roles, thereby generating an activity. This process is illustrated in a straightforward manner, as shown in the example screenshot of the Excel table in Figure 4.74. This classification resulted in a total of 226 RDG activities across 13 distinct roles.

[4] The dataset containing findings from the desk research is available at: Hazmi, Norzelatun Rodhiah (2024). Findings of the desk research analysis exploring research data governance activities. *Figshare*. Dataset. https://doi.org/10.6084/m9.figshare.26172280.v1.

Table 4.4 lists the terms used in the coding procedure along with the number of results obtained from analyzing the 36 policy documents.

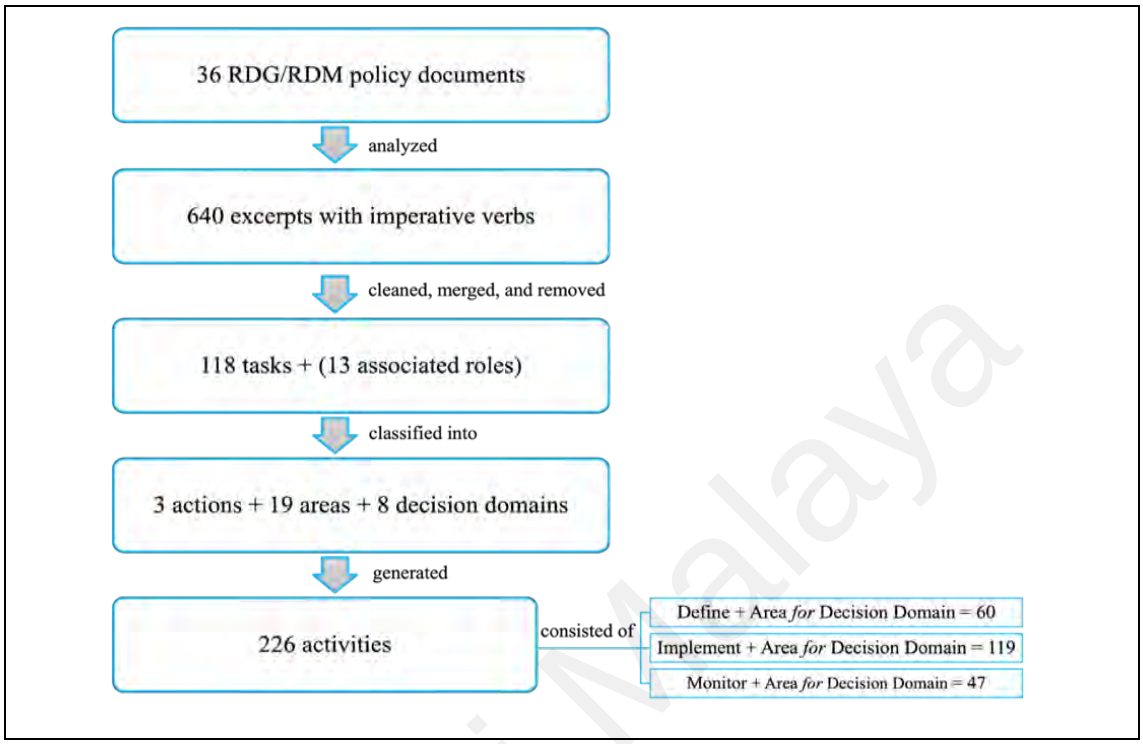
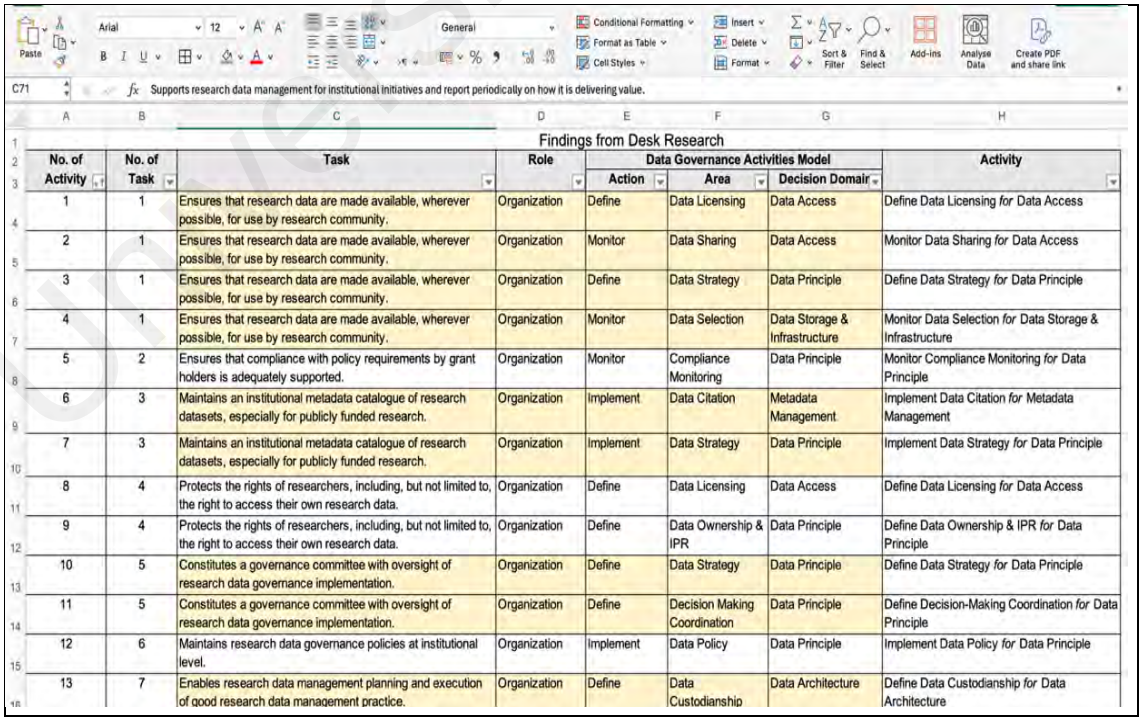


Figure 4.6: A complete overview of the desk research



Findings from Desk Research

No. of Activity	No. of Task	Task	Role	Data Governance Activities Model			Activity
				Action	Area	Decision Domain	
1	1	Ensures that research data are made available, wherever possible, for use by research community.	Organization	Define	Data Licensing	Data Access	Define Data Licensing *for* Data Access
2	1	Ensures that research data are made available, wherever possible, for use by research community.	Organization	Monitor	Data Sharing	Data Access	Monitor Data Sharing *for* Data Access
3	1	Ensures that research data are made available, wherever possible, for use by research community.	Organization	Define	Data Strategy	Data Principle	Define Data Strategy *for* Data Principle
4	1	Ensures that research data are made available, wherever possible, for use by research community.	Organization	Monitor	Data Selection	Data Storage & Infrastructure	Monitor Data Selection *for* Data Storage & Infrastructure
5	2	Ensures that compliance with policy requirements by grant holders is adequately supported.	Organization	Monitor	Compliance Monitoring	Data Principle	Monitor Compliance Monitoring *for* Data Principle
6	3	Maintains an institutional metadata catalogue of research datasets, especially for publicly funded research.	Organization	Implement	Data Citation	Metadata Management	Implement Data Citation *for* Metadata Management
7	3	Maintains an institutional metadata catalogue of research datasets, especially for publicly funded research.	Organization	Implement	Data Strategy	Data Principle	Implement Data Strategy *for* Data Principle
8	4	Protects the rights of researchers, including, but not limited to, the right to access their own research data.	Organization	Define	Data Licensing	Data Access	Define Data Licensing *for* Data Access
9	4	Protects the rights of researchers, including, but not limited to, the right to access their own research data.	Organization	Define	Data Ownership & IPR	Data Principle	Define Data Ownership & IPR *for* Data Principle
10	5	Constitutes a governance committee with oversight of research data governance implementation.	Organization	Define	Data Strategy	Data Principle	Define Data Strategy *for* Data Principle
11	5	Constitutes a governance committee with oversight of research data governance implementation.	Organization	Define	Decision Making Coordination	Data Principle	Define Decision-Making Coordination *for* Data Principle
12	6	Maintains research data governance policies at institutional level.	Organization	Implement	Data Policy	Data Principle	Implement Data Policy *for* Data Principle
13	7	Enables research data management planning and execution of good research data management practice.	Organization	Define	Data Custodianship	Data Architecture	Define Data Custodianship *for* Data Architecture

Figure 4.7: Sample of the data for desk research

Table 4.4: Terms included in coding procedures

Term	Count	Coding example
Statement (excerpt)	640	• Ensure an efficient transmission of general data management information between the central level and the research community (U8) • Ensure that all relevant role-players in their departments are aware of, and adhere to, the University's Research Data Management policy (U18) • Ensure that staff and students are aware of their responsibilities and obligations in effective management of Research Data and identify or promote training where gaps in these skills are identified (U19) • Ensuring that there is appropriate communication and organizational awareness of Research Data Management issues (U26) • Take appropriate steps to inform staff and PhD students about the policy and expected procedures (U27)
Task	118	Ensures an efficient transmission of general research data management information between the central level and the research community.
Role	13	Governor (Faculty level)
Action	3	Monitor
Area of governance	19	Communication
Decision domain	8	Data principle
Activity	226	Monitor communication *for* data principle

Meanwhile, nineteen (19) RDG areas were identified from the analysis and subsequently linked to the eight (8) predefined RDG decision domains, which were deductively identified through reviewing the existing literature, namely data principles, data architecture, data lifecycle, data storage and infrastructure, metadata, data quality, data security, and data access. Table 4.5 presents a detailed list of RDG roles, areas, and decision domains from the findings, following Table 4.6 that illustrates the alignment between RDG areas and decision domains.

Table 4.5: Detailed list of RDG roles, areas, and decision domains from desk research

RDG Roles	RDG Areas	RDG Decision Domain
1. Research Data Governance Committee 2. Funder 3. Organization 4. Library 5. Executive Sponsor 6. Research Data Governor 7. Research Data Steward 8. Data Governance Leader 9. Researcher	1. Communication 2. Compliance Monitoring 3. Data Citation 4. Data Custodianship 5. Data Integrity 6. Data Licensing 7. Data Ownership & IPR 8. Data Policy 9. Data Privacy 10. Data Repository	1. Data Access 2. Data Architecture 3. Data Lifecycle 4. Data Principle 5. Data Quality 6. Data Security 7. Data Storage & Infrastructure 8. Metadata Management

Table 4.5, continued

RDG Roles	RDG Areas	RDG Decision Domain
10. Research Office 11. Office of Research Data Governance 12. Research Data Consumer 13. Information Technology	11. Data Retention & Disposal 12. Data Selection 13. Data Sharing 14. Data Stewardship 15. Data Strategy 16. Decision Making Coordination 17. Issue And Risk Management 18. Performance Measurement 19. Training	

Table 4.6: The mapping of RDG areas with decision domains

No.	RDG Area	RDG Decision Domain								Supporting literature
		Data Principles	Data Architecture	Data Lifecycle	Data Storage & Infrastructure	Metadata	Data Quality	Data Security	Data Access	
1	Data Policy	√	-	-	-	-	-	-	-	Data policy provides guidelines and rules for the use and management of research data, closely connected to data principle (Allela & Mwai, 2019)
2	Data Strategy	√	-	-	-	-	-	-	-	Organizations need focused and tangible data strategies that align with their goals to treat data as valuable assets (Alhassan et al., 2019)
3	Compliance Monitoring	√	-	-	-	-	-	-	-	Compliance monitoring helps ensure that data principles are being followed and enforced (Goel et al., 2021)
4	Data Ownership & IPR	√	-	-	-	-	-	-	-	Defining data ownership and IPR is critical to ensuring that research data is managed and used appropriately (Boullenois, 2021; Sung et al., 2019)
5	Performance Measurement	√	-	-	-	-	-	-	-	Performance measurement is necessary to assess the effectiveness of data governance based on strategic business goals (Fauzy et al., 2021)
6	Decision-Making Coordination	√	-	-	-	-	-	-	-	Decision-making coordination is essential to ensuring that data governance decisions are made efficiently and effectively across departments (Ranathunga & Wickramarachchi, 2021)
7	Communication	√	-	-	-	-	-	-	-	Communication is necessary to ensure that everyone within an organization understands the data governance policies and procedures (Tan & Lim, 2022)
8	Training	√	-	-	-	-	-	-	-	Training is necessary to ensure that everyone within an organization is knowledgeable about data governance and how it affects their work (Alhassan et al., 2019; Parmiggiani & Grisot, 2020; Wang et al., 2018)
9	Data Custodianship		√							Data custodianship manages data architecture (Emam et al., 2019)

Table 4.6, continued

No.	RDG Area	RDG Decision Domain								Supporting literature
		Data Principles	Data Architecture	Data Lifecycle	Data Storage & Infrastructure	Metadata	Data Quality	Data Security	Data Access	
10	Data Stewardship	-	-	√	-	-	-	-	-	Data stewardship oversees the entire research data lifecycle, aiming to make research datasets findable, accessible, interoperable, and reusable (FAIR) (Dunning et al., 2018)
11	Data Selection	-	-	-	√	-	-	-	-	Data selection refers to the process of filtering and choosing the relevant data for use, while data storage and infrastructure is responsible for providing the necessary storage capacity and data management tools to support the selected data (Azeroual et al., 2022)
12	Data Repository	-	-	-	√	-	-	-	-	A data repository is a centralized location where data is stored, managed, and shared. It relies on a reliable data storage and infrastructure system to provide secure, efficient, and scalable storage for the data (Derakhshannia et al., 2020)
13	Data Retention & Disposal	-	-	-	√	-	-	-	-	Data retention & disposal involve setting up policies for how long data should be retained or permanently and data storage & infrastructure plays a crucial role in implementing these policies (Wang et al., 2021)
14	Data Citation	-	-	-	-	√	-	-	-	Metadata includes information that enables the discovery and reuse of research data (Lee et al., 2017), and data citation is one of the ways to ensure proper attribution of the data (Austin et al., 2021)
15	Data Integrity	-	-	-	-	-	√	-	-	Data integrity is a key component of ensuring data quality (Wibisono et al., 2022)
16	Issue & Risk Management	-	-	-	-	-	√	-	-	Issue and risk management plays a vital role in identifying, defining, and escalating data governance-related issues (DAMA International, 2017).
17	Data Privacy	-	-	-	-	-	-	√	-	Data security and data privacy are related as both are important aspects of protecting research data (Maniam & Singh, 2020)
18	Data Licensing	-	-	-	-	-	-	-	√	Data licensing determines the terms and conditions under which research data can be accessed (Grabus & Greenberg, 2019)
19	Data Sharing	-	-	-	-	-	-	-	√	Data sharing involves the dissemination of research data to a wider audience (Hao et al., 2020; Logan et al., 2021; Rousi, 2022)

The coding procedure found that RDG activities for various governance roles under the ‘implement’ action construct are the most frequently executed task among the selected RPOs, comprising 119 activities. Followed by RDG activities under the ‘define’ action construct with 60 activities. Remaining 47 RDG activities under the ‘monitor’ action

construct are the least frequently executed task. Regarding individual activities, the findings indicate that RPOs focus on 'Implement Compliance Monitoring for Data Principle' and 'Monitor Compliance Monitoring for Data Principle' the most, with 16 and 12 occurrences, respectively. Followed closely by 'Define Data Strategy for Data Principle', which had 15 occurrences. The coding details are outlined in Table 4.7.

Table 4.7: The code co-occurrence of the RDG activities identified from desk research

Area *for* Decision Domain	**Action**		
	Define Total=60	Implement Total =119	Monitor Total =47
Communication *for* Data Principle (n=9)	1	7	1
Compliance Monitoring *for* Data Principle (n=30)	2	16	12
Data Ownership & IPR *for* Data Principle (n=5)	1	3	1
Data Policy *for* Data Principle (n=16)	2	10	4
Data Strategy *for* Data Principle (n=29)	15	8	6
Decision-making Coordination *for* Data Principle (n=8)	2	3	3
Performance Measurement *for* Data Principle (n=6)	2	2	2
Training *for* Data Principle (n=12)	4	8	0
Data Custodianship *for* Data Architecture (n=8)	3	4	1
Data Stewardship *for* Data Lifecycle (n=18)	8	9	1
Data Selection *for* Data Storage & Infrastructure (n=4)	2	1	1
Data Repository *for* Data Storage & Infrastructure (n=14)	5	8	1
Data Retention & Disposal *for* Data Storage & Infrastructure (n=9)	2	6	1
Data Citation *for* Metadata (n=7)	1	6	0
Data Integrity *for* Data Quality (n=12)	1	10	1
Issue & Risk Management *for* Data Quality (n=8)	3	4	1
Data Privacy *for* Data Security (n=13)	1	8	4
Data Sharing *for* Data Access (n=8)	2	3	3
Data Licensing *for* Data Access (n=10)	3	3	4

The findings from the desk research were crucial in guiding the development of the Delphi instruments. The identified tasks and associated roles as well as the areas and its predefined decision domains from the desk research served as a foundation for exploring the level of importance and implementation of RDG activities within RPOs in Malaysia.

Additionally, the Delphi study aimed to achieve a priori consensus on the perceived importance of these tasks among the panelists, irrespective of their current implementation status. This insight is intended to assist in developing an RDG framework specifically tailored for RPOs, incorporating best practices for those interested in starting or improving their RDG activities.

4.5 Summary of Chapter Four

In this chapter, a thorough discussion was dedicated solely to the desk research process—a crucial phase aimed at gaining insights into the RDG activities implemented by leading RPOs worldwide. The findings from this phase informed the development of survey instruments for the subsequent Delphi study. This involved selecting policy documents on RDG/RDM for analysis. The entire process of data analysis was detailed, following the approach employed by Alhassan et al. (2018) in their Data Governance Activities Model. This analysis identified 118 tasks and 226 activities across 13 distinct roles. The chapter concludes with a summary that provides an overview of the key points discussed. The next Chapter Five discusses the findings of the Delphi study (Rounds I to III), focusing on the RDG activities being implemented in Malaysia RPOs and the tasks deemed important for implementation within these organizations.

CHAPTER 5: RESEARCH DATA GOVERNANCE ACTIVITIES IN MALAYSIA RESEARCH PERFORMING ORGANIZATIONS[5]

"Help one another in righteousness and piety, but do not help one another in sin and aggression".
(Al-Ma'idah, 5:2)

5.1 Introduction

This chapter presents the insights gained from panel experts through the modified Delphi study, focusing on Rounds I, II, and III. It integrates and synthesizes data from the literature review, desk research, and the Delphi process, aimed to explore RDG functional dimension of RDG activities. This aim guides the discussion in this chapter, which is organized around the two research questions (RQs) below:

RQ1: How are research data governance activities currently being implemented by data practitioners in research performing organizations?

RQ2: What research data governance activities do data practitioners consider significant for research performing organizations?

In addition to these findings, a collaborative effort to establish a consensus on descriptions for individual RDG areas, decision domains, and roles, along with determining the appropriate nomenclature for each role, was deemed essential for the development of the RDG framework. As a result, inquiries regarding these descriptions and nomenclatures were integrated into the Delphi instruments used in both Rounds III and IV. Thus, this chapter covers the RDG areas and decision domains, including its descriptions agreed upon by the panelists. In the meantime, Chapter Six discusses the findings related to the RDG roles.

[5] Several sections of this chapter have been previously published as a journal article entitled: Hazmi, N.R., Abrizah, A. and Yanti Idaya, A.M.K. (2023). Research data governance activities for implementation in Malaysia research performing organizations: Insights from data practitioners via Delphi study. *Malaysian Journal of Library and Information Science, 28*(3), 37-60. https://doi.org/10.22452/mjlis.vol28no3.3.

The findings from the Delphi study are specific to the Malaysian context and the practice of RPOs in Malaysia. Therefore, the identified priorities for RDG activities arise from this context. To offer clarity to the readers, the discussion commences with a demographic section, followed by the findings of both RQ1 and RQ2.

5.2 Demographic profiles of panelists

Two hundred ninety-two (292) individuals were invited to participate in Round I of the study via email; 67 for Set A (Strategic), 27 for Set B (Tactical), 51 for Set C (Operational), and 147 for Set D (Researcher). Forty-seven (47) data professionals who met the specified inclusion criteria (refer to Table 5.1), voluntarily participated in Round I of the study, following the outlined procedures in the consent form. Interestingly, 83 percent (n=39) of the panelists were linked to the MOSP initiative. While the remaining were identified either through snowballing, individuals in upper positions within their institutions, or researchers who have deposited their research data in the Dimensions data repository.

In Round I, three (3) eligible panelists responded to two different sets, namely Set A and Set B, resulting in a total of 50/292 (17.1%) responses. Detailed distributions of response rates in every set yielded a total of 15/67 responses in Set A and 11/27 responses in Set B. Furthermore, 13/51 respondents engaged with Set C, while 11/147 responded to Set D. In Round II, invitations were extended to the same data practitioners (n=47) who had taken part in Round I based on their assigned groups. However, 34 of them responded, resulting in a total of 37 (74.0%) responses out of the 50 responses received in Round I. This total includes the three panelists who had participated in two different sets (Set A and Set B) during the preceding round. In comparison to the total responses in Round I for each set, this breaks down to 13/15 from Set A, 8/11 from Set B, 10/13 from Set C, and 6/11 from Set D.

Table 5.1: Assessment of expert criteria for Round I

Expert per Set	Set	Expert ID	Current Position	Level of Governance/ Management			MOSP-related category			
				*S	*T	*O	*1	*2	*3	*4
1	A	P01	Deputy Chief Librarian	√	-	-	-	-	√	-
2	A	P04	Chief Librarian	√	√	-	√	-	-	-
3	A	P14	Public Health Medicine Specialist	√	-	-	√	-	-	-
4	A	P15	Deputy Director - IT Officer	√	√	-	√	-	-	-
5	A	P17	Professor	√	-	-	√	-	-	-
6	A	P23	Deputy Secretary, MOHE	√	-	-	√	-	-	-
7	A	P27	Chief Librarian	√	√	√				
8	A	P28	Deputy Chief Librarian	√	√	-	-	-	√	-
9	A	P29	Chief Librarian	√	√	-	-	-	√	-
10	A	P32	Senior Research Officer	√	-	-	√	-		-
11	A	P33	Chief Librarian	√	-	-	-	-	√	-
12	A	P34	Senior Lecturer	√	-	-	√	-		-
13/1	A/B	P05	Public Health Researcher	√	√	√	-	-	√	-
14/2	A/B	P20	Deputy Chief Librarian	√	√	-	-	-	√	-
15/3	A/B	P47	Statistician	√	√	-	-	-	√	-
4	B	P02	Senior Lecturer	√	√	-	√	-	-	-
5	B	P03	Senior Librarian	-	√	√	-	√	-	-
6	B	P06	Research Librarian	-	√	-	-	√	-	-
7	B	P07	Senior Lecturer	√	√	√	-		√	-
8	B	P16	Administrative Officer	-	√	-	-	√	-	-
9	B	P24	Senior Research Officer	√	√	-	-	√	-	-
10	B	P25	Senior Librarian	-	√	-	-	√	-	-
11	B	P35	Senior IT Officer	-	√	-	-	-	√	-
1	C	P08	Senior Librarian	-	-	√	-	-	√	-
2	C	P09	Senior Librarian	-	-	√	-	-	√	-
3	C	P10	Senior Librarian	-	√	√	-	-	√	-
4	C	P11	Senior Librarian	-		√	-	-	√	-
5	C	P12	Senior Librarian	-	√	√	-	-	√	-
6	C	P13	Deputy Chief Librarian	-	-	√	-	-	√	-
7	C	P18	Senior Librarian	-	-	√	-	-	√	-
8	C	P19	Librarian	-	-	√	-	-	√	-
9	C	P21	Research Officer	-	√	√	-	-	√	-
10	C	P22	Senior Librarian	-	√	√	-	-	√	-
11	C	P26	Research Officer	-	√	√	-	-	√	-
12	C	P30	Deputy Chief Librarian	-	√	√				
13	C	P31	Deputy Chief Librarian	-	-	√	-	-	√	-
1	D	P36	Senior Lecturer	-	-	√				
2	D	P37	Senior Lecturer	-	-	√				
3	D	P38	Senior Lecturer	√	-	-				
4	D	P39	Senior Lecturer	-	√	-				
5	D	P40	Professor	-	√	-				
6	D	P41	Public Health Researcher & Project Management Professional	√	-	√				
7	D	P42	Senior Lecturer	-	-	√	-	-	-	√
8	D	P43	Senior Medical Lecturer & Radiologist	-	-	√	-	-	-	√
9	D	P44	Senior Lecturer	-	-	√	-	-	-	√
10	D	P45	Senior Lecturer & Medical Officer	-	√	-	-	-	-	√
11	D	P46	Senior Lecturer	-	√	-	-	-	-	√

Notes.

*S-Strategic; T-Tactical; O-Operational; 1-MOSA/ Working group member; 2-Certified Data Steward; 3-TOT on Data Stewardship; 4-MOSP Depositor

In shaded cells, experts were identified using other approaches, e.g. snowballing, individual holding upper-level position, and depositor in Dimensions

All 47 data practitioners were contacted once more to take part in Round III. In this round, the instrument included queries aimed at anticipating different viewpoints on RDG practices and implementation. Thirty-four (34) individuals participated, but one participant's responses required resubmission, which unfortunately did not occur until the survey concluded. Consequently, responses from 33 panelists (70.2%) were accepted and analyzed. Unlike the previous rounds, they were not categorized into specific groups this time, aiming for a broader perspective on RDG activities. In the final round, the group that responded in Round III (n=33) was contacted to participate in this concluding round. Among them, 26 (78.8%) successfully submitted their responses. Table 5.2 provides an overview of the survey response rates across all rounds of this study.

Table 5.2: Survey response rate across study rounds

Round	Solicited	Returned	Total response rate %
Round I-Set A	67	15	-
Round I-Set B	27	11	-
Round I-Set C	51	13	-
Round I-Set D	147	11	-
TOTAL (ROUND I)	**292**	**[a]50**	**17.1%**
Round II-Set A	15	13	-
Round II-Set B	11	8	-
Round II-Set C	13	10	-
Round II-Set D	11	6	-
TOTAL (ROUND II)	**[a]50**	**[b]37**	**74.0%**
ROUND III	**47**	**33**	**70.2%**
ROUND IV	**33**	**26**	**78.8%**

Notes.
[a]Total no. of panelists=47
[b]Total no. of panelists=34

Throughout all rounds, no special follow-up engagement was conducted with individuals who withdrew from the study, understanding that the panelists dropped out due to other commitments. The participation requirement specified that those who took part in the initial round were eligible for involvement in Rounds I, II, and III, while those participating in Round III were eligible for inclusion in Round IV. The study secured the participation of 23 (49%) data practitioners who committed to participate in all four

rounds. Meanwhile, sixty percent (n=28) took part in Rounds I, II, and III, while 79 percent (n=26) of the 33 participants involved in Round III, participated in Round IV.

The study gathered personal and demographic information from the panelists, including details such as age, gender, management level, affiliation, position, and roles related to research data. However, it is noted that demographic information from 8.5 percent (n=4) of the panelists have not been received or provided. Numerous reminder emails had been sent, urging them to either completed and returned the forms or shared the required information via email. Sample of email is attached as Appendix U. As an alternative approach, two versions of the designated instrument were prepared in Round II, with one version specifically designed for the panelists to add their demographic information at the end of the instrument's section. Despite these efforts, 9 out of the 13 individuals who initially did not submit their demographic information have either emailed the required details or provided them through the instrument. Nevertheless, the priority should be on recruiting a diverse panel of subject matter experts. In the context of this study, some basic demographic information about the panelists is useful but not absolutely necessary. Consequently, there were limited information regarding their demographic profile.

Majority of the panelists, 25.5 percent (n=12), were in the 35-39 and 40-44 age category, and 21.3 percent (n=10) were in the 45-49 age category. In terms of gender distribution, most of the panelists were female, accounting for 68 percent (n=32) of the total, while males made up of 29.8 percent (n=14). In terms of organizational affiliation, the majority (80.9%, n=38) were affiliated with public universities. The remaining panelists represented research institutions, government agencies and private university. The panelists were asked to indicate their professional positions within their organizations, and they could choose multiple positions. Most of the panelists (44.7%,

n=21) were librarians, with various roles and experiences (including Chief Librarians, MOSP data stewards, liaison librarians, IT librarians, and archive librarians). Principal investigators and researchers each accounted for 23.4 percent (n=11). The remaining panelists included research officers, executives, heads of research and information technology officers, an administrative officer and an honorary professor.

In addition to asking about panelists' professional positions, questions about their specific roles related to research data were also included. Many panelists held diverse data-related responsibilities, covering a wide range of tasks and roles. Out of the 47 participants, 53.2 percent (n=25) were involved in activities such as accessing, analyzing, and manipulating research data, 38.3 percent (n=18) were providing support for the implementation of RDM policies, 34 percent (n=16) were engaged in conducting research, and 29.8 percent (n=14) were developing and leading research/publication data policies. Additionally, 21.3 percent (n=10) were responsible for overseeing the overall management of data and information governance, while another 21.3 percent (n=10) were ensuring the quality and compliance of RDM. Other reported roles included ensuring legal and regulatory compliance for research data; overseeing the implementation of research data governance policies; and having specific responsibilities for monitoring key risk indicators related to data misconduct. Under the category of "Other" data-related tasks and activities, the panelists mentioned tasks such as raising awareness about open data on campus; serving as a member of the RDM team; validating publications and grants; and establishing a unit dedicated to managing RDM and formulating policies. The demographic profiles of the panelists, along with their research data roles, are presented in Table 5.3.

Table 5.3: Demographic profiles of modified Delphi panelists

Characteristics	Participants	[a]n = 47	(%)
Age	30-34	2	4.3
	35-39	12	25.5
	40-44	12	25.5
	45-49	10	21.3
	50-54	4	8.5
	55-59	2	4.3
	>60	1	2.1
	Not known	4	8.5
Gender	Female	32	68.1
	Male	14	29.8
	Not known	1	2.1
Affiliation	Public University	38	80.9
	Private University	1	2.1
	Research Institution	4	8.5
	Government Agency	3	6.4
	Not known	1	2.1
[b]Position	Administrative Officer	1	2.1
	Executive	3	6.4
	Head of Research	2	4.3
	Honorary Professor	1	2.1
	Information Technology Officer	2	4.3
	Librarian	21	44.7
	Principal Investigator/Research Lead, Supervisor/Mentor	11	23.4
	Research Officer	4	8.5
	Researcher	11	23.4
[b]Research Data-Related Roles	Accessing, analyzing, and manipulating research data	25	53.2
	Developing and leading research/publication data policies	14	29.8
	Ensuring legal and regulatory compliance for research data	8	17.0
	Overseeing overall management of data governance	10	21.3
	Ensuring the quality and compliance of RDM	10	21.3
	Involving in conducting research	16	34.0
	Monitoring key risk indicators of data misconduct	4	8.5
	Ensuring the implementation of research data governance policy	6	12.8
	Providing support for the implementation of RDM policies	18	38.3
	Raising awareness about open data on campus	1	2.1
	Being a member of the RDM team	1	2.1
	Validating publications and grants	2	4.3
	Establishing a unit dedicated to RDM and preparing policies	1	2.1
Years of experience in research data-related roles	<3	20	42.6
	3–5	13	27.7
	6-10	8	17.0
	>10	3	6.4
	Not known	3	6.4

Notes:
[a]Sample size n=47 at Round I
[b]Participants were allowed to select more than one answer

5.3 Research Data Governance implementation among Malaysia Research Performing Organizations

How are research data governance activities currently being implemented by data practitioners in research performing organizations? To specifically address RQ1, the responses obtained from the modified Delphi study: Round I shed light on the implementation of RDG practices and enhance the understanding of how these practices are applied within Malaysia RPOs. Panelists assessed statements related to various RDG tasks, corresponding to different governance roles, using a 5-point scale ranging from 1 (Not Implemented) to 5 (Very Highly Implemented). This assessment aimed to uncover the tasks that have already been implemented and those that are not. Meanwhile, challenges and non-implementation factors are outside of the scope of the study.

It is worth noting that the desk research identified a total of 118 tasks. However, one task identified during the desk research was split into two distinct tasks within the Delphi instrument, resulting in a total of 119 task items. Additionally, the desk research initially highlighted thirteen (13) governance roles, but upon further consideration, some tasks performed by the 'Funder' role seemed more relevant, especially in the Malaysian context, when associated with 'External Bodies' under 'National/Institutional Repository' and 'National Journal Publisher'. This adjustment increased the total number of roles to fifteen (15). Consequently, the Delphi instruments were structured with a total of 119 tasks associated with 15 governance roles.

After concluding Round I, the first action involved exporting data from Cognito Forms to an Excel spreadsheet. Since the instruments utilized a 5-point scale for its questions, it was necessary to convert all panelists responses into numerical values. Subsequently, Excel was used to compute the percentage and mean values for each item response. Based on the results, it is evident that none of the tasks fall within the mean score range of 1.0-

1.49 and 4.50-5.0. These score ranges are interpreted as ‘not implemented’ and ‘very highly implemented’, respectively. However, the study identified 23 (19%) out of 119 tasks that are ‘highly implemented’ by various governance roles within RPOs in Malaysia. It is noteworthy that the ‘Researcher’ role emerged as the primary contributor, accounting for 52 percent (n=12) of the total highly implemented tasks and comprising 50 percent of the total tasks (n=24) associated with them. Meanwhile, ‘Organization’ and ‘Research Data Steward’ accounted for 22 percent (n=5), and ‘Research Data Governor’ represented 4 percent (n=1).

The task with the highest mean score (4.09) pertains to Researcher ‘*R6-protecting intellectual property according to funder or contractual obligations*’. Additionally, ‘*R17-protecting confidential and sensitive data in line with legal and ethical requirements*’ and ‘*R1-selecting research data for long-term preservation based on verification/replication and reuse needs*’, received mean ratings of 4.00 and 3.91, respectively. The findings indicate that all panelists (n=11) have implemented these three tasks. However, some of them do not have experience with the remaining tasks on the "highly implemented" list, selecting the “not implemented” option instead. It is evident that the panelists have prioritized sharing their research data by implementing governance tasks to ensure sensitivity, reproducibility, integrity, and quality, as well as preservation of the research data at hand. While the role of the Organization in ‘*ORG9-ensuring ethical use of research elements such as animals, human subjects, and materials’*, received a mean rating of 4.00, reflecting significant implementation. In addition to these tasks, there are other tasks that received higher mean ratings, ranging from 3.55 to 3.91. Table 5.4 enumerates tasks associated with various governance bodies categorized as highly implemented RDG tasks.

Table 5.4: Highly implemented RDG tasks

NO	[a]TASK ID	TASK	[b]N	[c]RESPONSE (n(%))					[d]M
				1	2	3	4	5	
1	DS5	Ensures the archival of data for long-term preservation.	11	1 (9.09)	0 (0.00)	2 (18.18)	5 (45.45)	3 (27.27)	3.82
2	DS3	Ensures appropriate classification and management of research data based on their sensitivity.	11	2 (18.18)	0 (0.00)	1 (9.09)	5 (45.45)	3 (27.27)	3.64
3	DS6	Ensures data licensing for reuse and dissemination.	11	1 (9.09)	1 (9.09)	2 (18.18)	4 (36.36)	3 (27.27)	3.64
4	DS11	Develops processes for data selection, storage, and protection.	11	2 (18.18)	0 (0.00)	1 (9.09)	6 (54.55)	2 (18.18)	3.55
5	DS8	Provides training and support for data management.	11	1 (9.09)	2 (18.18)	1 (9.09)	4 (36.36)	3 (27.27)	3.55
6	GOV1	Facilitates resources and support for research data management.	11	0 (0.00)	2 (18.18)	3 (27.27)	2 (18.18)	4 (36.36)	3.73
7	ORG9	Ensures ethical use of research elements like animals, human subjects, and materials.	47	4 (8.51)	2 (4.26)	3 (6.38)	19 (40.43)	19 (40.43)	4.00
8	ORG10	Facilitates investigations into scientific misconduct or conflict of interest.	47	3 (6.38)	3 (6.38)	8 (17.02)	18 (38.30)	15 (31.91)	3.83
9	ORG12	Owns all research data and associated intellectual property.	47	4 (8.51)	4 (8.51)	5 (10.64)	18 (38.30)	16 (34.04)	3.81
10	ORG4	Safeguards researchers' rights, ensuring access to their data.	47	3 (6.38)	7 (14.89)	5 (10.64)	14 (29.79)	18 (38.30)	3.79
11	ORG2	Ensures support for grant holders in policy compliance.	47	3 (6.38)	6 (12.77)	6 (12.77)	18 (38.30)	14 (29.79)	3.72
12	R6	Protects intellectual property according to funder or contractual obligations.	11	0 (0.00)	1 (9.09)	1 (9.09)	5 (45.45)	4 (36.36)	4.09
13	R17	Protects confidential and sensitive data in line with legal and ethical requirements.	11	0 (0.00)	2 (18.18)	1 (9.09)	3 (27.27)	5 (45.45)	4.00
14	R1	Selects research data for long-term preservation based on verification/replication and reuse needs.	11	0 (0.00)	2 (18.18)	1 (9.09)	4 (36.36)	4 (36.36)	3.91
15	R20	Regularly backs up research data in accordance with best practices.	11	1 (9.09)	1 (9.09)	1 (9.09)	3 (27.27)	5 (45.45)	3.91
16	R9	Guarantees integrity, quality, security, and persistent availability of research data.	11	1 (9.09)	0 (0.00)	2 (18.18)	4 (36.36)	4 (36.36)	3.91
17	R14	Manages data throughout the lifecycle in line with policies, guidelines, and requirements.	11	2 (18.18)	0 (0.00)	2 (18.18)	2 (18.18)	5 (45.45)	3.73
18	R23	Identifies and addresses research data integrity and quality issues.	11	2 (18.18)	0 (0.00)	1 (9.09)	4 (36.36)	4 (36.36)	3.73
19	R19	Upholds open-source file formats and types recommended for preservation.	11	2 (18.18)	0 (0.00)	2 (18.18)	3 (27.27)	4 (36.36)	3.64
20	R8	Publishes data to established repositories, maximizing research value.	11	1 (9.09)	3 (27.27)	0 (0.00)	2 (18.18)	5 (45.45)	3.64
21	R10	Disposes of data and materials securely.	11	1 (9.09)	2 (18.18)	1 (9.09)	4 (36.36)	3 (27.27)	3.55
22	R18	Participates in training on research data management and contractual obligations.	11	1 (9.09)	2 (18.18)	1 (9.09)	4 (36.36)	3 (27.27)	3.55
23	R21	Develops operating procedures to comply with research data governance.	11	2 (18.18)	0 (0.00)	2 (18.18)	4 (36.36)	3 (27.27)	3.55

Notes.
[a]DS-Research Data Steward; GOV-Research Data Governor; ORG-Organization; R-Researcher
[b]Different sets of instruments having different number of panelists
[c]Score ranging from: 1-Not Implemented; 2-Slightly Implemented; 3-Moderately Implemented; 4-Highly Implemented; 5-Very Highly Implemented
[d]M-Mean
Information was sorted by Task ID and Mean score from the highest to the lowest

Comparing to the qualitative dimension of 226 activities generated from the desk research for these 23 tasks, the findings revealed that the tasks contribute to 46 RDG activities. A majority of the 'highly implemented' tasks among RPOs in Malaysia fall under the 'implement' action construct, accounting for 30 activities. Following this, the 'monitor' and 'define' action constructs each account for 8 activities, respectively.

The highly implemented tasks involve four (4) key RDG roles: (i) the Organization, (ii) Research Data Governor, (iii) Research Data Steward, and (iv) Researcher. They cover fourteen (14) governance areas, namely: (i) Compliance Monitoring, (ii) Data Ownership & Intellectual Property Rights, (iii) Data Policy, (iv) Data Strategy, (v) Training, (vi) Data Stewardship, (vii) Data Selection, (viii) Data Repository, (ix) Data Retention & Disposal, (x) Data Integrity, (xi) Issue & Risk Management, (xii) Data Privacy, (xiii) Data Sharing, and (xiv) Data Licensing. These areas are related to six predefined (6) decision domains: (i) Data Principle, (ii) Data Lifecycle, (iii) Data Storage & Infrastructure, (iv) Data Quality, (v) Data Security, and (vi) Data Access.

Within the 'implement' action construct, eight activities are under '*Implement Compliance Monitoring for Data Principle*'. Following this, '*Implement Data Retention & Disposal for Data Storage & Infrastructure*' and '*Implement Data Integrity for Data Quality*' each constitute of four activities. However, it is noteworthy that the highly implemented tasks by RPOs in Malaysia do not cover the following five (5) areas and decision domains: (i) Communication for Data Principle, (ii) Decision-making Coordination for Data Principle, (iii) Performance Measurement for Data Principle, (iv) Data Custodianship for Data Architecture, and (v) Data Citation for Metadata. Table 5.5 compares the code co-occurrence of total activities with activities classified as 'highly implemented tasks'.

Table 5.5: Total RDG activities vs. Highly implemented RDG activities

Total Activities by 'Action' construct*			**Total Activities for Highly Implemented Tasks by 'Action' construct****			**Total Activities for Highly Implemented Tasks by Area *for* Decision Domain (Highly Implemented activity**/Total activity*)**
Define Total=60	Implement Total =119	Monitor Total =47	Define Total=8	Implement Total =30	Monitor Total =8	
1	7	1				Communication *for* Data Principle (n=0/9)
2	16	12	0	8	1	Compliance Monitoring *for* Data Principle (n=9/30)
1	3	1	1	2	0	Data Ownership & IPR *for* Data Principle (n=3/5)
2	10	4	0	2	0	Data Policy *for* Data Principle (n=2/16)
15	8	6	2	0	0	Data Strategy *for* Data Principle (n=2/29)
2	3	3				Decision-making Coordination *for* Data Principle (n=0/8)
2	2	2				Performance Measurement *for* Data Principle (n=0/6)
4	8	0	0	2	0	Training *for* Data Principle (n=2/12)
3	4	1				Data Custodianship *for* Data Architecture (n=0/8)
8	9	1	2	2	0	Data Stewardship *for* Data Lifecycle (n=4/18)
2	1	1	0	1	0	Data Selection *for* Data Storage & Infrastructure (n=1/4)
5	8	1	0	1	1	Data Repository *for* Data Storage & Infrastructure (n=2/14)
2	6	1	0	4	0	Data Retention & Disposal *for* Data Storage & Infrastructure (n=4/9)
1	6	0				Data Citation *for* Metadata (n=0/7)
1	10	1	0	4	0	Data Integrity *for* Data Quality (n=4/12)
3	4	1	1	1	0	Issue & Risk Management *for* Data Quality (n=2/8)
1	8	4	0	2	2	Data Privacy *for* Data Security (n=4/13)
2	3	3	0	1	1	Data Sharing *for* Data Access (n=2/8)
3	3	4	2	0	3	Data Licensing *for* Data Access (n=5/10)

Note:
* The shaded rows signify that no tasks have been implemented in any 'Action' constructs within this category

Meanwhile, the majority of the tasks within RPOs have seen 'moderate' implementation, comprising 77 (65%) of the total moderately implemented tasks (n=119). This includes 16 percent (n=12), constituting another 50 percent, associated with the Researcher role. The Organization role contributes 17 percent (n=13), and both the

Research Data Steward and Research Data Governor roles contribute 16 percent (n=12) each. The remaining tasks fall under the purview of other governance roles such as the Funder, Executive Sponsor, Library, and Research Management Office. In this category, tasks associated with Research Data Steward and Researcher roles such as '*DS17-arranging resources and support for research data management in the group*', '*DS9-ensuring data retention as per organizational guidelines*', '*R15-arranging safe and secure storage for research materials*', '*R16-understanding and addressing risks of third-party storage solutions'*, and '*R22-working with the information security team to ensure system controls*', all received the highest mean rating of 3.45. This suggests that these moderately implemented tasks have a higher level of implementation.

However, the tasks from various roles like '*IT3-providing secure access management following ICT security guidelines*', '*LIB2-managing research data metadata records and publishes them on a public catalog*', '*LIB5-providing high-quality infrastructure for data collection, storage, and sharing*', '*RO4-organizing training events on research data management*', and '*RO5-providing advice, guidance, and assistance to researchers in preparing data management plans*' achieved a mean score of 2.54. Similarly, the tasks of '*COMM1-providing a common vocabulary for primary research data entities and types essential to the organization*' and '*EXE1-providing necessary facilities and support for efficient research data management*', received a mean score of 2.53. These tasks, having among the lowest means in this ranking, are positioned closest to the borderline (2.49) indicating a slightly 'moderately implemented' status. Table 5.6 outlines the tasks corresponding to diverse governance bodies classified as moderately implemented RDG tasks.

Table 5.6: Moderately implemented RDG tasks

NO	[a]TASK ID	TASK	[b]N	[c]RESPONSE (n(%))					[d]M
				1	2	3	4	5	
1	COMM2	Develops tools, guidelines, principles, and policies for research data, covering classification, access, usage, integrity, retention, roles, incident response, and integration.	15	3 (20.00)	4 (26.67)	5 (33.33)	2 (13.33)	1 (6.67)	2.60
2	COMM1	Provides a common vocabulary for primary research data entities and types, essential to the organization.	15	4 (26.67)	3 (20.00)	5 (33.33)	2 (13.33)	1 (6.67)	2.53
3	CUST1	Acknowledges data sources and adhere to access terms and conditions.	13	3 (23.08)	1 (7.69)	3 (23.08)	6 (46.15)	0 (0.00)	2.92
4	CUST3	Enters into data use agreements for accessing embargoed research data.	13	3 (23.08)	4 (30.77)	1 (7.69)	5 (38.46)	0 (0.00)	2.62
5	DGO4	Provides advice, guidance, and reviews on research data management, emphasizing research integrity and ethics.	15	4 (26.67)	2 (13.33)	3 (20.00)	5 (33.33)	1 (6.67)	2.80
6	DGO2	Approves and implements processes supporting research data governance policies.	15	3 (20.00)	4 (26.67)	4 (26.67)	3 (20.00)	1 (6.67)	2.67
7	DGO1	Signs off policies, supports cultural and behavioural changes, and allocates resources to research data governance activities.	15	3 (20.00)	4 (26.67)	5 (33.33)	2 (13.33)	1 (6.67)	2.60
8	DGO3	Monitors compliance with research data governance policies and supporting processes.	15	4 (26.67)	3 (20.00)	4 (26.67)	3 (20.00)	1 (6.67)	2.60
9	DS17	Arranges resources and support for research data management in the group.	11	1 (9.09)	3 (27.27)	0 (0.00)	4 (36.36)	3 (27.27)	3.45
10	DS9	Ensures data retention as per organizational guidelines.	11	2 (18.18)	0 (0.00)	2 (18.18)	5 (45.45)	2 (18.18)	3.45
11	DS1	Securely store and protect active research data to meet compliance requirements.	11	2 (18.18)	2 (18.18)	1 (9.09)	2 (18.18)	4 (36.36)	3.36
12	DS7	Ensures management of custodial responsibilities for departing researchers.	11	3 (27.27)	0 (0.00)	2 (18.18)	2 (18.18)	4 (36.36)	3.36
13	DS12	Monitors group's data management regulations.	11	3 (27.27)	0 (0.00)	2 (18.18)	3 (27.27)	3 (27.27)	3.27
14	DS15	Ensures appropriate data access to project team members.	11	3 (27.27)	0 (0.00)	2 (18.18)	3 (27.27)	3 (27.27)	3.27
15	DS16	Establishes clear data management responsibilities in the group.	11	3 (27.27)	0 (0.00)	2 (18.18)	3 (27.27)	3 (27.27)	3.27
16	DS4	Ensures the security of confidential data in accordance with privacy laws.	11	3 (27.27)	0 (0.00)	2 (18.18)	3 (27.27)	3 (27.27)	3.27
17	DS10	Reviews and recommends improvements to data management plans.	11	2 (18.18)	1 (9.09)	1 (9.09)	7 (63.64)	0 (0.00)	3.18
18	DS13	Ensures the availability of data management plans for research projects.	11	3 (27.27)	1 (9.09)	1 (9.09)	4 (36.36)	2 (18.18)	3.09
19	DS2	Assists in external audits and provides data access.	11	4 (36.36)	0 (0.00)	1 (9.09)	3 (27.27)	3 (27.27)	3.09
20	DS14	Ensures the inclusion of data management costs in research proposals.	11	3 (27.27)	1 (9.09)	3 (27.27)	3 (27.27)	1 (9.09)	2.82
21	EXE1	Provides necessary facilities and support for efficient research data management.	15	3 (20.00)	4 (26.67)	6 (40.00)	1 (6.67)	1 (6.67)	2.53

Table 5.6, continued

NO	[a]TASK ID	TASK	[b]N	[c]RESPONSE (n(%))					[d]M
				1	2	3	4	5	
22	FUN2	Reviews implementation of research data management plans.	24	6 (25.00)	5 (20.83)	2 (8.33)	3 (12.50)	8 (33.33)	3.08
23	FUN1	Defines research data governance principles aligned with funding principles.	24	5 (20.83)	5 (20.83)	3 (12.50)	6 (25.00)	5 (20.83)	3.04
24	FUN3	Specifies retention periods for significant research data.	24	5 (20.83)	4 (16.67)	5 (20.83)	5 (20.83)	5 (20.83)	3.04
25	FUN4	Provides advice directly or through data services.	24	7 (29.17)	6 (25.00)	2 (8.33)	6 (25.00)	3 (12.50)	2.67
26	GOV2	Approves storage, disposal, and publication of research data.	11	1 (9.09)	1 (9.09)	4 (36.36)	3 (27.27)	2 (18.18)	3.36
27	GOV3	Models responsible data management behaviour.	11	2 (18.18)	1 (9.09)	3 (27.27)	2 (18.18)	3 (27.27)	3.27
28	GOV7	Reports IT security incidents and data breaches.	11	2 (18.18)	1 (9.09)	3 (27.27)	2 (18.18)	3 (27.27)	3.27
29	GOV8	Supervises adherence to regulations and procedures.	11	3 (27.27)	0 (0.00)	2 (18.18)	3 (27.27)	3 (27.27)	3.27
30	GOV11	Makes decisions in ambiguous data guidelines.	11	2 (18.18)	2 (18.18)	2 (18.18)	2 (18.18)	3 (27.27)	3.18
31	GOV13	Sponsors, secures, and/or influences resources for research data management.	11	1 (9.09)	3 (27.27)	3 (27.27)	1 (9.09)	3 (27.27)	3.18
32	GOV5	Ensures the requirement of data management plans for grant applications.	11	3 (27.27)	1 (9.09)	1 (9.09)	3 (27.27)	3 (27.27)	3.18
33	GOV6	Ensures compliance by principal investigators.	11	3 (27.27)	0 (0.00)	3 (27.27)	2 (18.18)	3 (27.27)	3.18
34	GOV12	Sets expectations for data classification and retention.	11	2 (18.18)	2 (18.18)	3 (27.27)	1 (9.09)	3 (27.27)	3.09
35	GOV4	Ensures effective communication on data management.	11	2 (18.18)	1 (9.09)	4 (36.36)	2 (18.18)	2 (18.18)	3.09
36	GOV10	Promotes a culture of data awareness and training.	11	3 (27.27)	1 (9.09)	3 (27.27)	2 (18.18)	2 (18.18)	2.91
37	GOV9	Approves attestations for researcher awareness.	11	4 (36.36)	1 (9.09)	1 (9.09)	3 (27.27)	2 (18.18)	2.82
38	IREP2	Specifies required contextual information and metadata for deposited data.	24	4 (16.67)	5 (20.83)	8 (33.33)	4 (16.67)	3 (12.50)	2.88
39	IREP1	Defines non-proprietary standards for access, use, and interpretation.	24	5 (20.83)	5 (20.83)	8 (33.33)	4 (16.67)	2 (8.33)	2.71
40	IT3	Provides secure access management following ICT security guidelines.	13	4 (30.77)	4 (30.77)	1 (7.69)	2 (15.38)	2 (15.38)	2.54
41	LEAD4	Develops and maintains a central repository for governance policies, guiding principles, and decisions.	15	3 (20.00)	4 (26.67)	4 (26.67)	3 (20.00)	1 (6.67)	2.67
42	LEAD5	Maintains the Research Data Governance Committee agenda and convenes meetings.	15	4 (26.67)	4 (26.67)	2 (13.33)	3 (20.00)	2 (13.33)	2.67
43	LIB1	Coordinates a network of data stewards.	13	3 (23.08)	4 (30.77)	2 (15.38)	3 (23.08)	1 (7.69)	2.62
44	LIB3	Maintains a research data repository for registering data and other outputs.	13	3 (23.08)	5 (38.46)	1 (7.69)	2 (15.38)	2 (15.38)	2.62
45	LIB4	Defines the institution's support for research data management.	13	3 (23.08)	5 (38.46)	1 (7.69)	2 (15.38)	2 (15.38)	2.62
46	LIB2	Manages research data metadata records and publishes them on a public catalog.	13	3 (23.08)	5 (38.46)	2 (15.38)	1 (7.69)	2 (15.38)	2.54
47	LIB5	Provides high-quality infrastructure for data collection, storage, and sharing.	13	3 (23.08)	4 (30.77)	3 (23.08)	2 (15.38)	1 (7.69)	2.54

Table 5.6, continued

NO	[a]TASK ID	TASK	[b]N	[c]RESPONSE (n(%))					[d]M
				1	2	3	4	5	
48	ORG17	Ensures storage facilities comply with legal and regulatory requirements.	47	5 (10.64)	10 (21.28)	7 (14.89)	13 (27.66)	12 (25.53)	3.36
49	ORG11	Takes custody of data when necessary for appropriate access.	47	9 (19.15)	6 (12.77)	8 (17.02)	14 (29.79)	10 (21.28)	3.21
50	ORG6	Maintains institutional-level research data governance policies.	47	7 (14.89)	9 (19.15)	11 (23.40)	10 (21.28)	10 (21.28)	3.15
51	ORG8	Collaborates with funders, policymakers, and stakeholders to align governance with sector requirements.	47	7 (14.89)	7 (14.89)	12 (25.53)	14 (29.79)	7 (14.89)	3.15
52	ORG14	Provides researchers with training on effective research data management.	47	8 (17.02)	6 (12.77)	13 (27.66)	12 (25.53)	8 (17.02)	3.13
53	ORG7	Enables planning and execution of good research data management practices.	47	7 (14.89)	12 (25.53)	5 (10.64)	15 (31.91)	8 (17.02)	3.11
54	ORG5	Establishes a governance committee overseeing data governance implementation.	47	4 (8.51)	13 (27.66)	14 (29.79)	9 (19.15)	7 (14.89)	3.04
55	ORG16	Offers facilities, advisory services, and resources for secure data storage and management.	47	6 (12.77)	12 (25.53)	11 (23.40)	11 (23.40)	7 (14.89)	3.02
56	ORG18	Acknowledges contributions of researchers who generate, preserve, and share key datasets.	47	10 (21.28)	12 (25.53)	6 (12.77)	9 19.15)	10 (21.28)	2.94
57	ORG15	Ensures long-term stewardship based on institutional/national data infrastructure.	47	6 (12.77)	15 (31.91)	14 (29.79)	6 (12.77)	6 (12.77)	2.81
58	ORG13	Develops field-specific data sharing best practices.	47	8 (17.02)	14 (29.79)	12 (25.53)	6 (12.77)	7 (14.89)	2.79
59	ORG1	Ensures research data availability for the research community.	47	7 (14.89)	14 (29.79)	15 (31.91)	8 (17.02)	3 (6.38)	2.70
60	ORG3	Manages an institutional metadata catalogue, especially for publicly funded research.	47	10 (21.28)	17 (36.17)	6 (12.77)	8 (17.02)	6 (12.77)	2.64
61	PUB1	Endorses research data deposits in established repositories.	24	5 (20.83)	6 (25.00)	7 (29.17)	3 (12.50)	3 (12.50)	2.71
62	R15	Arranges safe and secure storage of research materials.	11	2 (18.18)	0 (0.00)	2 (18.18)	5 (45.45)	2 (18.18)	3.45
63	R16	Understands and addresses risks of third-party storage solutions.	11	2 (18.18)	0 (0.00)	2 18.18)	5 (45.45)	2 (18.18)	3.45
64	R22	Works with the information security team to ensure system controls.	11	1 (9.09)	2 (18.18)	2 (18.18)	3 (27.27)	3 (27.27)	3.45
65	R2	Retains research data securely for a period determined by the organization.	11	1 (9.09)	2 (18.18)	3 (27.27)	2 (18.18)	3 (27.27)	3.36
66	R4	Ensures written agreements cover ownership, sharing, storage, and disposal of research data.	11	0 (0.00)	4 (36.36)	2 (18.18)	2 (18.18)	3 (27.27)	3.36
67	R13	Hands over data and materials after projects or leaving the institution.	11	2 (18.18)	1 (9.09)	3 (27.27)	2 (18.18)	3 (27.27)	3.27
68	R7	Provides sufficient metadata for discoverability and reusability of research data.	11	2 (18.18)	2 (18.18)	1 (9.09)	3 (27.27)	3 (27.27)	3.27
69	R5	Writes a comprehensive research data management plan.	11	2 (18.18)	0 (0.00)	4 (36.36)	4 (36.36)	1 (9.09)	3.18
70	R12	Reports security breaches affecting research data.	11	2 (18.18)	2 (18.18)	3 (27.27)	1 (9.09)	3 (27.27)	3.09
71	R11	Includes a data access statement in publications.	11	1 (9.09)	5 (45.45)	1 (9.09)	1 (9.09)	3 (27.27)	3.00
72	R24	Periodically reviews data access and usage agreements.	11	4 (36.36)	0 (0.00)	2 (18.18)	2 (18.18)	3 (27.27)	3.00
73	R3	Budgets costs for capturing, managing, archiving, and sharing research data.	11	2 (18.18)	3 (27.27)	1 (9.09)	3 (27.27)	2 (18.18)	3.00

Table 5.6, continued

NO	[a]TASK ID	TASK	[b]N	[c]RESPONSE (n(%))					[d]M
				1	2	3	4	5	
74	RO1	Advocates organizational awareness of research data management issues.	13	2 (15.38)	5 (38.46)	2 (15.38)	1 (7.69)	3 (23.08)	2.85
75	RO2	Ensures governance policies are updated based on the latest funder requirements and national directives.	13	2 (15.38)	5 (38.46)	1 (7.69)	4 (30.77)	1 (7.69)	2.77
76	RO4	Organizes training events on research data management.	13	3 (23.08)	4 (30.77)	3 (23.08)	2 (15.38)	1 (7.69)	2.54
77	RO5	Provides advice, guidance, and assistance to researchers in preparing data management plans.	13	5 (38.46)	2 (15.38)	2 (15.38)	2 (15.38)	2 (15.38)	2.54

Notes.
[a]COMM-Research Data Governance Committee; CUST-Research Data Consumer; DGO-Data Governance Office; DS-Research Data Steward; EXE-Executive Sponsor; FUN-National Funder; GOV-Research Data Governor; IREP-National and/or Institutional Repository; IT-Information Technology Office; LEAD-Data Governance Leader; LIB-Library; ORG-Organization; PUB-National Journal Publisher; R-Researcher; RO-Research Management Office
[b]Different set of instrument having different number of panelists
[c]Score ranging from: 1-Not Implemented; 2-Slightly Implemented; 3-Moderately Implemented; 4-Highly Implemented; 5-Very Highly Implemented
[d]M-Mean
Information was sorted by Task ID and Mean score from the highest to the lowest

Upon examining 'moderately implemented' tasks through the lens of the Data Governance Activities Model, it becomes evident that data practitioners in Malaysia RPOs predominantly carry out 'moderately implemented' tasks across all nineteen areas and eight predefined decision domains, involving all fifteen identified roles. From 77 tasks, a total of 145 activities were retrieved. This includes 40 activities under the 'define' action construct, 74 activities under the 'implement' action construct, and 31 activities under the 'monitor' action construct.

The highest score for moderately implemented tasks is observed in '*Monitor Compliance Monitoring for Data Principle*' activity, constituting 11 occurrences. Following this is '*Define Data Strategy for Data Principle*' with 9 occurrences, and the remaining activities having varying occurrences. Table 5.7 illustrates the code co-occurrence of RDG tasks classified as 'moderately implemented'.

Table 5.7: Total RDG activities vs. Moderately implemented RDG activities

Total Activities by 'Action' construct*			**Total Activities for Moderately Implemented Tasks by 'Action' construct****			**Total Activities for Moderately Implemented Tasks by Area *for* Decision Domain (Moderately Implemented activity**/Total activity*)**
Define Total=60	Implement Total =119	Monitor Total =47	Define Total=40	Implement Total =74	Monitor Total =31	
1	7	1	0	6	1	Communication *for* Data Principle (n=7/9)
2	16	12	2	6	11	Compliance Monitoring *for* Data Principle (n=19/30)
1	3	1	0	1	1	Data Ownership & IPR *for* Data Principle (n=2/5)
2	10	4	1	7	2	Data Policy *for* Data Principle (n=10/16)
15	8	6	9	7	4	Data Strategy *for* Data Principle (n=20/29)
2	3	3	1	2	3	Decision-making Coordination *for* Data Principle (n=6/8)
2	2	2	1	1	1	Performance Measurement *for* Data Principle (n=3/6)
4	8	0	3	3	0	Training *for* Data Principle (n=6/12)
3	4	1	3	3	1	Data Custodianship *for* Data Architecture (n=7/8)
8	9	1	4	6	0	Data Stewardship *for* Data Lifecycle (n=10/18)
2	1	1	2	0	1	Data Selection *for* Data Storage & Infrastructure (n=3/4)
5	8	1	5	5	0	Data Repository *for* Data Storage & Infrastructure (n=10/14)
2	6	1	2	1	1	Data Retention & Disposal *for* Data Storage & Infrastructure (n=4/9)
1	6	0	1	6	0	Data Citation *for* Metadata (n=7/7)
1	10	1	0	6	0	Data Integrity *for* Data Quality (n=6/12)
3	4	1	2	3	1	Issue & Risk Management *for* Data Quality (n=6/8)
1	8	4	1	6	2	Data Privacy *for* Data Security (n=9/13)
2	3	3	2	2	1	Data Sharing *for* Data Access (n=5/8)
3	3	4	1	3	1	Data Licensing *for* Data Access (n=5/10)

The remaining 16 percent (n=19) of the tasks are 'slightly implemented' by data practitioners who hold various governance roles within Malaysia RPOs. The mean scores for these tasks range from 2.13 to 2.46. The task with the highest mean score (2.46) for this ranking fall under the role of 'Research Data Consumer', indicating a lack of

'*CUST2-compliance with controls specified in the data management plan*'. Other tasks receiving the same score (2.46) come from Administrative Office; the 'Information Technology Office', which offers '*IT2-minimal technical support for data management, including storage, backup, and archiving*', and the 'Research Management Office', which infrequently '*RO3-facilitates and supports the implementation of RDG processes*'.

The tasks of the 'Research Data Governance Committee', which involve '*COMM4-overseeing initiatives to enhance effective and efficient research data utilization', 'supporting research data management for institutional initiatives',* and *'COMM5-cultivating a research data management culture that adds value to the institution*', also received a lower mean score (2.40), indicating a low level of implementation. Similarly, tasks under the responsibility of the 'Data Governance Leader' received the same mean score of 2.40. These tasks include '*LEAD1-ensuring the regular review and maintenance of research data governance policies in line with sector requirements', 'LEAD2-deciding on day-to-day matters related to research data governance',* and *'LEAD3-directing decision-making to appropriate stakeholders when needed*'. Other tasks under the purview of the 'Information Technology Office' and 'Research Management Office', such as '*IT1-providing a standardized, robust infrastructure for effective research data management' and 'RO6-offering templates for incoming and outgoing research data agreements'*, received mean scores of 2.38, respectively.

The remaining eight tasks, falling under the jurisdiction of various governance roles, scored mean values ranging from 2.13 to 2.33. For instance, the Executive Sponsor, who is expected to *'EXE2-oversee the implementation of research data governance as a framework for good research data management'*, and the Data Governance Leader, who is expected to *'LEAD6-communicate outcomes of the Research Data Governance Committee'* and *'LEAD7-serve as an expert on research data governance and*

recommend solutions', both scored within this range. Similarly, tasks that should be handled by the Library, including *'LIB6-reviewing research data management policies with governance bodies'*, *'LIB7-providing training on metadata standards, data management plans, legal/ethical issues'*, and *LIB8-'offering advice on research data costs throughout the project lifecycle'*, as well as tasks of the Research Data Committee, supposedly in *'COMM3-expressing strategic research data requirements reflecting organizational and management needs' and 'COMM6-periodically reporting on the value delivered by research data management'*, all fell within this mean score range, almost to the point of not being implemented. This indicates a quite lower level of implementation. The tasks for various governance roles falling under the slightly implemented RDG tasks are presented in Table 5.8.

Table 5.8: Slightly implemented RDG tasks

NO	[a]TASK ID	TASK	[b]N	[c]RESPONSE (%)					[d]M
				1	2	3	4	5	
1	COMM4	Oversees initiatives to enhance effective and efficient research data utilization.	15	5 (33.33)	5 (33.33)	1 (6.67)	2 (13.33)	2 (13.33)	2.40
2	COMM5	Supports research data management for institutional initiatives.	15	5 (33.33)	3 (20.00)	4 (26.67)	2 (13.33)	1 (6.67)	2.40
3	COMM7	Cultivates a research data management culture that adds value to the institution.	15	3 (20.00)	7 (46.67)	2 (13.33)	2 (13.33)	1 (6.67)	2.40
4	COMM3	Expresses strategic research data requirements reflecting organizational and management needs.	15	6 (40.00)	3 (20.00)	3 (20.00)	2 (13.33)	1 (6.67)	2.27
5	COMM6	Periodically reports on the value delivered by research data management.	15	5 (33.33)	6 (40.00)	2 (13.33)	1 (6.67)	1 (6.67)	2.13
6	CUST2	Complies with controls specified in the data management plan.	13	6 (46.15)	1 (7.69)	0 (0.00)	6 (46.15)	0 (0.00)	2.46
7	EXE2	Oversees the implementation of research data governance as a framework for good research data management.	15	4 (26.67)	6 (40.00)	2 (13.33)	2 (13.33)	1 (6.67)	2.33
8	IT2	Offers technical support for data management, including storage, backup, and archiving.	13	5 (38.46)	3 (23.08)	1 (7.69)	2 (15.38)	2 (15.38)	2.46
9	IT1	Provides a standardized, robust infrastructure for effective research data management.	13	5 (38.46)	3 (23.08)	1 (7.69)	3 (23.08)	1 (7.69)	2.38

Table 5.8, continued

NO	[a]TASK ID	TASK	[b]N	[c]RESPONSE (%)					[d]M
				1	2	3	4	5	
10	LEAD1	Ensures regular review and maintenance of research data governance policies in line with sector requirements.	15	4 (26.67)	5 (33.33)	3 (20.00)	2 (13.33)	1 (6.67)	2.40
11	LEAD2	Decides on day-to-day matters related to research data governance.	15	3 (20.00)	7 (46.67)	2 (13.33)	2 (13.33)	1 (6.67)	2.40
12	LEAD3	Directs decision-making to appropriate stakeholders when needed.	15	3 (20.00)	6 (40.00)	4 (26.67)	1 (6.67)	1 (6.67)	2.40
13	LEAD6	Communicates outcomes of the Research Data Governance Committee.	15	5 (33.33)	4 (26.67)	3 (20.00)	2 (13.33)	1 (6.67)	2.33
14	LEAD7	Serves as an expert on research data governance and recommends solutions.	15	5 (33.33)	5 (33.33)	2 (13.33)	2 (13.33)	1 (6.67)	2.27
15	LIB6	Reviews research data management policies with governance bodies.	13	4 (30.77)	4 (30.77)	3 (23.08)	1 (7.69)	1 (7.69)	2.31
16	LIB7	Offers training on metadata standards, data management plans, and legal/ethical issues.	13	4 (30.77)	4 (30.77)	2 (15.38)	3 (23.08)	0 (0.00)	2.31
17	LIB8	Advises on research data costs throughout the project lifecycle.	13	7 (53.85)	1 (7.69)	2 (15.38)	2 (15.38)	1 (7.69)	2.15
18	RO3	Facilitates and supports the implementation of research data governance processes.	13	4 (30.77)	3 (23.08)	3 (23.08)	2 (15.38)	1 (7.69)	2.46
19	RO6	Provides templates for incoming and outgoing research data agreements.	13	6 (46.15)	1 (7.69)	2 (15.38)	3 (23.08)	1 (7.69)	2.38

Notes.
[a]COMM-Research Data Governance Committee; CUST-Research Data Consumer; EXE-Executive Sponsor; IT-Information Technology Office; LEAD-Data Governance Leader; LIB-Library; RO-Research Management Office
[b]Different set of instrument having different number of panelists
[c]Score ranging from: 1-Not Implemented; 2-Slightly Implemented; 3-Moderately Implemented; 4-Highly Implemented; 5-Very Highly Implemented
[d]M-Mean
Information was sorted by Task ID and Mean score from the highest to the lowest

Analyzing the findings from a qualitative perspective reveals that 19 slightly implemented tasks resulted in a total of 35 activities: 12 activities fall under the 'define' construct, 15 activities under the 'implement' construct, and 8 activities under the 'monitor' construct. Furthermore, no tasks have been identified and listed under the 'slightly implemented tasks' category for the following six areas: (i) Data Ownership & IPR for Data Principle, (ii) Data Selection for Data Storage & Infrastructure, (iii) Data Citation for Metadata, (iv) Issue & Risk Management for Data Quality, (v) Data Privacy for Data Security, and (vi) Data Licensing for Data Access. This suggests that tasks within the aforementioned areas are either highly or moderately implemented by Malaysia

RPOs. Table 5.9 presents the code co-occurrence for tasks that are slightly implemented by RPOs in Malaysia.

Table 5.9: Total RDG activities vs. Slightly implemented RDG activities

Total Activities by 'Action' construct*			**Total Activities for Slightly Implemented Tasks by 'Action' construct****			**Total Activities for Slightly Implemented Tasks by Area *for* Decision Domain (Slightly Implemented activity**/Total activity*)**
Define Total=60	Implement Total =119	Monitor Total =47	Define Total=12	Implement Total =15	Monitor Total =8	
1	7	1	1	1	0	Communication *for* Data Principle (n=2/9)
2	16	12	0	2	0	Compliance Monitoring *for* Data Principle (n=2/30)
1	3	1				Data Ownership & IPR *for* Data Principle (n=0/5)
2	10	4	1	1	2	Data Policy *for* Data Principle (n=4/16)
15	8	6	4	1	2	Data Strategy *for* Data Principle (n=7/29)
2	3	3	1	1	0	Decision-making Coordination *for* Data Principle (n=2/8)
2	2	2	1	1	1	Performance Measurement *for* Data Principle (n=3/6)
4	8	0	1	3	0	Training *for* Data Principle (n=4/12)
3	4	1	0	1	0	Data Custodianship *for* Data Architecture (n=1/8)
8	9	1	2	1	1	Data Stewardship *for* Data Lifecycle (n=4/18)
2	1	1				Data Selection *for* Data Storage & Infrastructure (n=0/4)
5	8	1	0	2	0	Data Repository *for* Data Storage & Infrastructure (n=2/14)
2	6	1	0	1	0	Data Retention & Disposal *for* Data Storage & Infrastructure (n=1/9)
1	6	0				Data Citation *for* Metadata (n=0/7)
1	10	1	1	0	1	Data Integrity *for* Data Quality (n=2/12)
3	4	1				Issue & Risk Management *for* Data Quality (n=0/8)
1	8	4				Data Privacy *for* Data Security (n=0/13)
2	3	3	0	0	1	Data Sharing *for* Data Access (n=1/8)
3	3	4				Data Licensing *for* Data Access (n=0/10)

Note.
* The shaded rows signify that no tasks have been implemented in any 'Action' constructs within this category

Additionally, some of the panelists have provided some suggestions, comments or clarifications on certain questions to demonstrate the current implementation within their respective institutions, as indicated in the Table 5.10.

Table 5.10: Additional Suggestions/Comments/Clarifications provided by the Panelists on the Level of Implementation of RDG Tasks

Set	Task Statement	Suggestions/Comments/Clarifications
Set A (Strategic)	Ensures that research data governance policies are reviewed and maintained in accordance with sector requirements.	P34-Not yet but will be implemented.
	Communicates the Research Data Governance Committee outcomes.	P05-With the lack of data governance policies that transpire to all level of ministry, it is crucial the decision of the research data governance committee outcomes communicated with the rest.
	Signs off policy, supports appropriate cultural and behavioral changes, and allocates appropriate resources to research data governance activities	P05-It requires higher authority to oversee these changes and allocate appropriate resources.
	Provides advice, guidance, and reviews on research data management concerning research integrity and ethics.	P05-There are no other body that may advise and set standards to data management.
Set B (Tactical)	Store active research data securely and protect them from loss, unlawful or unethical access, and in accordance with all other applicable requirements.	P02-Data stewardship activities just started over a year ago, with less than 5 datasets being managed across the campus.
	Ensures any confidential data and material, including data and materials held in computing systems, are kept appropriately secure according to any applicable privacy laws.	P02-Based on the current deposition activities for research datasets, the data that have confidential issues are not shared by researchers.
	Ensures the ongoing custodial responsibilities for the research data upon researchers leaving the university.	P02-Deposition of research datasets just started, and the deposited datasets are all created by researchers whom all of them are still working at the University.
	Provides appropriate training and support to researchers with research data management, curation, or access queries and other relevant research data management regulations and procedures.	P02-Support is provided to researchers in terms of preparing Data Management Plan and advice on Data Management Policy, but training-wise is still in the planning stage.
	Examines the research data management plan and provides recommendations on its conformity with the research data governance policies.	P02-The University has Research Data Management Policy, but not research data governance policy.
	Arranges the availability of the necessary resources, facilities, and support for research data management in the research group.	P05- Since we don't have data stewards in our institution, research data are mainly managed by individual primary investigator.

Note.
* The exact wording of the task statement as presented in the questionnaire

In summary, while RDG is still in the early stage of implementation within RPOs in Malaysia, as indicated by P02 that in their institution "*Data stewardship activities just started over a year ago*", the results indicate that all governance tasks have been collectively initiated, although the implementation is not yet at its fullest. With that, all areas and decision domains identified from the desk research findings are covered, as

every governance task has been implemented. This underscores the potential for introducing and initiating RDM initiatives in individual RPO in Malaysia, considering their familiarity with specific RDG tasks.

5.4 Importance of Research Data Governance Activities: Insights from Data Practitioners

What research data governance activities do data practitioners consider significant for research performing organizations? To address this RQ2, the answers are derived from the findings of the modified Delphi study: Rounds I, II, and III. These findings will provide insights into the essential RDG tasks, areas, and decision domains deemed crucial for RPOs in Malaysia to establish effective RDM.

5.4.1 Findings for Round I

Following the questions on implementation, the focus shifted to indicating the level of importance for each governance task. The aim was to establish a consensus on which tasks associated with various governance roles are important and should be implemented by Malaysia RPOs, generally. The panelists could choose scale ranges from 1 (Not Important) to 5 (Very Highly Important) for this purpose. Statements on tasks achieving a priori consensus threshold, with group consensus level (CL) of 85 percent or higher, an interquartile deviation (IQD) of 0.5 or lower, and a median of 4 or higher were considered significant.

Earlier, the raw data was already exported from Cognito Forms to the Excel spreadsheet, and all the responses were converted into numeric values. Following that, descriptive statistics for the responses regarding the perceived importance were calculated in Excel to determine the percentage values indicative of the consensus level (CL). Additionally, RStudio was employed to calculate the median and interquartile range. The interquartile range represents the difference between the first quartile (the 25th percentile)

and the third quartile (the 75th percentile) of a dataset. In simple terms, it measures the spread of the middle 50 percent of values (IQR = Q3 – Q1). The built-in median() and IQR() functions in RStudio were used to compute the median and interquartile range of the values in the responses, as illustrated in Appendix V.

The value was then updated in the Excel spreadsheet and divided by two to obtain the interquartile deviation (IQD). IQD is defined as half of the distance between the third and the first quartile. The formula is IQD = Q3 – Q1 / 2. Following the execution of descriptive statistics, the Excel spreadsheet was appropriately updated. In Excel, the process enabled clear identification of items that achieved consensus and those that did not.

Overall, the results indicate a high level of consensus among the panelists regarding the importance of the RDG tasks assessed in this round. Out of 119 statements evaluated, 78 (65.5%) of them achieved the a priori consensus and 41 items (34.5%) were brought to the next round for re-evaluation.

Table 5.11 highlights among the tasks evaluated, 23 item statements pertaining to tasks across various RDG roles received unanimous agreement (CL=100%) on their importance, with the panelists rating them either as ‘highly important’ or ‘very highly important’. Among these, 18 tasks received the highest group consensus level with a median rating of 5 and an IQD of 0.5. This implies that most of the panelists consider these tasks to be of ‘very high importance’. Most of these tasks are linked to the Researcher role, with a few are attributed to the Research Data Steward role.

The tasks under the Researcher role that were rated as 'very highly important' by the majority of the panelists include tasks such as: *'R1-selecting data for long-term preservation based on verification, replication, and reuse needs', 'R6-handling intellectual property in research outputs as per obligations', 'R8-publishing research data in disciplinary, institutional, or established repositories', 'R9-guaranteeing the integrity, quality, security, and persistent availability of research data', 'R13-transferring data, materials, and records after projects or upon leaving the institution', 'R14-managing data throughout the research data lifecycle', 'R15-arranging secure storage for research data, records, and materials', 'R16-understanding the limitations and risks of third-party storage solutions', 'R17-protecting confidential, personal, and sensitive research data in accordance with legal and ethical requirements', 'R19-upholding open-source file formats and types recommended for preservation, 'R20-regularly backing up research data and records', 'R22- working with information security teams to ensure system controls for data protection', 'R23-identifying integrity and quality issues in research data',* and *'R24-periodically reviewing data access and usage agreements'*.

On the other hand, the 'very highly important' tasks attributed to the Research Data Steward role as unanimously agreed upon by the panelists, include *'DS5-ensuring that research data is archived for long-term preservation', 'DS6-ensuring that research data is licensed for reuse under permissible terms', 'DS7-ensuring ongoing custodial responsibilities for research data is maintained even after researchers leave the university',* and *'DS8-providing training and support for researchers in data management*'. This belief indicates that the panelists recognize the significance of these tasks in facilitating effective RDG. It suggests that there is a shared understanding among the panelists regarding the essential of these tasks in fostering ethical practices, ensuring responsible use, and safeguarding research data over the long term. This is particularly

noteworthy in relation to both roles, namely the Researcher and Research Data Steward, who play crucial roles in the production of research data and providing necessary support for RDM.

Table 5.11: Level of importance for research data governance tasks from Round I

NO.	[a]TASK ID	TASK	[b]N	[*]M	[*]IQD	[c*]CL
1	DS5	Ensures the archival of data for long-term preservation.	11	5	0.5	100%
2	DS6	Ensures data licensing for reuse and dissemination.	11	5	0.5	100%
3	DS7	Ensures management of custodial responsibilities for departing researchers.	11	5	0.5	100%
4	DS8	Provides training and support for data management.	11	5	0.5	100%
5	R1	Selects research data for long-term preservation based on verification/replication and reuse needs.	11	5	0.5	100%
6	R6	Protects intellectual property according to funder or contractual obligations.	11	5	0.5	100%
7	R8	Publishes data to established repositories, maximizing research value.	11	5	0.5	100%
8	R9	Guarantees integrity, quality, security, and persistent availability of research data.	11	5	0.5	100%
9	R13	Hands over data and materials after projects or leaving the institution.	11	5	0.5	100%
10	R14	Manages data throughout the lifecycle in line with policies, guidelines, and requirements.	11	5	0.5	100%
11	R15	Arranges safe and secure storage of research materials.	11	5	0.5	100%
12	R16	Understands and addresses risks of third-party storage solutions.	11	5	0.5	100%
13	R17	Protects confidential and sensitive data in line with legal and ethical requirements.	11	5	0.5	100%
14	R19	Upholds open-source file formats and types recommended for preservation.	11	5	0.5	100%
15	R20	Regularly backs up research data in accordance with best practices.	11	5	0.5	100%
16	R22	Works with the information security team to ensure system controls.	11	5	0.5	100%
17	R23	Identifies and addresses research data integrity and quality issues.	11	5	0.5	100%
18	R24	Periodically reviews data access and usage agreements.	11	5	0.5	100%
19	DS3	Ensures appropriate classification and management of research data based on their sensitivity.	11	4	0.5	100%
20	DS12	Monitors group's data management regulations.	11	4	0.5	100%
21	DS15	Ensures appropriate data access to project team members.	11	4	0.5	100%
22	DS16	Establishes clear data management responsibilities in the group.	11	4	0.5	100%
23	R21	Develops operating procedures to comply with research data governance.	11	4	0.5	100%
24	ORG2	Ensures support for grant holders in policy compliance.	47	4	0.5	96%
25	ORG6	Maintains institutional-level research data governance policies.	47	5	0.5	94%
26	EXE1	Provides necessary facilities and support for efficient research data management.	15	4	0.5	93%
27	LEAD4	Develops and maintains a central repository for governance policies, guiding principles, and decisions.	15	4	0.5	93%

Table 5.11, continued

NO.	[a]TASK ID	TASK	[b]N	[*]M	[*]IQD	[c*]CL
28	IT2	Offers technical support for data management, including storage, backup, and archiving.	13	5	0.5	92%
29	IT3	Provides secure access management following ICT security guidelines.	13	5	0.5	92%
30	LIB1	Coordinates a network of data stewards.	13	4	0.5	92%
31	RO1	Advocates organizational awareness of research data management issues.	13	4	0.5	92%
32	RO2	Ensures governance policies are updated based on the latest funder requirements and national directives.	13	4	0.5	92%
33	RO5	Provides advice, guidance, and assistance to researchers in preparing data management plans.	13	4	0.5	92%
34	ORG4	Safeguards researchers' rights, ensuring access to their data.	47	5	0.5	91%
35	ORG7	Enables planning and execution of good research data management practices.	47	5	0.5	91%
36	ORG9	Ensures ethical use of research elements like animals, human subjects, and materials.	47	5	0.5	91%
37	ORG12	Owns all research data and associated intellectual property.	47	5	0.5	91%
38	ORG13	Develops best practices for data sharing in different fields.	47	5	0.5	91%
39	ORG14	Provides researchers with training on effective research data management.	47	5	0.5	91%
40	ORG3	Manages an institutional metadata catalogue, especially for publicly funded research.	47	4	0.5	91%
41	ORG8	Collaborates with funders, policymakers, and stakeholders to align governance with sector requirements.	47	4	0.5	91%
42	GOV1	Facilitates resources and support for research data management.	11	5	0.5	91%
43	GOV2	Approves storage, disposal, and publication of research data.	11	5	0.5	91%
44	GOV4	Ensures effective communication on data management.	11	5	0.5	91%
45	GOV6	Ensures compliance by principal investigators.	11	5	0.5	91%
46	GOV7	Reports IT security incidents and data breaches.	11	5	0.5	91%
47	GOV8	Supervises adherence to regulations and procedures.	11	5	0.5	91%
48	DS1	Securely store and protect active research data to meet compliance requirements.	11	5	0.5	91%
49	DS17	Arranges resources and support for research data management in the group.	11	5	0.5	91%
50	R2	Retains research data securely for a period determined by the organization.	11	5	0.5	91%
51	R4	Ensures written agreements cover ownership, sharing, storage, and disposal of research data.	11	5	0.5	91%
52	R10	Disposes of data and materials securely.	11	5	0.5	91%
53	R12	Reports security breaches affecting research data.	11	5	0.5	91%
54	GOV3	Models responsible data management behavior.	11	4	0.5	91%
55	GOV5	Ensures the requirement of data management plans for grant applications.	11	4	0.5	91%
56	DS4	Ensures the security of confidential data in accordance with privacy laws.	11	4	0.5	91%
57	DS9	Ensures data retention as per organizational guidelines.	11	4	0.5	91%
58	DS13	Ensures the availability of data management plans for research projects.	11	4	0.5	91%

Table 5.11, continued

NO.	[a]TASK ID	TASK	[b]N	[*]M	[*]IQD	[c*]CL
59	R11	Includes a data access statement in publications.	11	4	0.5	91%
60	R18	Participates in training on research data management and contractual obligations.	11	4	0.5	91%
61	ORG10	Facilitates investigations into scientific misconduct or conflict of interest.	47	5	0.5	89%
62	ORG16	Offers facilities, advisory services, and resources for secure data storage and management.	47	5	0.5	89%
63	ORG17	Ensures storage facilities comply with legal and regulatory requirements.	47	5	0.5	89%
64	ORG5	Establishes a governance committee overseeing data governance implementation.	47	4	0.5	89%
65	FUN2	Reviews implementation of research data management plans.	24	4.5	0.5	88%
66	FUN1	Defines research data governance principles aligned with funding principles.	24	4	0.5	88%
67	FUN3	Specifies retention periods for significant research data.	24	4	0.5	88%
68	ORG1	Ensures research data availability for the research community.	47	4	0.5	87%
69	ORG11	Takes custody of data when necessary for appropriate access.	47	4	0.5	87%
70	EXE2	Oversees the implementation of research data governance as a framework for good research data management.	15	4	0.5	87%
71	ORG15	Ensures long-term stewardship based on institutional/national data infrastructure.	47	4	0.5	85%
72	ORG18	Acknowledges contributions of researchers who generate, preserve, and share key datasets.	47	4	0.5	85%
73	CUST1	Acknowledges data sources and adhere to access terms and conditions.	13	5	0.5	85%
74	LIB6	Reviews research data management policies with governance bodies.	13	4	0.5	85%
75	IT1	Provides a standardized, robust infrastructure for effective research data management.	13	4	0.5	85%
76	RO4	Organizes training events on research data management.	13	4	0.5	85%
77	RO6	Provides templates for incoming and outgoing research data agreements.	13	4	0.5	85%
78	CUST2	Complies with controls specified in the data management plan.	13	4	0.5	85%

Note:
[a]CUST-Research Data Consumer; DS-Research Data Steward; EXE-Executive Sponsor; FUN-National Funder; GOV-Research Data Governor; IT-Information Technology Office; LEAD-Data Governance Leader; LIB-Library; ORG-Organization; R-Researcher
RO-Research Management Office
[b]Different set of instrument having different number of panelists
[c]Percentages were rounded to the nearest whole number
[*]M – Median; IQD – Interquartile deviation; CL – Consensus Level
Information was sorted based on a priori criteria (M, IQD, CL)
The highlighted items reached the highest group consensus

However, there are panelists who view tasks like *'ORG15-ensuring long-term stewardship based on institutional/national data infrastructure', 'ORG18-acknowledging contributions of researchers who generate, preserve, and share key datasets', 'CUST1-acknowledging research data sources and adhering to access terms and conditions', 'LIB6-reviewing research data management policies with governance bodies', 'IT1-providing a standardized, robust infrastructure for effective RDM', 'RO4-organizing training events on RDM',* as well as *'RO6-providing templates for incoming and outgoing research data agreements',* and *'CUST2-complying with controls specified in the data management plan*' as lesser priorities with a CL of 85%. Nevertheless, it is important to note that collectively, these tasks fulfil the minimum a priori consensus criteria.

5.4.2 Findings for Round II

Moving on to Round II, the panelists were allowed to re-evaluate their ratings for the importance of task statements that had not reached a consensus in the initial round. This phase enabled the panelists to examine the group responses, allowing them to decide whether to uphold their original responses or amend them based on insights from the group's collective feedback. Similar to Round I, after the closure of the Round II survey, the raw data was exported from Cognito Forms to the same Excel file and sheet. The approach to obtaining descriptive statistics, including the median, interquartile deviation (IQD), and percentage values, remained consistent across both RStudio and Excel.

For Round II, invitations were extended to the same group of data practitioners (n=47) who participated in Round I. However, only 34 (72%) of them responded, including those who had the privilege of answering two sets of instruments, resulting in 37 received responses. During Round II, the remaining 41 items that did not reach consensus in Round I were re-evaluated. Out of the 37 who responded, seven of them (19%) have amended their responses, with at least one item being re-rated. Among the

suggestions/comments/clarifications of changing their responses from the previous round are influenced by several factors as outlined in Table 5.12. Meanwhile, most of the panelists maintained their previous responses, either because their rating was aligned with the majority group rating or because they wanted to adhere to it, believing it represented the best rating for the respective item.

Table 5.12: Panelists' Suggestions/Comments/Clarifications on changing Round I' responses

Set	Task Statement*	Suggestions/Comments/Clarifications
Set A (Strategic)	Ensures that research data governance policies are reviewed and maintained in accordance with sector requirements.	P05-Recent discussion on data governance within the ministry shows there are more room for improvement, hence the lack of policies proven the task is more difficult than I initially assume.
	Communicates the Research Data Governance Committee outcomes.	P05-With the lack of data governance policies that transpire to all level of ministry, it is crucial the decision of the research data governance committee outcomes communicated with the rest.
	Signs off policy, supports appropriate cultural and behavioral changes, and allocates appropriate resources to research data governance activities	P05-It requires higher authority to oversee these changes and allocate appropriate resources.
	Provides advice, guidance, and reviews on research data management concerning research integrity and ethics.	P05-There are no other body that may advise and set standards to data management.
Set B (Tactical)	Approves annual attestations to ensure researchers are aware of the relevant research data usage requirements.	P05-It is agreeable that higher body requires to coordinate proper use of research data.
	Makes decisions where standard research data guidelines are ambiguous or not applicable.	P05-The higher authority shall play this role.
	Assists the institution in the event of an external audit, including granting access to research data as required.	P05-The data stewards will ensure the proper data management plan is in place.
	Examines the research data management plan and provides recommendations on its conformity with the research data governance policies.	P05-Data steward will ensure proper plan in place and will be great help to the researchers and team.
	Develops and records procedures and processes for selection, collection, storage, use (including reuses), access, and retention of research data related to their research programme, including protection of essential records in the event of a natural disaster or other emergencies.	P05-Data steward can represent the secondary data usage better.
	Ensures that research data management requirements are costed in the research proposals.	P05-These requirements ~costing RDM in research proposal~ are proven crucial throughout the data management journey. Absent/lack thereof these will cause missed opportunity for the data to be reuse in the future.

Table 5.12, continued

Set	Task Statement*	Suggestions/Comments/Clarifications
Set B (Tactical)	Fosters a culture and training regime for research data awareness, including data usage agreements, disclosure process, and security controls.	P25-Less training available related to research data.
Set C (Operational)	The Library acts as the custodian of centrally managed research data collection metadata records by publishing sufficient appropriate metadata describing every shared dataset on a publicly available catalogue.	P21-Library has human resources with relevant knowledge.
	The Library provides high-quality infrastructure and facilities for collecting, storing, accessing, sharing, and archiving research data.	P21-Actually, to develop very high-quality infrastructure and facilities, it is up to the institutions/university/organization budget. But then how to manage that system keep on going utilized is the main point.
	Research Office facilitates, coordinates, and supports the execution of the processes in the research data governance policies.	P26-Since the library at our university lead the RDM initiatives, hence I think the library needs to facilitate, coordinates, and supports the execution of the processes in the research data governance policies.
	The Library provides training and advice on all aspects of research data management including minimum standards for metadata description, the writing of research data management plans, funders requirements, ownership, copyright and licenses, linked identifiers, data citation and legal and ethical issues related to the collection, storage, access, sharing and archiving of research data.	P30-It will not only the library, but it depends on the roles associated with particular offices. For example, storage, the IT Department need to provide training and advice in that aspect.
	The Library provides high-quality infrastructure and facilities for collecting, storing, accessing, sharing, and archiving research data.	P30-Since the library at our university need to develop and manage the research data repository, hence I change my response to very highly important.
Set D (Researcher)	National and/or institutional repository defines a preference for non-proprietary international and community standards that facilitate access, use, and interpretation of research data.	P43-Has not attained to very highly important scale due to limited repositories and lack of awareness of standards required.

Note.
* The exact wording of the task statement as presented in the questionnaire

After the conclusion of Round II, 28 (23.5%) additional items, successfully reached consensus among the panelists. This time, 11 of these items received unanimous agreement (CL=100%) from all panelists. Some panelists revised their initial responses and recognized that these items are either 'highly important' or 'very highly important'. Nevertheless, 13 items (10.9%) failed to achieve consensus and were subsequently

excluded from the list. '*Monitors compliance with the research data governance policies and supporting processes*' is the item that received the lowest rate (69%) where many panelists do not agree on the importance of this task which is under the responsibility of the Office of Research Data Governance. The exclusion of these items does not necessarily suggest their lack of importance; rather, it signifies that, at this point, they are deemed less critical for implementation within RPOs. Furthermore, the removal of these items has not impacted the coverage of areas and decision domains identified in the earlier desk research. The panelists' responses for Round II are summarized in Table 5.13.

Table 5.13: Level of importance for research data governance tasks from Round II

NO	[a]TASK ID	TASK	[b]N	[*]M	[*]IQD	[c*]CL
1	LIB2	Manages research data metadata records and publishes them on a public catalog.	10	5	0	100%
2	R7	Provides sufficient metadata for discoverability and reusability of research data.	6	5	0	100%
3	LIB3	Maintains a research data repository for registering data and other outputs.	10	5	0.5	100%
4	R5	Writes a comprehensive research data management plan.	6	5	0.5	100%
5	GOV9	Approves attestations for researcher awareness.	8	4.5	0.5	100%
6	GOV11	Makes decisions in ambiguous data guidelines.	8	4.5	0.5	100%
7	GOV13	Sponsors, secures, and/or influences resources for research data management.	8	4.5	0.5	100%
8	DS2	Assists in external audits and provides data access.	8	4.5	0.5	100%
9	DS10	Reviews and recommends improvements to data management plans.	8	4.5	0.5	100%
10	DS11	Develops processes for data selection, storage, and protection.	8	4.5	0.5	100%
11	LIB7	Offers training on metadata standards, data management plans and legal/ethical issues.	10	4.5	0.5	100%
12	IREP2	Specifies required contextual information and metadata for deposited data.	16	5	0.5	94%
13	LEAD1	Ensures regular review and maintenance of research data governance policies in line with sector requirements.	13	5	0.5	92%
14	DGO1	Signs off policies, supports cultural and behavioral changes, a allocates resources to research data governance activities.	13	5	0.5	92%
15	LEAD6	Communicates outcomes of the Research Data Governance Committee.	13	4	0.5	92%
16	DGO4	Provides advice, guidance, and reviews on research data management, emphasizing research integrity and ethics.	13	4	0.5	92%
17	LIB4	Defines the institution's support for research data management.	10	5	0.5	90%
18	LIB5	Provides high-quality infrastructure for data collection, storage, and sharing.	10	5	0.5	90%
19	LIB8	Advises on research data costs throughout the project lifecycle.	10	4.5	0.5	90%

Table 5.13, continued

NO	[a]TASK ID	TASK	[b]N	[*]M	[*]IQD	[c*]CL
20	GOV10	Promotes a culture of data awareness and training.	8	4.5	0.5	88%
21	GOV12	Sets expectations for data classification and retention.	8	4.5	0.5	88%
22	DS14	Ensures the inclusion of data management costs in research proposals.	8	4	0.5	88%
23	COMM1	Provides a common vocabulary for primary research data entities and types, essential to the organization.	13	5	0.5	85%
24	LEAD5	Maintains the Research Data Governance Committee agenda and convenes meetings.	13	4	0.5	85%
25	LEAD7	Serves as an expert on research data governance and recommends solutions.	13	4	0.5	85%
26	COMM4	Oversees initiatives to enhance effective and efficient research data utilization.	13	4	0.5	85%
27	COMM5	Supports research data management for institutional initiatives.	13	4	0.5	85%
28	COMM7	Cultivates a research data management culture that adds value to the institution.	13	4	0.5	85%
-1	R3	Budgets costs for capturing, managing, archiving, and sharing research data.	6	4.5	0.5	83%
-2	IREP1	Defines non-proprietary standards for access, use, and interpretation.	16	4	0.5	81%
-3	RO3	Facilitates and supports the implementation of research data governance processes.	10	4	0.5	80%
-4	LEAD2	Decides on day-to-day matters related to research data governance.	13	4	0.75	77%
-5	LEAD3	Directs decision-making to appropriate stakeholders when needed.	13	4	0.75	77%
-6	COMM2	Develops tools, guidelines, principles, and policies for research data, covering classification, access, usage, integrity, retention, roles, incident response, and integration.	13	4	0.75	77%
-7	COMM3	Expresses strategic research data requirements reflecting organizational and management needs.	13	4	0.75	77%
-8	COMM6	Periodically reports on the value delivered by research data management.	13	4	0.75	77%
-9	DGO2	Approves and implements processes supporting research data governance policies.	13	4	0.75	77%
-10	PUB1	Endorses research data deposits in established repositories.	16	4	0.75	75%
-11	FUN4	Provides advice directly or through data services.	16	4	1	75%
-12	CUST3	Enters into data use agreements for accessing embargoed research data.	10	4	1	70%
-13	DGO3	Monitors compliance with research data governance policies and supporting processes.	13	4	1	69%

Notes:
[a]COMM-Research Data Governance Committee; CUST-Research Data Consumer; DGO-Data Governance Office; DS-Research Data Steward; FUN-National Funder; GOV-Research Data Governor; LEAD-Data Governance Leader; LIB-Library; R-Researcher; RO-Research Management Office
[b]Different set of instrument having different number of panelists
[c]Percentages were rounded to the nearest whole number
[*]M – Median; IQD – Interquartile deviation; CL – Consensus Level
Information was sorted based on a priori criteria (M, IQD, CL)
The highlighted items in yellow reached CL=100
The highlighted items in grey did not reach a consensus and were eliminated

5.4.3 Findings for Round III

In Round III, the primary focus was on gaining broader insights into RDG activities. This involved determining the significance of specific areas and evaluating the importance of different decision domains for RDG. The Round III instrument was developed in accordance with the findings from the desk research and was informed by insights gained in previous rounds. In essence, the statements for each item that had achieved consensus in the earlier rounds were formulated to assess the perceived importance of RDG areas and decision domains.

Table 5.14 displays the results for Round III, providing descriptive analysis for each RDG area and decision domain derived from Excel and RStudio. It is worth noting that Research Data Integrity and Compliance Monitoring are areas that achieved unanimous agreement with a 100 percent CL, indicating a robust consensus on their critical importance. Similarly, Research Data Ownership & Intellectual Property Rights, Issue & Risk Management, Research Data Stewardship, Research Data Custodianship, Research Data Privacy, and Training received high approval, each securing a 97 percent CL, indicating a collective acknowledgement of their significance. The CL for Research Data Policy, Research Data Repository, and Research Data Sharing, all at 94 percent, emphasized a shared recognition of their importance in the RDG landscape. While, Research Data Strategy, Performance Measurement, Research Data Retention & Disposal, and Research Data Licensing, all with a 91 percent CL, indicate a shared acknowledgment of these areas in the RDG framework. Research Data Citation and Decision-Making Coordination, achieved an 88 percent CL each, signifying a notable level of agreement on its importance. Research Data Selection and Communication, each with an 85 percent CL, demonstrate a substantial but slightly less unanimous consensus.

Table 5.14: Level of importance for RDG areas and decision domains from Round III

NO	[a]ITEM ID	ITEM	N	[b]M	[b]IQD	[bc]CL
1	A14	Research Data Integrity	33	5	0.25	100%
2	A6	Compliance Monitoring	33	4	0.5	100%
3	A3	Research Data Ownership & IPR	33	5	0.5	97%
4	A5	Issue & Risk Management	33	5	0.5	97%
5	A7	Research Data Stewardship	33	5	0.5	97%
6	A8	Research Data Custodianship	33	5	0.5	97%
7	A12	Research Data Privacy	33	5	0.5	97%
8	A19	Training	33	5	0.5	97%
9	A1	Research Data Policy	33	5	0.5	94%
10	A10	Research Data Repository	33	5	0.5	94%
11	A16	Research Data Sharing	33	5	0.5	94%
12	A2	Research Data Strategy	33	5	0.5	91%
13	A4	Performance Measurement	33	4	0.5	91%
14	A11	Research Data Retention & Disposal	33	4	0.5	91%
15	A15	Research Data Licensing	33	4	0.5	91%
16	A13	Research Data Citation	33	4	0.5	88%
17	A18	Decision-Making Coordination	33	4	0.5	88%
18	A9	Research Data Selection	33	4	0.5	85%
19	A17	Communication	33	4	0.5	85%
1	D5	Research Data Security	33	5	0.25	100%
2	D2	Research Data Lifecycle	33	5	0.5	100%
3	D6	Metadata Management	33	5	0.5	100%
4	D1	Research Data Principle	33	4	0.5	100%
5	D4	Research Data Storage & Infrastructure	33	5	0	97%
6	D3	Research Data Architecture	33	5	0.5	94%
7	D7	Research Data Quality	33	5	0.5	94%
8	D8	Research Data Access	33	5	0.5	94%

Notes.
[a]A- RDG Area; D-RDG Decision Domain
[b]M – Median; IQD – Interquartile deviation; CL – Consensus Level
[c]Percentages were rounded to the nearest whole number

Meanwhile, within the RDG decision domain, Research Data Principle, Research Data Lifecycle, Research Data Security, and Metadata Management achieved a CL of 100 percent, indicating unanimous agreement on their fundamental importance. Additionally, Research Data Storage & Infrastructure received a strong CL approval rating of 97 percent, underlining its significance. The 94 percent CL for Research Data Architecture, Research Data Quality, and Research Data Access reveal a broad acknowledgement of these decision domains in the RDM landscape. These ratings collectively indicate a high degree of consensus regarding the importance of these principles and practices within the framework of RDG. The assigned importance ratings to various RDG areas and decision domains emphasize the need to prioritize tasks aimed at safeguarding data integrity,

ensuring compliance, and enhancing the efficient management of data within research settings.

5.5 Research-Derived Descriptions for Research Data Governance Areas and Decision Domains

The initial descriptions were derived from literature and web search, and these foundational descriptions were also included in the glossary of terms within the instruments, serving as a point of reference for the panelists.

The findings on panelists' assessment and evaluation of these initial descriptions of RDG areas and decision domains are presented in Table 5.15. During Round III, consensus was defined as unanimous agreement (100%) among all panelists. In instances where consensus was not attained due to having new suggestions or modifications from the panelists, the descriptions were revisited in Round IV for re-assessment.

The results for Round III indicate widespread acceptance and agreement with the majority of the initial descriptions. Out of the 27 descriptions related to RDG areas (n=18) and decision domains (n=9), two from RDG areas underwent revision by at least a single participant, who proposed new descriptions. These revisions were finalized in the subsequent round, allowing the panel to re-evaluate and decide whether to accept or reject the revised descriptions. Meanwhile, the remaining descriptions were endorsed as standard research descriptions through this Delphi study.

Table 5.15: Assessment of descriptions of RDG areas and decision domains in Round III

*ID	ITEM	INITIAL DESCRIPTION	ACCEPTED		REVISED		REVISED DESCRIPTION
			Count	%	Count	%	
A1	Research Data Policy	All procedures, guidelines, standards, and requirements that have been developed within an institution or partnership for the management, archiving and sharing of research data.	32	97%	1	3%	All policies, procedures, guidelines, standards, and requirements that have been developed within an institution or partnership for the management, archiving and sharing of research data.
A2	Research Data Strategy	How an organization improves specific business objectives by strategically using its research data as assets.	33	100%	-	-	-
A3	Research Data Ownership & IPR	The legal or moral rights that gives individuals, groups, or organizations the authority to determine storage, retention, disposal, publication, or licensing arrangements of ideas, inventions, and creative expressions.	33	100%	-	-	-
A4	Performance Measurement	Evaluating how well research data are managed and the value they deliver for users and other stakeholders.	33	100%	-	-	-
A5	Issue & Risk Management	The ability to identify and monitor risks (threats and opportunities), to plan and implement responses to those risks, and respond to other issues that affect the change initiative.	33	100%	-	-	-
A6	Compliance Monitoring	The observation of the environment to identify gaps between the actual operations, the internal policies and standards, and the requirements as they derive from external regulations, laws, and orders.	33	100%	-	-	-
A7	Research Data Stewardship	The oversight of the entire research data lifecycle, aiming to ensure that the right processes are put in place and that appropriate decisions are made to make research datasets Findable, Accessible, Interoperable and Reusable (FAIR).	33	100%	-	-	-

Table 5.15, continued

*ID	ITEM	INITIAL DESCRIPTION	ACCEPTED		REVISED		REVISED DESCRIPTION
			Count	%	Count	%	
A8	Research Data Custodianship	A set of activities required to ensure that important datasets are developed, maintained, and are accessible within their defined specifications.	33	100%	-	-	-
A9	Research Data Selection	The process of choosing research datasets, which are considered worth long-term preservation by a data archive.	33	100%	-	-	-
A10	Research Data Repository	A digital platform where research data may be stored and made available. This includes but is not limited to an institutional data repository, a discipline specific repository, a funder repository, or a commercial system.	33	100%	-	-	-
A11	Research Data Retention & Disposal	The long-term storage of research data and records after the completion of a research activity/project, for the purposes of meeting legal obligations or other purposes. Meanwhile, research data disposal describes how research data is destroyed.	33	100%	-	-	-
A12	Research Data Privacy	The rights and obligations of individuals and the organization with respect to the collection, use, retention, and disclosure of personal information.	33	100%	-	-	-
A13	Research Data Citation	A reference to research data for the purpose of credit attribution and facilitation of access.	33	100%	-	-	-
A14	Research Data Integrity	The maintenance of, and the assurance of the completeness, accuracy, and consistency of research data over its entire lifecycle from inconsistencies, accidental or malicious alteration or destruction.	33	100%	-	-	-
A15	Research Data Licensing	A license applied to research data or research datasets that clearly defines how and in which conditions it can be reused and guides future re-users.	33	100%	-	-	-
A16	Research Data Sharing	The process of taking any type of research data and making it available for other researchers to examine or use.	33	100%	-	-	-
A17	Communica-tion	Communication attempts to raise awareness of the research data governance program among stakeholders constantly.	33	100%	-	-	-

Table 5.15, continued

*ID	ITEM	INITIAL DESCRIPTION	ACCEPTED		REVISED		REVISED DESCRIPTION
			Count	%	Count	%	
A18	Decision-Making Coordination	The observation of the environment to identify gaps between the actual operations, the internal policies and standards, and the requirements as they derive from external regulations, laws, and orders.	33	100%	-	-	-
A19	Training	Sets of activities that are specifically designed to improve research data management skills thereby enhancing increased performance or productivity.	32	97%	1	3%	A set of activities designed to improve research data management skills, enhancing performance, productivity, capacity building, and knowledge transfer.
D1	Research Data Principle	Setting the boundaries for the intended uses of research data, the organization's standards for data quality, and ultimately the foundation for how users will access and interpret the data.	33	100%	-	-	-
D2	Research Data Lifecycle	The process through which research data flow from creation, to processing, analysis, preservation, distribution/sharing and re-use.	33	100%	-	-	-
D3	Research Data Architecture	A set of rules, policies, standards, and models that govern and define the type of data collected and how it is used, stored, managed, and integrated within an organization and its database systems.	33	100%	-	-	-
D4	Research Data Storage & Infrastructure	The technology, policies, standards, and human resources necessary to acquire, process, store, distribute, and improve utilization of research data, including the retention of research data using technology specifically developed to keep that data and have it as accessible as necessary.	33	100%	-	-	-
D5	Research Data Security	The protection of research data from loss, unauthorized access, and unauthorized modification. Security must be maintained while research data is both at rest and in transit.	33	100%	-	-	-

Table 5.15, continued

*ID	ITEM	INITIAL DESCRIPTION	ACCEPTED		REVISED		REVISED DESCRIPTION
			Count	%	Count	%	
D6	Metadata Management	A set of processes for determining structured information about the attributes of a dataset that enables the research data to be identified, retrieved, and managed over time.	33	100%	-	-	-
D7	Research Data Quality	A measure of the condition of research data based on factors such as accuracy, completeness, consistency, reliability and whether it is up to date.	33	100%	-	-	-
D8	Research Data Access	The degree to which the research data collected is made available to individuals, groups and/or the public.	33	100%	-	-	-

Notes.
*A- RDG Areas; D-RDG Decision Domains

The panelists were given another opportunity in Round IV to re-evaluate responses for items or statements that had not received a group consensus during Round III, aiming to establish a confirmed and definitive consensus. The earlier two items related to the descriptions of RDG areas were brought into this round for re-evaluation, as new revised descriptions were available. Since this was a concluding round, consensus was determined by the majority agreement among panelists, who either endorsed the revised descriptions or opted to retain the initial descriptions, reflecting the peak of the collaborative process.

Based on the results, the panelists achieved a majority agreement on the revised descriptions for all RDG areas listed in this round. As a result, these descriptions, along with those that obtained unanimous agreement in the previous round, were officially endorsed as standard descriptions within the scope of this research. The results are presented in Table 5.16:

Table 5.16: Re-assessment of definitions pertaining to RDG areas and decision domains in Round IV

NO.	ITEM ID	REVISED DESCRIPTION	USED REVISED DESCRIPTION		RETAINED INITIAL DESCRIPTION	
			Count	Percentage	Count	Percentage
1	A1	All policies, procedures, guidelines, standards, and requirements that have been developed within an institution or partnership for the management, archiving and sharing of research data.	15	57.7%	11	42.3%
2	A19	A set of activities designed to improve research data management skills, enhancing performance, productivity, capacity building, and knowledge transfer.	19	73.1%	7	26.9%

In conclusion, the provision of standardized descriptions for each element of RDG activities, derived from the Delphi study, is crucial. Through continuous refinement, these descriptions will contribute to establishing a foundation for developing the RDG framework.

5.6 Gap Analysis: Perceived Importance vs. Implementation

Since there were data available on the actual implementation and perceived importance of RDG activities, the connection between both aspects was then analyzed in more detail. Therefore, a gap analysis was conducted to examine these activities by measuring the differences between panelists' perceived importance and actual implementation. The hypotheses are presented as follows:

- Null Hypothesis (H_0): There is no significant difference between the actual implementation and perceived importance of RDG tasks among data practitioners within RPOs.
- Alternative Hypothesis (H_1): There is a significant difference between the actual implementation and perceived importance of RDG tasks among data practitioners within RPOs.

The data were derived from the responses received in Round I, and the sample size varies based on different groups of panelists. Figure 5.1 displays mean comparisons between the perceived importance and actual implementation of RDG tasks among data practitioners in Malaysia RPOs, as responded by 47 panelists. The panelists' perceived importance means range from 4.170 to 4.489, with an overall mean of 4.3. Among the 18 tasks for the Organization role, twelve of them (67%) show higher mean values than the overall mean, including ORG2, ORG3, ORG4, ORG6, ORG7, ORG8, ORG10, ORG12, ORG13, ORG14, ORG16, and ORG17. ORG4 receives the highest mean rating for perceived importance (M=4.489), followed by ORG6 and ORG7 (M=4.468), ORG14 (M=4.447), and ORG2 (M=4.404). Tasks such as ORG5 (M=4.255), ORG15 & ORG18 (M=4.191), and ORG1 (M=4.170) receive relatively lower importance mean scores.

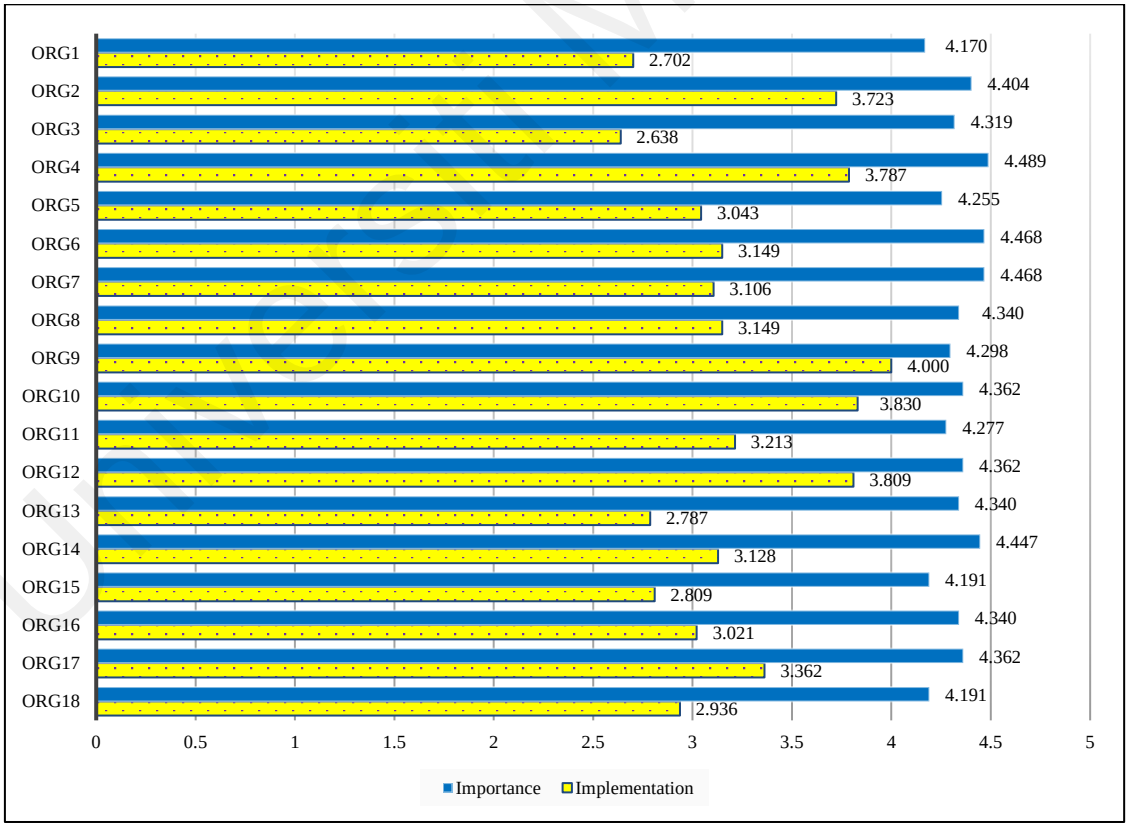


Figure 5.1: Comparisons between importance and implementation means for RDG tasks related to the Organization Role

On the other hand, panellists' actual implementation means scores range from 2.638 to 4. The overall mean for actual implementation is 3.2, notably lower than the overall mean for perceived importance. The highest ratings are for ORG9 (M=4), followed by ORG10 (M=3.830), and ORG12 (M=3.809). Conversely, lower scores are observed for ORG15 (M=2.809), ORG13 (M=2.787), ORG1 (M=2.702), and ORG3 (M=2.638).

The paired-sample t-test was conducted to analyze the gap scores between the perceived importance and actual implementation for each individual RDG task. The results of the paired-sample t-test, focusing on the perceived importance and actual implementation of RDG tasks associated with the Organization role, are presented in Table 5.17. All these 18 RDG tasks exhibit lower mean ratings for actual implementation than perceived importance, resulting in negative mean differences. In the paired-sample t-test, statistically significant mean differences are observed in 17 tasks, except for ORG9 (M.D=-0.298, t=-2.003, p=0.051), which has a slightly larger p-value just above 0.05. The tasks with the highest negative mean differences are ORG3 (M.D=-1.681, t=-8.321), followed by ORG13 (M.D=-1.553, t=-8.417), ORG1 (M.D=-1.468, t=-8.422), and ORG15 (M.D=-1.383, t=-6.890). These areas are identified as focal points that necessitate further attention for improvement.

In Figure 5.2, mean comparisons are presented illustrating the perceived importance and actual implementation of RDG tasks associated with the roles such as Executive Sponsor (EXE), Data Governance Leader (LEAD), Research Data Governance Committee (COMM), and Data Governance Office (DGO), as outlined in the survey instrument: Set A (Strategic). This information is based on the responses from 15 panelists who participated in this set.

Table 5.17: Gap analysis between perceived importance and actual implementation of RDG tasks related to the organization role

TASK ID	TASK	Implementation			Importance			Paired-sample t-test		
		M	S.D.	Rank	M	S.D.	Rank	M.D	t-value	Sig.
ORG1	Ensures research data availability for the research community.	2.702	1.121	16	4.170	0.868	12	-1.468	-8.422	*0.000
ORG2	Ensures support for grant holders in policy compliance.	3.723	1.210	5	4.404	0.648	4	-0.681	-4.468	*0.000
ORG3	Manages an institutional metadata catalogue, especially for publicly funded research.	2.638	1.342	17	4.319	0.783	7	-1.681	-8.321	*0.000
ORG4	Safeguards researchers' rights, ensuring access to their data.	3.787	1.284	4	4.489	0.718	1	-0.702	-3.907	*0.000
ORG5	Establishes a governance committee overseeing data governance implementation.	3.043	1.197	11	4.255	0.820	10	-1.213	-6.845	*0.000
ORG6	Maintains institutional-level research data governance policies.	3.149	1.367	8	4.468	0.654	2	-1.319	-7.218	*0.000
ORG7	Enables planning and execution of good research data management practices.	3.106	1.371	10	4.468	0.654	2	-1.362	-7.522	*0.000
ORG8	Collaborates with funders, policymakers, and stakeholders to align governance with sector requirements.	3.149	1.285	8	4.340	0.760	6	-1.191	-6.857	*0.000
ORG9	Ensures ethical use of research elements like animals, human subjects, and materials.	4.000	1.198	1	4.298	1.020	8	-0.298	-2.003	**0.051
ORG10	Facilitates investigations into scientific misconduct or conflict of interest.	3.830	1.148	2	4.362	0.919	5	-0.532	-3.925	*0.000
ORG11	Takes custody of data when necessary for appropriate access.	3.213	1.429	7	4.277	0.852	9	-1.064	-5.878	*0.000
ORG12	Owns all research data and associated intellectual property.	3.809	1.245	3	4.362	0.845	5	-0.553	-3.331	*0.002
ORG13	Develops best practices for data sharing in different fields.	2.787	1.301	15	4.340	0.841	6	-1.553	-8.417	*0.000
ORG14	Provides researchers with training on effective research data management.	3.128	1.329	9	4.447	0.746	3	-1.319	-7.026	*0.000
ORG15	Ensures long-term stewardship based on institutional/national data infrastructure.	2.809	1.209	14	4.191	0.924	11	-1.383	-6.890	*0.000
ORG16	Offers facilities, advisory services, and resources for secure data storage and management.	3.021	1.277	12	4.340	0.915	6	-1.319	-7.026	*0.000
ORG17	Ensures storage facilities comply with legal and regulatory requirements.	3.362	1.358	6	4.362	0.792	5	-1.000	-5.334	*0.000
ORG18	Acknowledges contributions of researchers who generate, preserve, and share key datasets.	2.936	1.480	13	4.191	0.924	11	-1.255	-6.260	*0.000

Notes.

1. Paired-sample t-test with probability *p < 0.05 and **p > 0.05.
2. Mean: mean scores were measured on a 5-point scale; the higher the score, the greater implementation or importance of RDG task.
3. S.D.: standard deviation.
4. Ranking was based on the highest to the lowest mean
5. Mean difference (M.D): implementation mean – importance mean

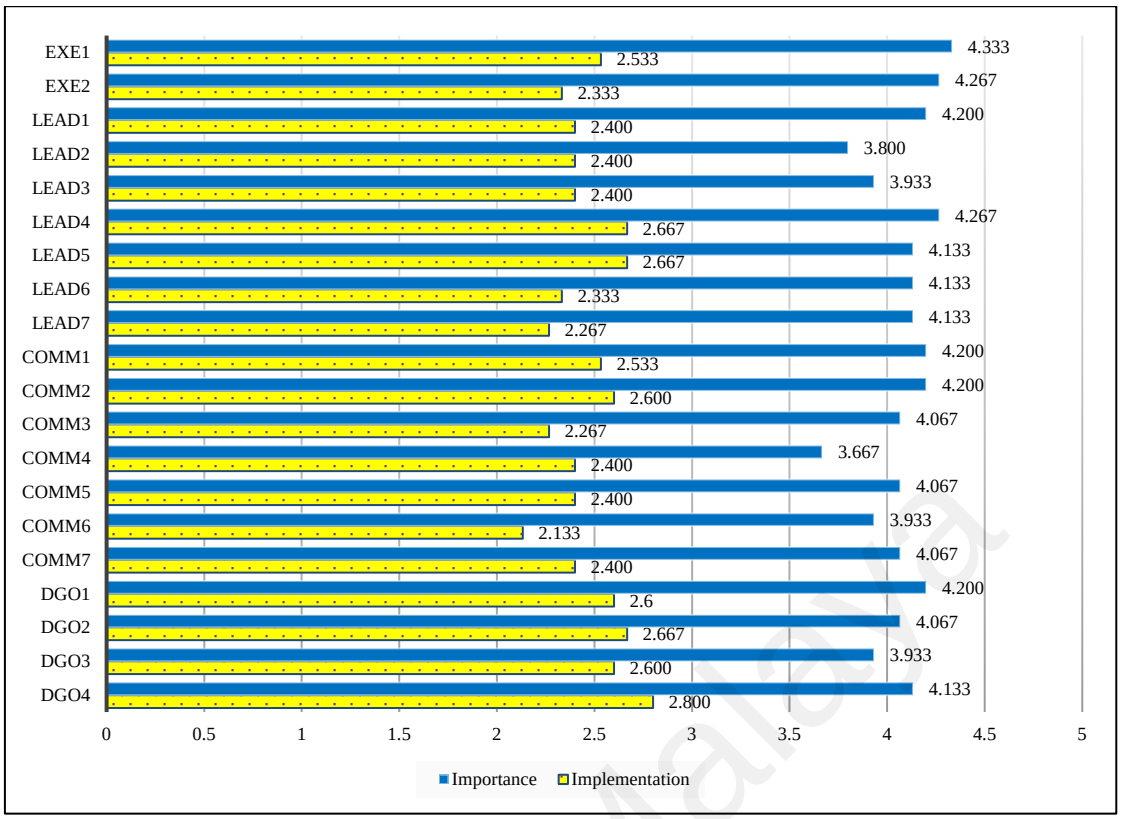


Figure 5.2: Comparisons between importance and implementation means for RDG tasks related to the roles queried in Set A (Strategic)

The panelists' perceived importance means range from 3.667 to 4.333, with an overall mean of 4.09. Out of the 20 tasks evaluated, eleven showed mean values for perceived importance that exceeded the overall mean. These tasks include EXE1, EXE2, LEAD1, LEAD4, LEAD5, LEAD6, LEAD7, COMM1, COMM2, DGO1, and DGO4. Notably, EXE1 receives the highest mean rating (M=4.333), followed by EXE2 and LEAD4 (M=4.267), as well as LEAD1, COMM1, COMM2, and DGO1 (M=4.200). Conversely, tasks such as LEAD3, COMM6, and DGO3 (M=3.933), LEAD2 (M=3.800), and COMM4 (M=3.667) receive relatively lower importance mean scores.

On the other hand, the mean scores for the actual implementation range relatively low, spanning from 2.133 to 2.800. The overall mean for actual implementation is 2.47, notably much lower than the overall mean for perceived importance. Nine tasks obtain mean scores higher than the overall mean. These tasks include EXE1, LEAD4, LEAD5, COMM1, COMM2, DGO1, DGO2, DGO3, and DGO4. The highest ratings are observed

for DGO4 (M=2.800), followed by LEAD4, LEAD5, and DGO2 (M=2.667), as well as COMM2, DGO1, and DGO3 (M=2.600). Conversely, lower scores are noted for EXE2 and LEAD6 (M=2.333), LEAD7 and COMM3 (M=2.267), and COMM6 (M=2.133).

Meanwhile, the results presented in Table 5.18 underscore the outcomes of the paired-sample t-test, revealing the disparity between perceived importance and the actual implementation of RDG tasks associated with the roles specified in Set A. For all 20 RDG tasks aligned with these roles, the mean ratings for actual implementation consistently fall below those for perceived importance, as indicated by the negative mean differences across the board. Notably, statistically significant mean differences are evident in every task, with each task exhibiting higher mean ratings for perceived importance than for actual implementation. Tasks with the most pronounced negative mean differences are EXE2 (M.D=-1.933, t=-4.882), LEAD7 (M.D=-1.867, t=-4.802), and EXE1 and LEAD1 (M.D=-1.800, t=-5.077), LEAD6 (M.D=-1.800, t=-4.583), COMM3 (M.D=-1.800, t=-4.006), and COMM6 (M.D=-1.800, t=-4.447), emerge as priority areas for improvement.

Table 5.18: Gap analysis between perceived importance and actual implementation of RDG tasks related to the roles queried in Set A (Strategic)

TASK ID	TASK	Implementation			Importance			Paired-sample t-test		
		M	S.D.	Rank	M	S.D.	Rank	M.D	t-value	Sig.
EXE1	Provides necessary facilities and support for efficient research data management.	2.533	1.125	1	4.333	0.816	1	-1.800	-5.077	*0.000
EXE2	Oversees the implementation of research data governance as a framework for good research data management.	2.333	1.234	2	4.267	0.884	2	-1.934	-4.882	*0.000
LEAD1	Ensures regular review and maintenance of research data governance policies in line with sector requirements.	2.400	1.242	2	4.200	0.941	2	-1.800	-5.077	*0.000
LEAD2	Decides on day-to-day matters related to research data governance.	2.400	1.183	2	3.800	1.082	5	-1.400	-3.609	*0.003
LEAD3	Directs decision-making to appropriate stakeholders when needed.	2.400	1.121	2	3.933	1.033	4	-1.533	-4.075	*0.001
LEAD4	Develops and maintains a central repository for governance policies, guiding principles, and decisions.	2.667	1.234	1	4.267	0.799	1	-1.600	-4.413	*0.001
LEAD5	Maintains the Research Data Governance Committee agenda and convenes meetings.	2.667	1.447	1	4.133	0.915	3	-1.467	-3.372	*0.005
LEAD6	Communicates outcomes of the Research Data Governance Committee.	2.333	1.291	3	4.133	0.915	3	-1.800	-4.583	*0.000
LEAD7	Serves as an expert on research data governance and recommends solutions.	2.267	1.280	4	4.133	0.915	3	-1.867	-4.802	*0.000
COMM1	Provides a common vocabulary for primary research data entities and types, essential to the organization.	2.533	1.246	2	4.200	0.941	1	-1.667	-3.953	*0.001
COMM2	Develops tools, guidelines, principles, and policies for research data, covering classification, access, usage, integrity, retention, roles, incident response, and integration.	2.600	1.183	1	4.200	0.775	1	-1.600	-3.886	*0.002
COMM3	Expresses strategic research data requirements reflecting organizational and management needs.	2.267	1.335	4	4.067	0.961	2	-1.800	-4.006	*0.001
COMM4	Oversees initiatives to enhance effective and efficient research data utilization.	2.400	1.454	3	3.667	1.175	4	-1.267	-3.106	*0.008
COMM5	Supports research data management for institutional initiatives.	2.400	1.298	3	4.067	0.884	2	-1.667	-4.315	*0.001

Table 5.18, continued

TASK ID	TASK	Implementation			Importance			Paired-sample t-test		
		M	S.D.	Rank	M	S.D.	Rank	M.D	t-value	Sig.
COMM6	Periodically reports on the value delivered by research data management.	2.133	1.187	5	3.933	1.033	3	-1.800	-4.447	*0.001
COMM7	Cultivates a research data management culture that adds value to the institution.	2.400	1.183	3	4.067	0.884	2	-1.667	-4.315	*0.001
DGO1	Signs off policies, supports cultural and behavioral changes, and allocates resources to research data governance activities.	2.600	1.183	3	4.200	1.014	1	-1.600	-4.000	*0.001
DGO2	Approves and implements processes supporting research data governance policies.	2.667	1.234	2	4.067	0.961	3	-1.400	-3.500	*0.004
DGO3	Monitors compliance with research data governance policies and supporting processes.	2.600	1.298	3	3.933	1.100	4	-1.333	-3.081	*0.008
DGO4	Provides advice, guidance, and reviews on research data management, emphasizing research integrity and ethics.	2.800	1.373	1	4.133	0.915	2	-1.333	-3.162	*0.007

Notes.
1. Paired-sample t-test with probability *p < 0.05 and **p > 0.05.
2. Mean: mean scores were measured on a 5-point scale; the higher the score, the greater implementation or importance of RDG task.
3. S.D.: standard deviation.
4. Ranking was based on the highest to the lowest mean.
5. Mean difference (M.D): implementation mean – importance mean.

Panelists (n=11) from Set B (Tactical) offered insights into the disparity between actual implementation and perceived importance for tasks associated with the roles of Research Data Governor (GOV) and Research Data Steward (DS), involving a total of 13 and 17 tasks, respectively. In terms of the Research Data Governor tasks (Figure 5.3), the perceived importance is consistently high, with mean scores ranging from 4.182 to 4.545. The overall mean for perceived importance is 4.35, indicating a generally elevated level of significance. Eight tasks (GOV1 to GOV8) surpass the overall mean, with the highest mean scores recorded for GOV1 and GOV8 at 4.545. However, when examining the actual implementation of RDG tasks, the mean scores are notably lower, falling within the range of 2.818 to 3.727. The overall mean for actual implementation is 3.200. Despite five tasks (GOV1, GOV2, GOV3, GOV7, and GOV8) scoring above the overall mean for

implementation, the discrepancy between perceived importance and actual implementation is evident. Tasks with lower scores in actual implementation include GOV9 (M=2.818), GOV10 (M=2.909), GOV4 and GOV12 (M=3.091), GOV5, GOV6, GOV11, and GOV13 (M=3.182).

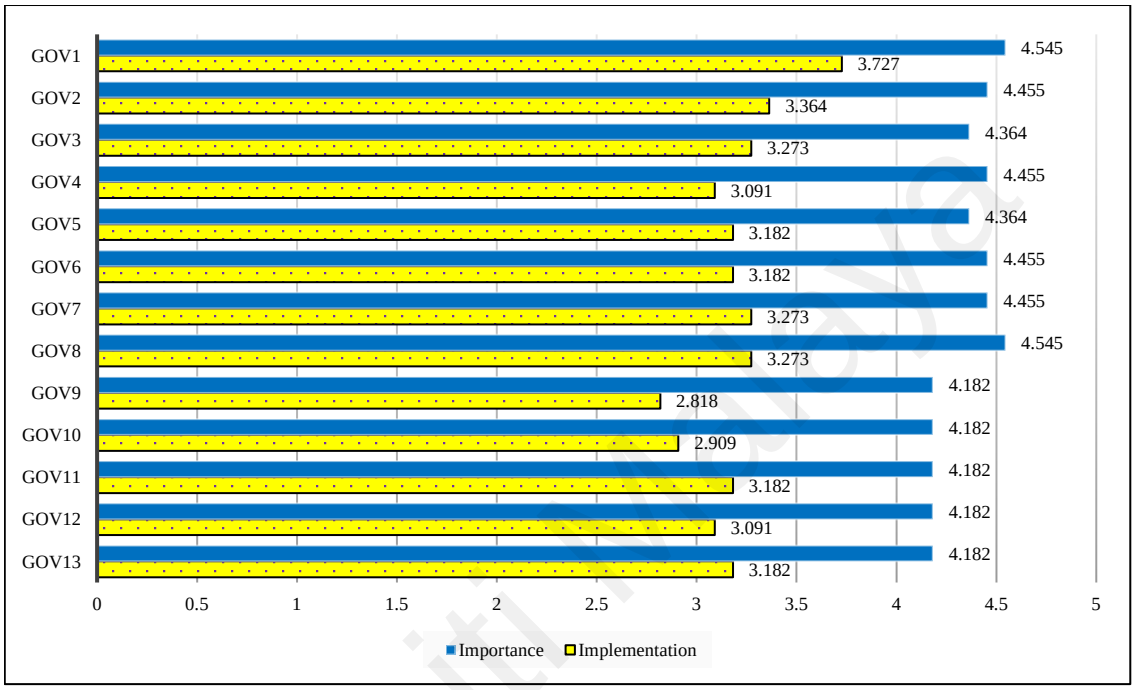


Figure 5.3: Comparisons between importance and implementation means for RDG tasks related to the Research Data Governor role

Meanwhile, for the role of Research Data Steward, the perceived importance mean scores range between 4.091 and 4.636 (Figure 5.4), resulting in an overall mean of 4.42. Out of the total tasks associated with the Research Data Steward, 59 percent (n=10) surpass the overall mean score. These tasks include DS1, DS3, DS5, DS6, DS7, DS8, DS12, DS15, DS16, and DS17. Notably, DS5 and DS7 (M=4.636) have the highest mean scores of perceived importance, while DS14 has the lowest mean score at 4.091.

In contrast, for the actual implementation, the mean scores are lower than the mean scores for the perceived importance. The mean scores range from 2.818 to 3.818, with an overall mean of 3.36. Only 53 percent (n=9) of tasks are beyond the overall mean (DS1, DS3, DS5, DS6, DS7, DS8, DS9, DS11, and DS17). The tasks with higher mean scores

for actual implementation are DS5 (M=3.818), DS3, and DS6 (M=3.636), as well as DS8 and DS11 (M=3.545). In contrast, tasks with lower mean scores include DS10 (M=3.182), DS2, and DS13 (M=3.091), along with DS14 (M=2.818).

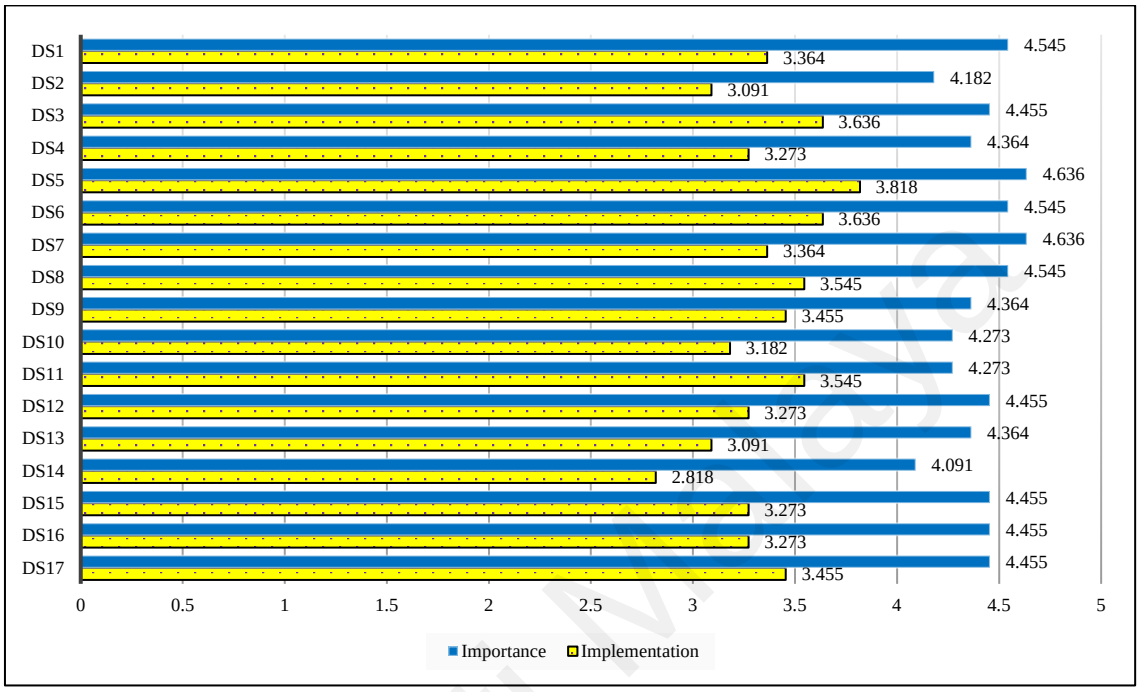


Figure 5.4: Comparisons between importance and implementation means for RDG tasks related to the Research Data Steward role

The perceived importance and actual implementation of RDG tasks for both roles reveal a considerable gap. According to the paired-sample t-test, as shown in Table 5.19, all tasks exhibit negative mean differences. Notably, three tasks associated with the Research Data Steward role—DS3, DS5, and DS6—do not show significant p-values, with M.D=-0.818, t=-2.043, p=0.068 for DS3 and DS5, and M.D=-0.909, t=-2.193, p=0.053 for DS6. In contrast, the remaining tasks, excluding these three, demonstrate a significant difference between actual implementation and perceived importance.

This discrepancy highlights a notable gap in implementation, suggesting that, although the RDG tasks for Research Data Governor and Research Data Steward roles are recognized as highly important, their implementation is not reflecting the importance adequately. As a result, certain tasks for the Research Data Governor role, marked by

higher mean differences, highlight specific areas that need improvement. These include GOV4 (M.D=-1.364, t=-4.404), GOV6 (M.D=-1.273, t=-3.545), GOV8 (M.D=-1.273, t=-3.825), GOV9 (M.D=-1.364, t=-4.038), and GOV10 (M.D=-1.273, t=-3.825). Similarly, for the Research Data Steward, tasks such as DS7 (M.D=-1.273, t=-2.715), DS13 (M.D=-1.273, t=-3.545), DS14 (M.D=-1.273, t=-3.131), as well as DS1 (M.D=-1.182, t=-2.952), DS12, DS15, and DS16 (M.D=-1.182, t=-2.797) exhibit higher mean differences, pinpointing specific aspects that necessitate targeted enhancements.

Table 5.19: Gap analysis between perceived importance and actual implementation of RDG tasks by the Research Data Governor and Research Data Steward roles

TASK ID	TASK	Implementation			Importance			Paired-sample t-test		
		M	S.D.	Rank	M	S.D.	Rank	M.D	t-value	Sig.
GOV1	Facilitates resources and support for research data management.	3.727	1.191	1	4.545	0.688	1	-0.818	-3.615	*0.005
GOV2	Approves storage, disposal, and publication of research data.	3.364	1.206	2	4.455	0.688	2	-1.091	-3.833	*0.003
GOV3	Models responsible data management behavior.	3.273	1.489	3	4.364	0.674	3	-1.091	-2.963	*0.014
GOV4	Ensures effective communication on data management.	3.091	1.375	5	4.455	0.688	2	-1.364	-4.404	*0.001
GOV5	Ensures the requirement of data management plans for grant applications.	3.182	1.662	4	4.364	0.674	3	-1.182	-2.952	*0.015
GOV6	Ensures compliance by principal investigators.	3.182	1.601	4	4.455	0.688	2	-1.273	-3.545	*0.005
GOV7	Reports IT security incidents and data breaches.	3.273	1.489	3	4.455	0.688	2	-1.182	-3.634	*0.005
GOV8	Supervises adherence to regulations and procedures.	3.273	1.618	3	4.545	0.688	1	-1.273	-3.825	*0.003
GOV9	Approves attestations for researcher awareness.	2.818	1.662	7	4.182	1.250	4	-1.364	-4.038	*0.002
GOV10	Promotes a culture of data awareness and training.	2.909	1.514	6	4.182	1.250	4	-1.273	-3.825	*0.003
GOV11	Makes decisions in ambiguous data guidelines.	3.182	1.537	4	4.182	0.982	4	-1.000	-2.622	*0.026
GOV12	Sets expectations for data classification and retention.	3.091	1.514	5	4.182	0.982	4	-1.091	-2.782	*0.019
GOV13	Sponsors, secures, and/or influences resources for research data management.	3.182	1.401	4	4.182	0.982	4	-1.000	-2.345	*0.041
DS1	Securely store and protect active research data to meet compliance requirements.	3.364	1.629	5	4.545	0.688	2	-1.182	-2.952	*0.015
DS2	Assists in external audits and provides data access.	3.091	1.758	8	4.182	1.250	6	-1.091	-3.185	*0.010
DS3	Ensures appropriate classification and management of research data based on their sensitivity.	3.636	1.433	2	4.455	0.522	3	-0.818	-2.043	**0.068
DS4	Ensures the security of confidential data in accordance with privacy laws.	3.273	1.618	6	4.364	0.674	4	-1.091	-2.782	*0.019
DS5	Ensures the archival of data for long-term preservation.	3.818	1.168	1	4.636	0.505	1	-0.818	-2.043	**0.068

Table 5.19, continued

TASK ID	TASK	Implementation			Importance			Paired-sample t-test		
		M	S.D.	Rank	M	S.D.	Rank	M.D	t-value	Sig.
DS6	Ensures data licensing for reuse and dissemination.	3.636	1.286	2	4.545	0.522	2	-0.909	-2.193	**0.053
DS7	Ensures management of custodial responsibilities for departing researchers.	3.364	1.690	5	4.636	0.505	1	-1.273	-2.715	*0.022
DS8	Provides training and support for data management.	3.545	1.368	3	4.545	0.522	2	-1.000	-2.622	*0.026
DS9	Ensures data retention as per organizational guidelines.	3.455	1.368	4	4.364	0.674	4	-0.909	-2.319	*0.043
DS10	Reviews and recommends improvements to data management plans.	3.182	1.250	7	4.273	0.786	5	-1.091	-3.185	*0.010
DS11	Develops processes for data selection, storage, and protection.	3.545	1.368	3	4.273	0.786	5	-0.727	-2.026	**0.070
DS12	Monitors group's data management regulations.	3.273	1.618	6	4.455	0.522	3	-1.182	-2.797	*0.019
DS13	Ensures the availability of data management plans for research projects.	3.091	1.578	8	4.364	0.674	4	-1.273	-3.545	*0.005
DS14	Ensures the inclusion of data management costs in research proposals.	2.818	1.401	9	4.091	0.831	7	-1.273	-3.131	*0.011
DS15	Ensures appropriate data access to project team members.	3.273	1.618	6	4.455	0.522	3	-1.182	-2.797	*0.019
DS16	Establishes clear data management responsibilities in the group.	3.273	1.618	6	4.455	0.522	3	-1.182	-2.797	*0.019
DS17	Arranges resources and support for research data management in the group.	3.455	1.440	4	4.455	0.688	3	-1.000	-2.803	*0.019

Notes.
1. Paired-sample t-test with probability *p < 0.05 and **p > 0.05.
2. Mean: mean scores were measured on a 5-point scale; the higher the score, the greater implementation or importance of RDG task.
3. S.D.: standard deviation.
4. Ranking was based on the highest to the lowest mean.
5. Mean difference (M.D): implementation mean – importance mean.

The discussion now shifts to the results of the comparison between mean scores for the perceived importance and actual implementation of RDG tasks (n=27) associated with the Administrative Offices, including the Library (LIB), Information Technology Office (IT), and Research Management Office (RO) in Set C (Operational). Additionally, the roles of Research Data Consumer (CUST) and External Bodies, comprising the Funder (FUN), National Publisher (PUB), and National/Institutional Repository (IREP), were also examined. All these roles were queried in Set C (Operational), with responses

obtained from 13 panelists. However, it is important to note that questions related to External Bodies were also answered by panelists who participated in Set D (Researcher), expanding the sample size to include these individuals (n=11). Thus, it results in a total of 24 panelists.

The mean scores for the perceived importance of all tasks related to the mentioned roles range from 3.846 to 4.462, with an overall mean of 4.15. Notably, fifteen tasks (56%) have surpassed the overall mean, including LIB1, LIB2, LIB6, IT1, IT2, IT3, RO1, RO2, RO4, RO5, CUST1, FUN1, FUN3, IREP1, and IREP2. Tasks with the highest perceived importance are led by IT3 (M=4.462), followed by IT2 (M=4.385), LIB1, IT1, RO1, and RO2 (M=4.308). On the other hand, tasks with lower mean scores include CUST2, RO3, RO6, and LIB7 (M=4.077), LIB5, LIB3, LIB4, and FUN4 (M=4.000), CUST3 (M=3.923), and LIB8 (M=3.846).

Meanwhile, for the level of implementation, the tasks associated with these roles exhibit mean scores ranging from 2.154 to 3.154, with an overall mean of 2.63—considerably lower than their perceived importance. Ten tasks have surpassed this overall mean: CUST1 (M=3.154), FUN2 (M=3.083), FUN1 and FUN3 (M=3.042), IREP2 (M=2.875), RO1 (M=2.846), RO2 (M=2.769), IREP1 and PUB1 (M=2.708), and FUN4 (M=2.667). Conversely, tasks with mean scores below the overall mean include LIB1, LIB3, LIB4, and CUST3 (M=2.615), IT3, LIB2, RO5, RO4, CUST2, and LIB5 (M=2.538), IT2 and RO3 (M=2.462), IT1 and RO6 (M=2.385), LIB6 and LIB7 (M=2.308), and LIB8 (M=2.154). Figure 5.5 shows the comparisons of importance means and implementation means for RDG tasks related to the Administrative Offices, Research Data Consumer, and External Bodies roles.

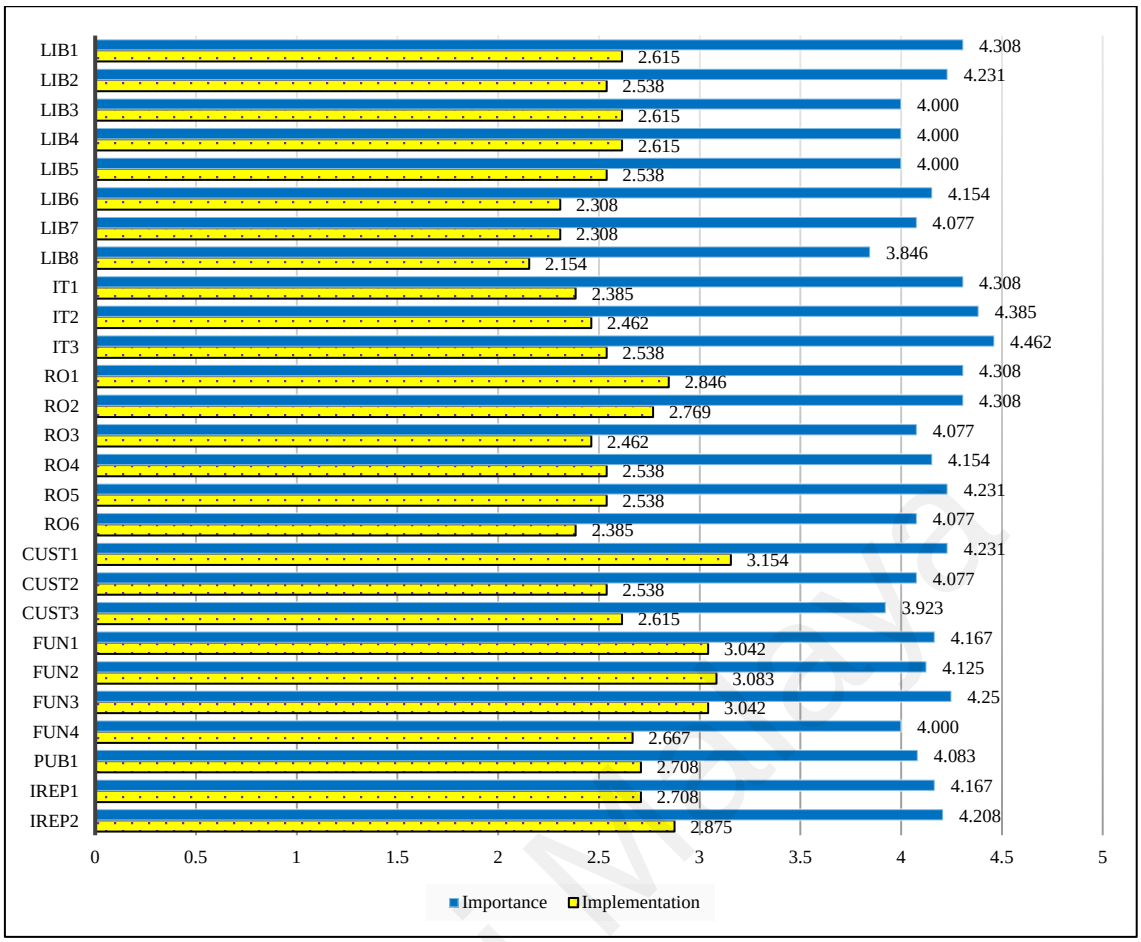


Figure 5.5: Comparisons between importance and implementation means for RDG tasks related to the Administrative Offices, Research Data Consumer, and External Bodies roles

The paired-sample t-test results between perceived importance and actual implementation for tasks assigned to the Administrative Offices, Research Data Consumer, and External Bodies reveal significant differences (Table 5.20). Negative mean differences were observed across all tasks, with the highest mean differences found in IT1 (M.D=-1.923, t=-4.319), IT2 (M.D=-1.923, t=-4.466), IT3 (M.D=-1.923, t=-4.811), LIB6 (M.D=-1.846, t=-5.482), LIB7 (M.D=-1.769, t=-4.679), LIB1 (M.D=-1.692, t=-4.879), LIB2 (M.D=-1.692, t=-4.430), LIB8 (M.D=-1.692, t=-3.941), RO5 (M.D=-1.692, t=-4.638), and RO6 (M.D=-1.692, t=-4.430). Despite being perceived as important, these tasks exhibit a significant implementation gap, highlighting a need for improvement.

Table 5.20: Gap Analysis between perceived importance and actual implementation of RDG tasks by Administrative Offices, Research Data Consumer, and External Bodies

TASK ID	TASK	Implementation			Importance			Paired-sample t-test		
		M	S.D.	Rank	M	S.D.	Rank	M.D	t-value	Sig.
LIB1	Coordinates a network of data stewards.	2.615	1.325	1	4.308	0.855	1	-1.692	-4.879	*0.000
LIB2	Manages research data metadata records and publishes them on a public catalog.	2.538	1.391	2	4.231	1.013	2	-1.692	-4.430	*0.001
LIB3	Maintains a research data repository for registering data and other outputs.	2.615	1.446	1	4.000	1.155	5	-1.385	-3.960	*0.002
LIB4	Defines the institution's support for research data management.	2.615	1.446	1	4.000	1.291	5	-1.385	-4.185	*0.001
LIB5	Provides high-quality infrastructure for data collection, storage, and sharing.	2.538	1.266	2	4.000	1.080	5	-1.462	-3.787	*0.003
LIB6	Reviews research data management policies with governance bodies.	2.308	1.251	3	4.154	1.144	3	-1.846	-5.482	*0.000
LIB7	Offers training on metadata standards, data management plans, and legal/ethical issues.	2.308	1.182	3	4.077	1.188	4	-1.769	-4.679	*0.001
LIB8	Advises on research data costs throughout the project lifecycle.	2.154	1.463	4	3.846	1.281	6	-1.692	-3.941	*0.002
IT1	Provides a standardized, robust infrastructure for effective research data management.	2.385	1.446	3	4.308	0.855	3	-1.923	-4.319	*0.001
IT2	Offers technical support for data management, including storage, backup, and archiving.	2.462	1.561	2	4.385	0.870	2	-1.923	-4.466	*0.001
IT3	Provides secure access management following ICT security guidelines.	2.538	1.506	1	4.462	0.877	1	-1.923	-4.811	*0.000
RO1	Advocates organizational awareness of research data management issues.	2.846	1.463	1	4.308	0.855	1	-1.462	-4.163	*0.001
RO2	Ensures governance policies are updated based on the latest funder requirements and national directives.	2.769	1.301	2	4.308	0.855	1	-1.538	-4.170	*0.001
RO3	Facilitates and supports the implementation of research data governance processes.	2.462	1.330	4	4.077	0.954	4	-1.615	-4.029	*0.002
RO4	Organizes training events on research data management.	2.538	1.266	3	4.154	0.899	3	-1.615	-4.395	*0.001
RO5	Provides advice, guidance, and assistance to researchers in preparing data management plans.	2.538	1.561	3	4.231	1.092	2	-1.692	-4.638	*0.001
RO6	Provides templates for incoming and outgoing research data agreements.	2.385	1.502	5	4.077	1.115	4	-1.692	-4.430	*0.001
CUST1	Acknowledges data sources and adhere to access terms and conditions.	3.154	1.345	1	4.231	1.092	1	-1.077	-2.694	*0.020
CUST2	Complies with controls specified in the data management plan.	2.538	1.450	3	4.077	1.038	2	-1.538	-3.826	*0.002
CUST3	Enters into data use agreements for accessing embargoed research data.	2.615	1.261	2	3.923	1.115	3	-1.308	-3.770	*0.003

Table 5.20, continued

TASK ID	TASK	Implementation			Importance			Paired-sample t-test		
		M	S.D.	Rank	M	S.D.	Rank	M.D	t-value	Sig.
FUN1	Defines research data governance principles aligned with funding principles.	3.042	1.488	2	4.167	1.007	2	-1.125	-4.776	*0.000
FUN2	Reviews implementation of research data management plans.	3.083	1.666	1	4.125	1.296	3	-1.042	-4.139	*0.000
FUN3	Specifies retention periods for significant research data.	3.042	1.459	2	4.250	0.897	1	-1.208	-4.608	*0.000
FUN4	Provides advice directly or through data services.	2.667	1.465	3	4.000	1.063	4	-1.333	-4.995	*0.000
PUB1	Endorses research data deposits in established repositories.	2.708	1.301	2	4.083	1.060	3	-1.375	-5.571	*0.000
IREP1	Defines non-proprietary standards for access, use, and interpretation.	2.708	1.233	2	4.167	1.050	2	-1.458	-5.420	*0.000
IREP2	Specifies required contextual information and metadata for deposited data.	2.875	1.262	1	4.208	1.103	1	-1.333	-5.426	*0.000

Notes.
1. Paired-sample t-test with probability *p < 0.05 and **p > 0.05.
2. Mean: mean scores were measured on a 5-point scale; the higher the score, the greater implementation or importance of RDG task.
3. S.D.: standard deviation.
4. Ranking was based on the highest to the lowest mean.
5. Mean difference (M.D): implementation mean – importance mean.

The last roles were presented to the panelists (n=11) engaged in Set D (Researcher). Alongside inquiries about the External Bodies, they were also questioned about the RDG tasks (n=24) pertaining to the Researcher role, focusing on both the implementation and perceived importance of each task. Since the discussion on External Bodies was conducted in the preceding section, the focus in this section is solely on the Researcher role.

The mean score on the importance of each task ranged from 4.273 to 4.727 (as shown in Figure 5.6), with an overall mean of 4.55. Notably, 42 percent (n=10) of the tasks exceeded the overall mean, with 6 tasks (R10, R14, R15, R17, R19, and R22) averaging 4.636, and 4 tasks (R23, R6, R8, and R9) averaging 4.727. Conversely, 3 tasks (R11, R18, and R5) have lower means of 4.364, and R3 has the lowest mean at 4.273, indicating a lower perceived level of importance.

On the other hand, the mean scores for actual implementation ranged from 3.000 to 4.091, with an overall mean of 3.5. Notably, 50 percent (n=12) of the tasks exceeded the overall mean. Specifically, 2 tasks (R18 and R21) have a mean of 3.545, another 2 tasks (R8 and R19) have a mean of 3.636, 2 more tasks (R14 and R23) have a mean of 3.727, 3 tasks (R1, R9, and R20) have a mean of 3.909, and R17 has a mean of 4.000. R6 achieved the highest mean among all tasks at 4.091. Conversely, 3 tasks (R3, R11, and R24) have lower mean scores of 3.000, indicating a comparatively lower level of actual implementation.

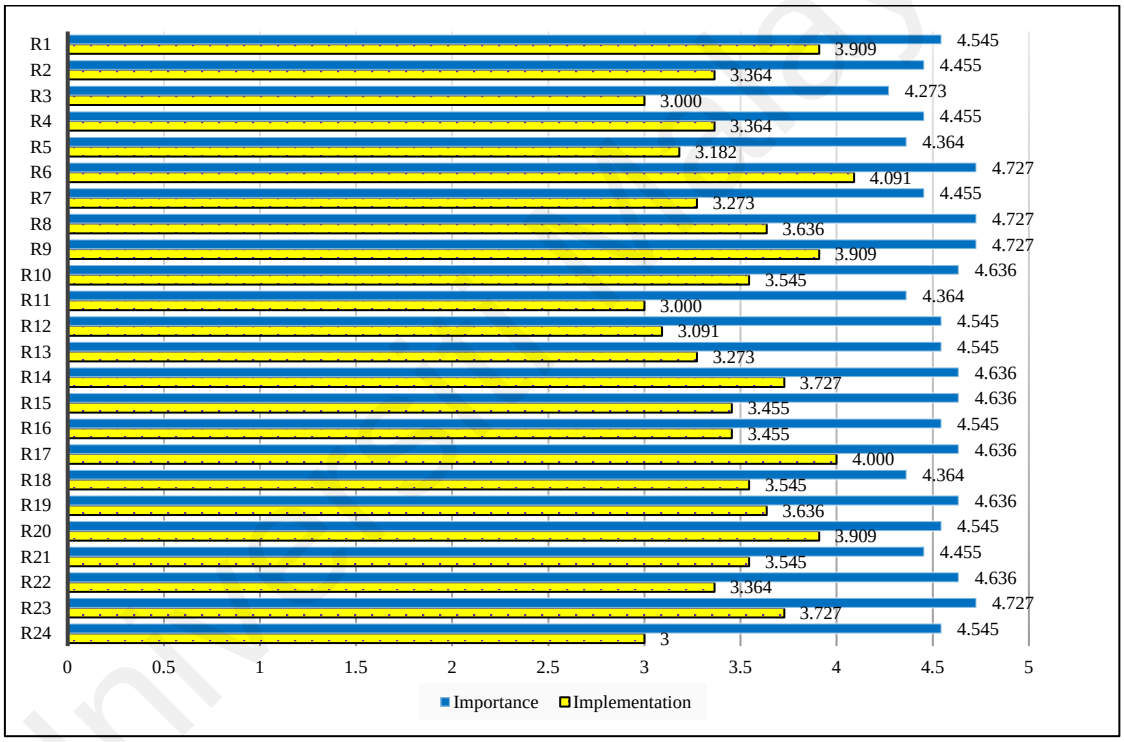


Figure 5.6: Comparisons between importance and implementation means for RDG tasks related to the Researcher role

The paired-sample t-test suggests a significant difference between the perceived importance of tasks associated with the Researcher role and their actual implementation, except for R20 with M.D = -0.636 and t = -2.055, yielding a higher p-value than the significance level of 0.05. This suggests that, for these particular tasks, there is no strong evidence to reject the null hypothesis that the perceived importance is equals to the actual implementation.

While certain tasks are considered highly important, they may not be implemented to the same extent, as all mean differences show negative values. The highest mean differences are for R24 (M.D = -1.545, t = -3.746), R12 (M.D = -1.455, t = -3.730), R11 (M.D = -1.364, t = -3.750), R3, R13, and R22 (M.D = -1.273, t = -3.545), highlighting potential areas for improvement in aligning perceived importance with actual implementation. Details regarding the gap analysis between perceived importance and actual implementation of RDG tasks by researchers are presented in Table 5.21.

Table 5.21: Gap analysis between perceived importance and actual implementation of RDG tasks by Researcher role

TASK ID	TASK	Implementation			Importance			Paired-sample t-test		
		M	S.D.	Rank	M	S.D.	Rank	M.D	t-value	Sig.
R1	Selects research data for long-term preservation based on verification/replication and reuse needs.	3.909	1.136	3	4.545	0.522	3	-0.636	-2.609	*0.026
R2	Retains research data securely for a period determined by the organization.	3.364	1.362	8	4.455	0.688	4	-1.091	-3.464	*0.006
R3	Budgets costs for capturing, managing, archiving, and sharing research data.	3.000	1.483	12	4.273	0.786	6	-1.273	-3.545	*0.005
R4	Ensures written agreements cover ownership, sharing, storage, and disposal of research data.	3.364	1.286	8	4.455	0.688	4	-1.091	-3.464	*0.006
R5	Writes a comprehensive research data management plan.	3.182	1.250	10	4.364	0.809	5	-1.182	-3.993	*0.003
R6	Protects intellectual property according to funder or contractual obligations.	4.091	0.944	1	4.727	0.467	1	-0.636	-2.609	*0.026
R7	Provides sufficient metadata for discoverability and reusability of research data.	3.273	1.555	9	4.455	0.820	4	-1.182	-3.357	*0.007
R8	Publishes data to established repositories, maximizing research value.	3.636	1.567	5	4.727	0.467	1	-1.091	-2.631	*0.025
R9	Guarantees integrity, quality, security, and persistent availability of research data.	3.909	1.221	3	4.727	0.467	1	-0.818	-2.764	*0.020
R10	Disposes of data and materials securely.	3.545	1.368	6	4.636	0.674	2	-1.091	-3.185	*0.010
R11	Includes a data access statement in publications.	3.000	1.483	12	4.364	0.674	5	-1.364	-3.750	*0.004
R12	Reports security breaches affecting research data.	3.091	1.514	11	4.545	0.688	3	-1.455	-3.730	*0.004
R13	Hands over data and materials after projects or leaving the institution.	3.273	1.489	9	4.545	0.522	3	-1.273	-3.545	*0.005

Table 5.21, continued

TASK ID	TASK	Implementation			Importance			Paired-sample t-test		
		M	S.D.	Rank	M	S.D.	Rank	M.D	t-value	Sig.
R14	Manages data throughout the lifecycle in line with policies, guidelines, and requirements.	3.727	1.555	4	4.636	0.505	2	-0.909	-2.469	*0.033
R15	Arranges safe and secure storage of research materials.	3.455	1.368	7	4.636	0.505	2	-1.182	-3.634	*0.005
R16	Understands and addresses risks of third-party storage solutions.	3.455	1.368	7	4.545	0.522	3	-1.091	-2.963	*0.014
R17	Protects confidential and sensitive data in line with legal and ethical requirements.	4.000	1.183	2	4.636	0.505	2	-0.636	-2.283	*0.046
R18	Participates in training on research data management and contractual obligations.	3.545	1.368	6	4.364	0.674	5	-0.818	-2.516	*0.031
R19	Upholds open-source file formats and types recommended for preservation.	3.636	1.502	5	4.636	0.505	2	-1.000	-2.622	*0.026
R20	Regularly backs up research data in accordance with best practices.	3.909	1.375	3	4.545	0.522	3	-0.636	-2.055	**0.067
R21	Develops operating procedures to comply with research data governance.	3.545	1.440	6	4.455	0.522	4	-0.909	-2.469	*0.033
R22	Works with the information security team to ensure system controls.	3.364	1.433	8	4.636	0.505	2	-1.273	-3.545	*0.005
R23	Identifies and addresses research data integrity and quality issues.	3.727	1.489	4	4.727	0.467	1	-1.000	-2.803	*0.019
R24	Periodically reviews data access and usage agreements.	3.000	1.732	12	4.545	0.522	3	-1.545	-3.746	*0.004

Notes.
1. Paired-sample t-test with probability *p < 0.05 and **p > 0.05.
2. Mean: mean scores were measured on a 5-point scale; the higher the score, the greater implementation or importance of RDG task.
3. S.D.: standard deviation.
4. Ranking was based on the highest to the lowest mean.
5. Mean difference (M.D): implementation mean – importance mean.

It is worth noting that only five RDG tasks received p-values higher than the significance threshold of p=0.05. In these cases, the null hypothesis failed to be rejected, suggesting that there is insufficient evidence to conclude a difference between the perceived importance and actual implementation for these specific RDG tasks. In general, the analysis suggests that there is a significant discrepancy between the perceived importance and actual implementation of RDG tasks within RPOs, with the tasks generally being perceived as more important than they are currently implemented.

Further investigation into the reasons behind this misalignment and potential corrective actions may be conducted in the future. The observed gap may indicate challenges in resource allocation or implementation strategies, potentially due to constraints in budget, expertise, or technology, as emphasized by one panelist, "*P21-actually, to develop very high-quality infrastructure and facilities, it is up to the institution/university/organization budget*". Furthermore, the negative mean difference may reflect a lack of awareness or a cultural barrier within RPOs regarding the importance of specific RDG tasks. This notion is supported by a comment from one of the panelists, stating that there is a "*P43-lack of awareness of standards required*". Additionally, the identified gap could signal a need for training and skill development among personnel responsible for RDG task implementation. This is substantiated by the statement of one panelist who mentioned, "*P25-less training available related to research data*".

5.7 Summary of Chapter Five

The chapter focuses on findings for RQ1 and RQ2, aiming to investigate the functional dimensions of RDG activities, including tasks, areas, and decision domains. This study considers both the actual implementation and perceived importance of RDG tasks among data practitioners within RPOs in Malaysia, conducted through the modified Delphi study. The chapter provides an overview of the demographic profiles of participants involved in the study. It then examines the implementation of RDG among RPOs in Malaysia. Following this, a subsequent section explores the importance of RDG activities, presenting insights gathered from the panelists through the first three Delphi rounds. Additionally, the chapter addresses the progression towards formulating research-derived descriptions for RDG areas and decision domains. Moreover, a gap analysis is conducted within the chapter, examining the disparities between the perceived importance and the actual implementation of RDG activities within Malaysia RPOs. Figure 5.7 illustrates an overview of Delphi process in addressing RO1, i.e. to investigate RDG activities in

relation to both actual implementation and perceived importance among data practitioners. The following Chapter Six explores the stakeholders that the panelists consider responsible for RDG roles and identifies the structural positions of these roles within RPOs. Additionally, it presents findings on the descriptions of individual RDG roles and determines the appropriate nomenclature for each role.

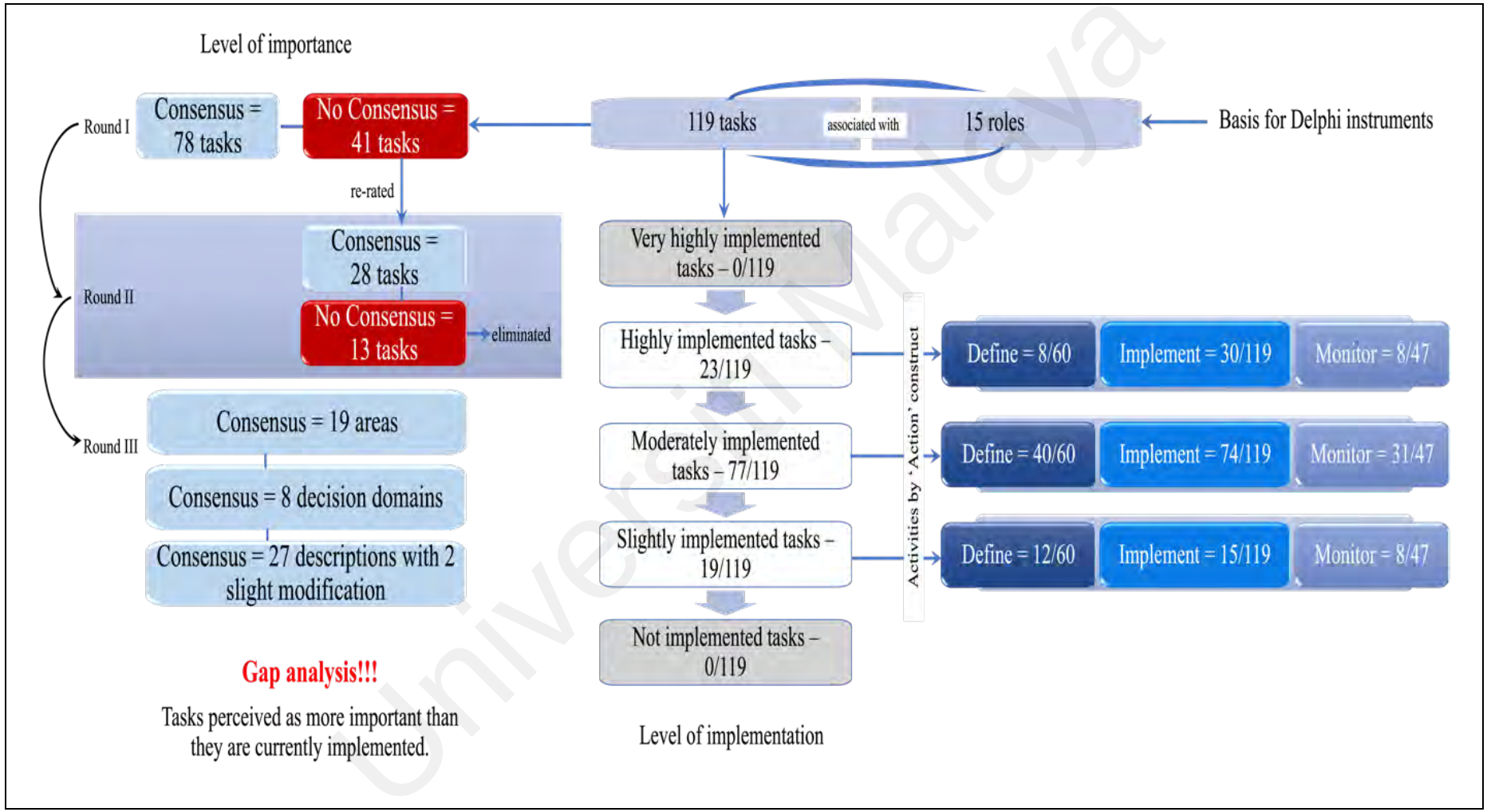


Figure 5.7: An overview of Delphi process in addressing RO1

CHAPTER 6: RESEARCH DATA GOVERNANCE ROLES AND STRUCTURAL POSITIONS IN MALAYSIA RESEARCH PERFORMING ORGANIZATIONS

"It is He who has made you successors (khalifah) upon the Earth…"
(Al-Baqarah, 2:30)

6.1 Introduction

Chapter Six presents the key findings corresponding to Research Objective 2, i.e. defining stakeholders and their structural positions for RDG roles within RPOs. This chapter respectively addresses the following research questions:

RQ3: Which stakeholders do data practitioners in research performing organizations consider responsible for research data governance roles?

RQ4: How do data practitioners identify the structural positions of research data governance roles within research performing organizations?

This chapter explores which stakeholders, data practitioners in RPOs consider responsible for RDG roles, and identify the structural positions of RDG roles within RPOs. Besides that, the findings on descriptions for individual RDG roles, along with determining the appropriate nomenclature for each role, are also presented.

6.2 Identifying Research Data Governance Stakeholders

To answer RQ3: *"Which stakeholders do data practitioners in research performing organizations consider responsible for research data governance roles?"*, it is necessary to utilize findings from Rounds I and II. In all Delphi rounds, besides questions on RDG activities, there were various questions and statements related to the allocation of RDG roles among different stakeholders. This is important for gaining insights into the appropriate stakeholders, whether individuals or bodies, that should hold specific roles to govern research data. The identification of appropriate stakeholders was fulfilled in

Rounds I and II. The discussion highlights are organized according to each role, deviating from the previous approach of organizing them based on the priority of RDG activities. This arrangement aims to offer a clearer and more easily understandable presentation. Given that the descriptive statistics only involved frequencies and percentages, Excel was utilized as the primary tool for statistical analysis.

Although the Delphi panel was divided into four groups and responded to specific questions within each set of the instrument during Rounds I and II, all panelists were assigned the responsibility of addressing eight questions, corresponding to eight identified roles. These questions comprised seven multiple-choice questions and one open-ended question, with a focus on identifying key players for RDG roles. Additionally, the panelists had the option to supplement their responses by adding more answers based on their organization's implementation or their own knowledge in the "Other" category. In Round I, the primary objective was to gather insights on RDG key players associated with specific roles by extracting as much information as possible from the panels. Therefore, achieving a consensus was not a focal aim during the initial round for this particular set of questions.

The eight questions from the initial round were subsequently carried over to Round II, where all existing and new stakeholders – that have been added by the panelists for each question – were listed in the answer options for the panelists to re-rate the best possible answer to be adapted in the RDG framework. During Round II, each panelist had the opportunity to either retain or revise their responses. This process was guided by group responses, aiming to achieve a consensus on the appropriate stakeholders holding the specified RDG roles within the context of RPOs in Malaysia. Twenty-one out of 34 panelists (62%) have amended at least one of their responses from the previous round.

One of the factors was influenced by – as stated by one of the panelists – *"P41-based on new options provided and comments from the group"*.

6.2.1 Stakeholders in the Role of Executive Sponsor

In response to Question 1, the panelists were required to specify the stakeholder who (should) hold the role of an Executive Sponsor. They had the option to choose more than one answer that applied. During Round I, the results reveal a variety of preferences for the Executive Sponsor role, with the majority (72.3%, n=34) indicating that the role (should be) held by the Deputy Vice Chancellor (Research)/Deputy Rector (Research). Additionally, 21.3 percent (n=10) opted for the Director of Institute/Director General/Chief Executive Officer, 10.6 percent (n=5) chose Deputy Vice Chancellors/Deputy Rectors, 8.5 percent (n=4) specified Vice Chancellor/Rector, and 6.4 percent (n=3) selected Deputy Director of Institute/Deputy Director General/Deputy Chief Executive Officer. Moreover, 10.6 percent (n=5) identified the Chief Information Officer as (should) holding the position, while one participant each (2.1%) pointed out the Chief Statistician, Head of Centre, Honorary Professor, and the Malaysian Ministry of Higher Education. Comments or clarifications from the panelists included noting that the task might sometimes be shared with other senior management within the institute, for example the Centre Head. They emphasized the importance of someone with an understanding of data management and research for the role.

In Round II, the analysis reveals a predominant presence of Deputy Vice Chancellors or Deputy Rectors (Research), constituting 79.4 percent (n=27) of the panelists. This suggests a substantial emphasis on research leadership within academic institutions. Additionally, individuals holding the position of Director of Institute, Director General, Chief Executive Officer, or Chief Statistician account for 2.6 percent (n=7) of the executive sponsors, signifying a specialized focus on specific institutes or organizations

involved in research. This suggests that both groups are comprised of senior management individuals who could serve as Executive Sponsors, depending on the context of an academic or research institution and its suitability for the respective organization. Larger organizations may require more executive sponsors to oversee and ensure the implementation of RDG.

In addition to the individuals mentioned above, the involvement of Deputy Vice Chancellors or Deputy Rectors, constituting 11.8 percent (n=4) of the panelists, indicates a broader distribution of responsibilities, potentially covering various academic and administrative duties beyond research. Similarly, Vice Chancellors or Rectors, also at 11.8 percent (n=4), signify a high-level commitment to research within university leadership, reflecting a strategic integration of research into the overall mission of the institution. Chief Information Officers, at 8.8 percent (n=3), play a role in the executive sponsorship, highlighting the recognition of the importance of information management in research activities. Their involvement suggests a focus on data governance, technology infrastructure, and security in research processes. Individuals in Deputy Director positions (5.9%, n=2) may support primary executive sponsors and manage specific aspects of research initiatives.

The inclusion of an Honorary Professor (2.9%, n=1) in an Executive Sponsor role indicates the involvement of individuals with significant academic or research credentials, even if they hold honorary positions. Notably, a representative from the Malaysian Ministry of Higher Education (2.9%, n=1) participates in executive sponsorship, suggesting external oversight and influence on RDG to ensure compliance with national policies and standards. However, Heads of Centre are absent from the executive sponsor role, implying that primary responsibility for RDG may lie outside the scope of individual research centres.

In summary, the findings indicate a strong preference for the Deputy Vice Chancellor (Research)/Deputy Rector (Research) to assume the role of an Executive Sponsor. However, in cases where RPOs do not have a Deputy Vice Chancellor (Research)/Deputy Rector (Research), it is recommended that the position be held by individuals such as the Director of Institute/Director General/Chief Executive Officer or any C-level management personnel within the organization, such as the Chief Statistician. The detailed breakdown of responses for this question is presented in Table 6.1.

Table 6.1: Panelist responses on stakeholders in the Executive Sponsor role in Round I and II

Stakeholder in Executive Sponsor Role	Round I			Round II		
	Rank	Count	Percentage	Rank	Count	Percentage
Deputy Vice Chancellor (Research)/Deputy Rector (Research)	1	34	72.3%	1	27	79.4%
Deputy Vice Chancellors/Deputy Rectors	3	5	10.6%	3	4	11.8%
Deputy Director of Institute/Deputy Director General/Deputy Chief Executive Officer	5	3	6.4%	5	2	5.9%
Director of Institute/Director General/Chief Executive Officer/**Chief Statistician	2	10	21.3%	2	7	20.6%
Vice Chancellor/Rector	4	4	8.5%	3	4	11.8%
*Chief Information Officer	3	5	10.6%	4	3	8.8%
**Chief Statistician	6	1	2.1%	-	-	-
*Head of Centre	6	1	2.1%	7	0	0.0%
*Honorary Professor	6	1	2.1%	6	1	2.9%
*Malaysian Ministry of Higher Education	6	1	2.1%	6	1	2.9%
Comment, feedback or rewording for clarification:	-	2	4.3%	-	2	5.9%

From Round I

1. P05-Sometime the task will be shared with other senior management within institute especially center head.
2. P45-Better someone who has an understanding about data management and for research.

From Round II

1. P06-This department is research management center.
2. P41-Based on new option provided and comments of the group.

Notes.
*Stakeholder identified by panelist in Round I was included as new option for group reassessment in Round II.
**Stakeholder identified by panelist in Round I was incorporated into an existing option during Round II.

6.2.2 Stakeholders in the Role of Data Governance Leader

For Question 2, the panelists were asked about who holds the role of a Data Governance Leader in their institution, or if the position has not been implemented, then who is the most appropriate person to hold the position. The panelists were required to select only one best answer for this item. The findings reveal that various stakeholders currently hold or are deemed responsible for the role of a Data Governance Leader. A comment from one of the panelists highlighted that there are still institutions where this position has not been established, as stated by "*P03-the position is not available yet, but I think maybe Deputy VC (Research) is the right person for that position*". Specifically, 36.2 percent (n=17) of the panelists specified the Deputy Vice Chancellor (Research)/Deputy Rector (Research) as fulfilling this role, followed by the Chief Information Officer at 31.9 percent (n=15). Additionally, 14.9 percent (n=7) mentioned the Chief Data Officer, 6.4 percent (n=3) identified the Chief Librarian, and 2.1 percent (n=1) specified a Primary Investigator. Furthermore, some panelists added other stakeholders for the Data Governance Leader, including 4.3 percent (n= 2) for the Director of Research Planning Division, and 2.1 percent (n=1) each for the Chief Statistician and Chief Digital Officer. Some of them have provided additional clarifications on what is being implemented at their institution.

In Round II, the analysis of Data Governance Leaders indicates that Chief Information Officers play a predominant role, comprising 38.2 percent (n=13) of the leadership responsible for overseeing RDG. Following closely are Deputy Vice Chancellors (Research)/Deputy Rectors (Research) at 35.3 percent (n=12), emphasizing the significant influence of academic leadership in shaping policies and strategies for data governance. In the initial round, the situation was reversed, with Deputy Vice Chancellors (Research)/Deputy Rectors (Research) being predominant, followed closely by the Chief Information Officer. This implies that for research institutions without the positions of

Deputy Vice Chancellors (Research) or Deputy Rectors (Research), they have the option to designate the Chief Information Officer to assume the role. Conversely, in cases where the Deputy Vice Chancellors (Research) or Deputy Rectors (Research) positions are established, the organization has the discretion to assign either of these roles, including the Chief Information Officer, to fulfil the responsibilities.

Meanwhile, Chief Data Officers were identified by 14.7 percent (n=5) of the panelists, contribute a specialized focus on data management, ensuring data quality, and fostering a data-driven culture within the research environment. Leaders with titles such as Director of Institute, Director General, Chief Executive Officer, or Chief Statistician, specified by 8.9 percent (n=3), indicate a strategic approach to data management within specific entities, potentially emphasizing statistical accuracy and organizational data needs. Chief Digital Officers, at 2.9 percent (n=1), highlight the recognition of digital strategies in data governance, overseeing the digital transformation of data management processes within the research context. However, positions such as Chief Librarian, Head of Centre, and Principal Investigator/Research Lead do not appear in the role of Data Governance Leader in this context, suggesting that these roles may not be directly responsible for overall data governance. In comment sections, some of the panelists have provided additional clarifications on why they made their preference choices.

Therefore, the results indicate a lack of clear consensus among stakeholders regarding who should assume the role of a Data Governance Leader. As a result, it can be inferred that either the Chief Information Officer or the Deputy Vice Chancellor (Research)/Deputy Rector (Research) could take the lead in spearheading the RDG initiatives. Table 6.2 sheds light on the diverse representation of stakeholders connected to the role of Data Governance Leader.

Table 6.2: Panelist responses on stakeholders in the Data Governance Leader role in Round I and II

Stakeholder in Data Governance Leader Role	Round I			Round II		
	Rank	Count	Percentage	Rank	Count	Percentage
Deputy Vice Chancellor (Research)/Deputy Rector (Research)	1	17	36.2%	2	12	35.3%
Chief Information Officer	2	15	31.9%	1	13	38.2%
Chief Data Officer	3	7	14.9%	3	5	14.7%
Chief Librarian	4	3	6.4%	6	0	0.0%
Principal Investigator/Research Lead	6	1	2.1%	6	0	0.0%
**Director, Institutional Planning & Strategic Center/Director of Research Planning Division	5	2	4.3%	-	-	-
**Chief Statistician	6	1	2.1%	-	-	-
*Chief Digital Officer	6	1	2.1%	5	1	2.9%
***Director of Institute/Director General/Chief Executive Officer/Chief Statistician	-	-	-	4	3	8.9%
***Head of Centre	-	-	-	6	0	0.0%
TOTAL	-	**47**	**[a]100%**	-	**34**	**[a]100%**
Comment, feedback or rewording for clarification:	-	3	6.4%	-	6	17.6%

From Round I

1. P03-The position is not available yet, but I think maybe Deputy VC (Research) is the right person for that position.
2. P18-Is it chief information officer is same with head of the branch library? if yes, at my place, partially data of campus being handled by library especially for archive. But for current data, no.
3. P-30-Director OSIC (Office of Strategy and Institutional Change) is currently in charge of IIUM organization's data under the supervision of Chief Digital Officer.

From Round II

1. P05-Looking at the broader aspect of how ministry of health do things, the original position of CDO (held by KSU) will ensure all the data governance principals and the decision will cover more grounds/ more nuance than merely primary investigator.
2. P06-As research management center they also manage all research data.
3. P14-CIO has wider strategic role than CDO.
4. P27-Data storage policy under the supervision of data centre under CIO.
5. P34-My opinion is that the CIO should be the leader. My original response is based just USM open science.
6. P44-Currently, in my institution, the management of research data is handled by the library, with the Chief Librarian being responsible for data management. However, upon further consideration, I have decided to choose the Chief Information Officer as the responsible party for Research Data Governance.

Notes.
*Stakeholder identified by panelist in Round I was included as new option for group reassessment in Round II
**Stakeholder identified by panelist in Round I was incorporated into an existing option during Round II
***More options added in Round II based on panelists' responses in Round I
[a]Percentages were rounded to the nearest whole number

6.2.3 Stakeholders in the Role of the Research Data Governance Committee

Given that the Research Data Governance Committee comprises individuals in various roles, reflecting a diverse distribution of responsibilities, Question 3 allowed panelists to select more than one appropriate answer. Based on the responses received, 76.6 percent (n=36) of the panelists identified the Head of Research as one of the entities holding this role, followed by 55.3 percent (n=26) for Deputy Vice Chancellor/Deputy Rector and 53.2 percent (n=25) for the Research Data Steward. Meanwhile, 31.9 percent (n=15) of the panelists asserted that the Dean/Head of School and Head of Department are also included, while 29.8 percent (n=14) of the panelists mentioned a Professional Staff Member. Other roles contributing to the committee include Research Data Owner (27.7%, n=13), Senior Academic (19.1%, n=9), Chief Librarian (12.8%, n=6), Research Data Consumer (8.5%, n=4), and Research Data Creator (2.1%, n=1). Additionally, the panelists indicated involvement from various roles as implemented in their institution. For instance, the Director of Research Centre (4.3%, n=2), Deputy Dean (Research), Research Ethics Committee, Chief Statistician, Head of Information Technology Department, Head of Institute of Postgraduate Studies, and Centre for Knowledge, Communication & Technology (PPKT) each had 2.1 percent (n=1).

The composition of the Research Data Governance Committee reveals a significant representation of key individuals. In Round II of the analysis, it is evident that the Head of Research (Faculty) remains predominant, with 73.5 percent (n=25) of panelists specifying this role. This underscores a strong focus on faculty-level leadership, demonstrating a commitment to aligning data practices with the specific needs of the faculty. Deputy Vice Chancellors or Deputy Rectors, identified by 61.8 percent (n=21), play a vital role, signaling a high-level institutional commitment to data governance. They actively contribute to shaping policies and strategies for RDM. Research Data Steward, chosen by 58.8 percent (n=20) of panelists, is instrumental in overseeing the practical

aspects of data management. Their inclusion underscores the importance of dedicated individuals responsible for ensuring the quality, integrity, and compliance of research data.

Besides that, the individuals like Heads of Departments, specified by 38.2 percent (n=13), provide a departmental perspective to the committee, emphasizing a decentralized approach to address unique data governance needs within various academic departments. Deans or Heads of Faculties, at 32.4 percent (n=11), contribute with broader academic and administrative perspectives, suggesting a holistic approach to data governance aligned with the overall mission of the faculty. Professional Staff Members and Senior Academics, each at 26.5 percent (n=9), ensure a balance between administrative and academic expertise within the committee, supporting a comprehensive understanding of RDG. The Chief Librarian, at 23.5 percent (n=8), brings library expertise to the committee, highlighting the library's role in managing and preserving research data, especially concerning data access and archiving.

Research Data Owner, specified by 20.6 percent (n=7) actively involved, represent individuals responsible for specific datasets, ensuring the active participation of those intimately familiar with the data in its governance. Individuals in leadership positions, such as the Director of Research Management Office (14.7%, n=5), Director of Institute/Director General/Chief Statistician, Head of Centre, and Research Data Creator (11.8%, n=4), along with Deputy Dean (Research), Research Data Consumer, and Research Ethics Committee, each at 8.8 percent (n=3), collectively offer a diverse array of perspectives within the committee. This inclusive approach ensures a well-rounded and comprehensive approach to RDG. Finally, the Head of the Information Technology Department and the Head of the Institute of Postgraduate Studies each have received no

response. Several justifications were added by the panelists, expressing their thoughts on why they modified their responses from the previous round.

The outcomes of the findings suggest a prevailing agreement in favour of forming the Research Data Governance Committee with a combination of Heads of Research (Faculty), Deputy Vice Chancellors/Deputy Rectors, and Research Data Stewards. In cases where these stakeholders are not present in any RPOs, inclusion of Heads of Departments and/or Deans/Heads of Faculty as committee members is recommended. While this inclusion may not achieve a majority consensus, it is noteworthy that they are among the top five ranks in the list. Table 6.3 offers insight into the diverse distribution of stakeholders associated with the position of Research Data Governance Committee.

6.2.4 Stakeholders in the Role of the Office of Research Data Governance

According to responses from Question 4, the central management of RDG responsibilities is distributed among various offices and departments within RPOs. 29.8 percent (n=14) of the panelists specified the Office of Research Integrity and Ethics as the designated RDG office. In addition, the Office of Deputy Vice Chancellor (Research)/Office of Deputy Rector (Research) accounted for 23.4 percent (n=11), while the Research Management Office was chosen by 21.3 percent (n=10) of the panelists. Notably, 14.9 percent (n=7) of the panelists specified that the Library holds this role. Other offices identified to handle the responsibility of the RDG office include the Human Resource Department and the Office of Manager, each with 4.3 percent (n=2). Furthermore, one panelist indicated the involvement of the Office of the Chief Statistician in his institution, representing 2.1 percent (n=1).

Table 6.3: Panelist responses on stakeholders in the Research Data Governance Committee role in Round I and II

Stakeholder in Research Data Governance Committee Role	Round I			Round II		
	Rank	Count	Percentage	Rank	Count	Percentage
Head of Research (Faculty)	1	36	76.6%	1	25	73.5%
Deputy Vice Chancellor/Deputy Rector	2	26	55.3%	2	21	61.8%
Research Data Steward	3	25	53.2%	3	20	58.8%
Dean/Head of Faculty	4	15	31.9%	5	11	32.4%
Head of Department	4	15	31.9%	4	13	38.2%
Professional Staff Member	5	14	29.8%	6	9	26.5%
Research Data Owner	6	13	27.7%	8	7	20.6%
Senior Academic	7	9	19.1%	6	9	26.5%
Chief librarian	8	6	12.8%	7	8	23.5%
Research Data Consumer	9	4	8.5%	6	3	8.8%
Research Data Creator	11	1	2.1%	10	4	11.8%
**Directors of Research Centres	10	2	4.3%	-	-	-
*Deputy Dean (Research)	11	1	2.1%	11	3	8.8%
*Research Ethics Committee	11	1	2.1%	11	3	8.8%
**Chief Statistician	11	1	2.1%	-	-	-
**Head of Information Technology Department	11	1	2.1%	-	-	-
*Head of Institute of Postgraduate Studies	11	1	2.1%	12	0	0.0%
**Centre for Knowledge, Communication & Technology (PPKT)	11	1	2.1%	-	-	-
***Director of Research Management Office	-	-	-	9	5	14.7%
***Director of Institute/Director General/ Chief Executive Officer/Chief Statistician	-	-	-	10	4	11.8%
***Head of Centre	-	-	-	10	4	11.8%
Comment, feedback or rewording for clarification:	-	0	0.0%	-	4	11.8%

From Round II

1. P30-I added the Chief Librarian because at our university, the Deputy Rector Responsible Research and Innovation requested the Library to lead RDM. Hence, I think the Chief Librarian must be one of the Research Data Governance Committee.
2. P44-Upon further consideration, the following key positions are responsible for implementing policies and practices for research data governance.
3. P18-Because I think back that there's another post that need to play role in RDG Committee especially for the beginner campus to start RDG.
4. P19-Because I think there's another post that need to play role in RDG Committee.

Notes.
*Stakeholder identified by panelist in Round I was included as new option for group reassessment in Round II
**Stakeholder identified by panelist in Round I was incorporated into an existing option during Round II
***More options added in Round II based on panelists' responses in Round I

In Round II, the analysis of the distribution of offices or departments responsible for the role of the Office of Research Data Governance still predominantly shows the Office of Research Integrity and Ethics, similar to Round I, with 38.2 percent (n=13) of the panelists agreed that the office holds the specified role. This highlights a clear alignment of RDG with ethical considerations, emphasizing the importance of integrity in data management. Eight (23.5%) of the panelists agreed that the offices of Deputy Vice Chancellors or Deputy Rectors (Research) play a significant role, reflecting a top-level institutional approach to integrating RDG within broader research leadership.

Research Management Offices were selected by 20.6 percent (n=7) of the panelists for the role, underscoring the recognition of the need for dedicated offices focused on managing and overseeing various aspects of research, including RDG. Meanwhile, the library assumed the role of the Office of Research Data Governance, as suggested by 8.8 percent (n=3) of the panelists. This highlights the evolving role of libraries in managing research data, particularly in terms of data access, preservation, and archiving.

In addition, 2.9 percent (n=1) of the panelists viewed that the Human Resource Department is identified as playing the role, suggesting a recognition that human resources are integral to the effective governance of research data, potentially focusing on training, compliance, and workforce development. Other than that, offices associated with top leadership, including Directors of Institutes or Chief Statisticians, were identified by 2.9 percent (n=1) of the panelists. This indicates a strategic approach, aligning data governance with high-level organizational leadership. Another 2.9 percent (n=1) of them chosen that a general “Office of Manager” may play the role, suggesting a more decentralized or department-specific approach to the role of the Office of Research Data Governance.

However, based on the data, there is no prevailing consensus among panelists on which office should hold the role as the Office of Research Data Governance. Consequently, the responsibility for the Office of Research Data Governance can be assigned to either the Office of Research Integrity and Ethics, the Office of Deputy Vice Chancellor (Research)/Office of Deputy Rector (Research), the Research Management Office, or the Library. Based on the comment made by P36, the Office of Research Integrity and Ethics is ideally suited to hold this role. However, the decision can be made based on the organizations' perception of suitability and, if necessary, considering the availability of the respected office(s). Insights into the varied distribution of stakeholders linked to the role of Office of Research Data Governance are provided in Table 6.4.

Table 6.4: Panelist responses on stakeholders in the Office of Research Data Governance role in Round I and II

Stakeholder in Office of Research Data Governance Role	Round I			Round II		
	Rank	Count	Percentage	Rank	Count	Percentage
Office of Research Integrity and Ethics	1	14	29.8%	1	13	38.2%
Office of Deputy Vice Chancellor (Research)/ Office of Deputy Rector (Research)	2	11	23.4%	2	8	23.5%
Research Management Office	3	10	21.3%	3	7	20.6%
Library	4	7	14.9%	4	3	8.8%
Human Resource Department	5	2	4.3%	5	1	2.9%
Office of Manager	5	2	4.3%	5	1	2.9%
**Office of Chief Statistician	6	1	2.1%	-	-	-
***Office of Director of Institute/Director General/Chief Executive Officer/Chief Statistician	-	-	-	5	1	2.9%
TOTAL	-	**47**	**[a]100%**	-	**34**	**[a]100%**
Comment, feedback or rewording for clarification:	-	0	0%	-	1	3%

From Round II

1. P36-Office of Research Integrity and Ethics would be more specific to address this matter.

Notes.
**Stakeholder identified by panelist in Round I was incorporated into an existing option during Round II
***More options added in Round II based on panelists' responses in Round I
[a]Percentages were rounded to the nearest whole number

6.2.5 Stakeholders in the Role of Research Data Owner

Given that many parties and entities are involved in dealing with research data, various individuals share ownership and have the responsibility to ensure that all legal, regulatory, and policy requirements are met for the research data they handle. Therefore, Research Data Owners are dispersed among different individuals and entities, and Question 5 aimed to unravel the answer. Realizing the co-ownership possibility, the panelists were allowed to provide more than one possible answer.

In Round I, the survey results indicate that, besides the Organization, several stakeholders involved in research data are regarded as Research Data Owners. The majority, 85.1 percent (n=40), identified the Principal Investigator/Research Lead as holding the Research Data Owner role. The Head of the Research Institute represented 40.4 percent (n=19), while the Dean/Head of School was specified by 10.6 percent (n=5). Moreover, the panelists offered additional perspectives on those fulfilling the role of Research Data Owner such as the Library research support team, Research Supervisor, and the Chief Statistician were each mentioned by 2.1 percent (n=1), and the Director of the Research Management Office was suggested by 4.3 percent (n=2).

While the findings for Round II indicate that the first two groups of individuals, such as Principal Investigators/Research Leads and Heads of Research Institutes, remain predominant entities, with 82.4 percent (n=28) and 26.5 percent (n=9) of the panelists agreeing with it, respectively. The former emphasizes a decentralized and hands-on approach to ownership, and the latter reflects institutional ownership where leaders at the institute level actively oversee and take responsibility for research data. This is supported by the comment made by P44, who stated that Principal Investigators hold more responsibilities towards the produced research data. Directors of Research Management Offices take on the role as specified by 5.9 percent (n=2) of the panelists, indicating a

specialized administrative involvement in data ownership aligned with the office's broader responsibility for managing various aspects of research.

Top leadership, including Directors of Institutes or Chief Statisticians, may play the role, as suggested by 2.9 percent (n=1) of the panelists, demonstrating a strategic and high-level involvement in data ownership in alignment with organizational leadership responsibilities. Research Supervisors were identified as playing the role by 2.9 percent (n=1) of the panelists, suggesting active involvement in owning and overseeing the data generated within the scope of research projects. However, Dean/Head of Faculty, previously selected by 5 panelists, and Library Research Support Team, with 1 panelist selected in the initial round, changed their response to another option. Thus, they did not appear in the role of Research Data Owner in Round II, indicating that ownership may not be assigned to these positions in this specific context. This contradicts the viewpoint expressed by P02, who suggested that the Dean/Head of Faculty should be the owner of the research data, given their authority in accessing information on legal matters.

While various stakeholders received responses regarding who should hold the role of the Research Data Owner, the consensus suggests that Principal Investigators/Research Leads are considered the owners of the research data. Table 6.5 displays a detailed breakdown of responses pertaining to this question.

Table 6.5: Panelist responses on stakeholders in the Research Data Owner role in Round I and II

Stakeholder in Research Data Owner Role	Round I			Round II		
	Rank	Count	Percentage	Rank	Count	Percentage
Principal Investigator/Research Lead	1	40	85.1%	1	28	82.4%
Head of Research Institute	2	19	40.4%	2	9	26.5%
Dean/Head of Faculty	3	5	10.6%	5	0	0.0%
*Director of Research Management Office	4	2	4.3%	3	2	5.9%
*Library research support team	5	1	2.1%	5	0	0.0%
*Research Supervisor	5	1	2.1%	4	1	2.9%
**Chief Statistician	5	1	2.1%	-	-	-
***Director of Institute/Director General/ Chief Executive Officer/Chief Statistician	-	-	-	4	1	2.9%
Comment, feedback or rewording for clarification:	-	0	0.0%	-	3	8.8%

From Round II

1. P02-Head of research institute and dean/head of faculty are unlikely to have access (especially time) to check all legal, regulatory, and policy requirements.
2. P38-The keyword research data owner.
3. P44-After further consideration, the Principal Investigator (PI) holds more responsibilities.

Notes.
*Stakeholder identified by panelist in Round I was included as new option for group reassessment in Round II
**Stakeholder identified by panelist in Round I was incorporated into an existing option during Round II
***More options added in Round II based on panelists' responses in Round I

6.2.6 Stakeholders in the Role of Research Data Governor

The role of Research Data Governor(s) within RPOs is carried out by a variety of individuals, as evidenced by the survey responses to Question 6. The Head of the Research Institute was identified as fulfilling this role by 66 percent of responses, totalling 31 panelists. The Head of Research within faculties was mentioned by 51.1 percent, with 24 panelists, while the Principal Investigator/Research Lead represented 38.3 percent, with 18 panelists. Additionally, the role was attributed to the Dean/Head of School by 21.3 percent, with 10 panelists, and to Senior Administrators by 8.5 percent, with 4 panelists. Furthermore, some panelists suggested additional contributors to the role of Research Data Governor(s), such as Deputy Dean of Research (2.1%, n=1), Chief Statistician (2.1%, n=1), and Librarian (2.1%, n=1).

The results for Round II indicate that the first four stakeholders remain in the same rank of selection, where the Head of Research Institute emerges as a prominent figure, identified by 82.4 percent (n=28) of the panelists. This suggests that at the school or department level, leaders of research institutes play a pivotal role in governing research data, overseeing its management and integrity. This notion is supported by P14 and P43, who emphasized the importance of the Head of Research Institute in ensuring that policies are adhered to. Following closely, Heads of Research within faculties were identified by 58.8 percent (n=20) of the panelists. This emphasizes a decentralized approach, where faculty-level leadership actively engages in governing research data to align with specific academic needs. Besides the Head of Research Institute, the Heads of Research are also identified as crucial entities who should hold the role, ensuring the administration and implementation of research data governance policies, as expressed by P44. Principal Investigators or Research Leads were recognized as potential Research Data Governors by 29.4 percent (n=10) of the panelists. This highlights the importance of those directly involved in generating research data taking an active role in its governance at the faculty level. Deans or Heads of Faculty, identified by 14.7 percent (n=5) of them, contribute to the governance of research data. Their involvement signifies a broader administrative perspective, aligning data governance with overall faculty missions.

Meanwhile, top-level leadership, including Directors of Institutes or Chief Statisticians, were mentioned by 5.9 percent (n=2) of the panelists. This indicates a strategic and high-level involvement in governing research data, aligning with broader organizational goals. Senior Administrators were mentioned by 5.9 percent (n=2) of the panelists, while Deputy Deans (Research) were mentioned by 2.9 percent (n=1), suggesting a shared responsibility for data governance at the faculty level. However, librarians did not appear in the identified role for Research Data Governor in this context.

This may suggest a specific distribution of responsibilities, focusing more on academic and administrative leadership.

Overall, the results highlight a collaborative and distributed approach to governing research data at the faculty (departmental) level, involving a mix of researcher, administrative, and leadership roles. The prominence of Heads of Research Institutes and faculty-level leaders like Head of Research (Faculty) emphasizes the importance of aligning data governance with the specific needs and objectives of individual faculties (or departments). Table 6.6 provides a detailed perspective on who should hold the role of Research Data Governor.

Table 6.6: Panelist responses on stakeholders in the Research Data Governor role in Round I and II

Stakeholder in Research Data Governor Role	Round I			Round II		
	Rank	Count	Percentage	Rank	Count	Percentage
Head of Research Institute	1	31	66.0%	1	28	82.4%
Head of Research (Faculty)	2	24	51.1%	2	20	58.8%
Principal Investigator/Research Lead	3	18	38.3%	3	10	29.4%
Dean/Head of Faculty	4	10	21.3%	4	5	14.7%
Senior Administrator	5	4	8.5%	5	2	5.9%
*Deputy Dean Research	6	1	2.1%	6	1	2.9%
**Chief Statistician	6	1	2.1%	-	-	-
*Librarian	6	1	2.1%	7	0	0.0%
***Director of Institute/Director General/ Chief Executive Officer/Chief Statistician	-	-	-	5	2	5.9%
Comment, feedback or rewording for clarification:	-	0	0.0%	-	5	14.7%

From Round II

1. P14-Head of research institute is in a more suitable position to ensure policies are adhered.
2. P26-The new response is more accurate.
3. P30-Maybe I was confused during the first time.
4. P43-Head of Research Institute is in a better leadership role towards administering and implementing applicable research data governance policies.
5. P44-After further consideration, the two above selected positions are crucial in ensuring the administration and implementation of research data governance's policies.

Notes.
*Stakeholder identified by panelist in Round I was included as new option for group reassessment in Round II
**Stakeholder identified by panelist in Round I was incorporated into an existing option during Round II
***More options added in Round II based on panelists' responses in Round I

6.2.7 Stakeholders in the Role of Research Data Steward

The survey responses to Question 7 provide evidence that a diverse group of stakeholders is responsible for fulfilling the role of Research Data Steward(s) within RPOs. In Round I, librarians are identified as fulfilling this role in 70.2 percent of responses, with 33 panelists noting their involvement. Principal Investigators/Research Leads closely followed at 61.7 percent (n=29), and Research Officers at 57.4 percent (n=27). Senior Academics were specified by 29.8 percent (n=14), while Mentors/Supervisors were mentioned by 21.3 percent (n=10). Additionally, 14.9 percent (n=7) of the panelists specified Senior Administrators' involvement in research data stewardship, and there is minor representation from Supervisees and a Statistician, each accounting for 2.1 percent, with 1 panelist each. Two comments received from the panelists (P05 & P08) for clarification on few points.

In determining the stakeholder of Research Data Steward role, the Round II findings indicate a varied distribution, with librarians still prominently identified by most of the panelists (82.4%, n=28). This underscores the evolving role of librarians in the management and curation of research data, aligning with their expertise in information organization and access. Principal Investigators or Research Leads were also recognized as Research Data Stewards, as suggested by 61.8 percent (n=21) of the panelists. This emphasizes the need for an active involvement of those generating data in stewardship responsibilities, ensuring a hands-on approach to data management. Similarly, Research Officers were identified as Research Data Stewards by 61.8 percent (n=21) of the panelists, highlighting the importance of dedicated roles specifically focused on managing and overseeing various aspects of research data. In addition, Senior Academics were viewed as playing a role in data stewardship by 32.4 percent (n=11) of them. This suggests that individuals with significant academic experience contribute to ensuring the quality and integrity of research data.

Furthermore, 17.6 percent (n=6) of the panelists specified that Mentors or Supervisors were also identified as Research Data Stewards. Their involvement reflects a mentorship approach to stewardship, guiding and overseeing the responsible management of research data by their supervisees. Senior Administrators were mentioned as Research Data Stewards by 14.7 percent (n=5) of the panelists, signifying a role for administrative staff in the stewardship of research data at a higher level. A Statistician and Supervisees were each identified as a Research Data Steward in 2.9 percent (n=1) of instances.

In general, the findings illustrate a collaborative approach among stakeholders and a diverse range of perspectives on holding the role of Research Data Steward within RPOs. The majority consensus leans towards the top two positions in the rank, namely Librarians and Principal Investigators/Research Leads, along with Research Officers. This aligns with the viewpoints highlighted by some of the panelists (P27, P36, and P43). Table 6.7 showcases a comprehensive breakdown, illustrating the varied distribution of stakeholders serving as Research Data Stewards.

Table 6.7: Panelist responses on stakeholders in Research Data Steward role in Round I and II

Stakeholder in Research Data Steward Role	Round I			Round II		
	Rank	Count	Percentage	Rank	Count	Percentage
Librarian	1	33	70.2%	1	28	82.4%
Principal Investigator/Research Lead	2	29	61.7%	2	21	61.8%
Research Officer	3	27	57.4%	2	21	61.8%
Senior Academic	4	14	29.8%	3	11	32.4%
Mentor/Supervisor	5	10	21.3%	4	6	17.6%
Senior Administrator	6	7	14.9%	5	5	14.7%
*Supervisee	7	1	2.1%	6	1	2.9%
*Statistician	7	1	2.1%	6	1	2.9%
Comment, feedback or rewording for clarification:	-	2	4.3%	-	3	8.8%

From Round I

1. P05-Research team may have their own data manager.
2. P08-Currently, RDG key players and the roles and responsibilities among my colleagues has been large played by journal publishing companies, especially on research data that being attached to articles. Unconsciously, most of the institutional research data shared for scholarly publication has already been managed by publishing company staff and they are taking over on the task of data stewards. These research data are sent along with the publication manuscript on which being uploaded on open repositories such as Zenodo and Figshare (e.g.: Springer account - https://springernature.figshare.com/).

From Round II

1. P27-Lecturers also being appointed as data steward.
2. P36-Librarian may also be important in the data management.
3. P43-Research officer has a more specific job scope and role in this.

Notes.
*Stakeholder identified by panelist in Round I was included as new option for group reassessment in Round II

6.2.8 Administrative Offices supporting RDG

In Question 8, panelists were presented with an open-ended question, asking them to identify other administrative offices within their institution that currently support researchers and various departments in implementing RDG and related procedures within their respective areas of responsibility. In instances where RDG is not yet in place, panelists were also asked to suggest offices they believe should take on this role in supporting RDG initiatives. However, this question was not mandatory. From the responses, as shown in Table 6.8, certain offices showed commonalities, while others were specific to each institution. A list of 13 administrative offices was compiled by consolidating and simplifying office names into more generic categories.

Table 6.8: Panelist responses on Administrative Offices supporting Research Data Governance in Round I

Responses from the Panelists	Administrative Offices supporting RDG (based on more generic names)
1) Center of Excellence 2) Centre for Corporate Strategy and Relations 3) Corporate Strategy and Communications Office for Promoting 4) Data Governance Committee Meeting (JKTDU) / Registrar Office 5) Data Section 6) Dean/Head of School 7) Deputy Dean of Research's Office 8) Deputy Vice Chancellor (Research) 9) Deputy Vice Chancellor (Research) Office 10) Deputy Vice-Chancellor for Research 11) Ethics Committee 12) Ethics Committee, Scientific Dissemination Unit 13) Finance, in term of management of research funds 14) Funders 15) HCM - Human Resource Management & President 16) Human resource 17) Human Resource Department 18) Human Resource Department 19) Human Resources Department 20) Institution and Strategic Planning Center 21) Kulliyah/Faculty - Head of Research Department 22) Office of Registrar 23) Office of Strategic Planning 24) Principal Investigator/Research Lead 25) Project Management Office 26) Project manager 27) R and D Department/Incubation Department 28) RCMO (Pejabat Pengurusan dan Kreativiti Penyelidikan) 29) Research Data Management Centre, Principal Investigator 30) Research Dept of the University 31) Research Integrity and Ethics 32) Research Management Centre/ Office 33) The researcher (data creator) 34) Vice Chancellor (Research) 35) We have data repository system managed by biostatistics sector under officer of manager, NIH. But the role to be central administrator is yet to be implement.	1) Academic Administration and Services Department 2) Centre of Excellence/Centre for Development of Academic Excellence 3) Corporate Strategy and Communications Office 4) Faculty/Kulliyyah/School 5) Human Resources Department 6) Incubation Department/Centre for Innovation/Research and Development (R&D) Department 7) Legal Unit/Office/Department/Division 8) Office of Deputy Dean (Research) 9) Office of Deputy Vice Chancellor (Research) 10) Office of Research Integrity and Ethics 11) Project Management Office 12) Quality and Accreditation Unit/Office/Department/Division 13) Research Data Management Unit/Centre/Department

The consolidation was conducted for re-evaluation in Round II serves as a guide for developing an RDG framework for Malaysia RPOs. The findings indicate that offices or departments supporting RDG as part of the Central Administration are diverse and encompass crucial functions within an organization. The scores are nearly equal, signifying that these offices and departments carry similar weight in supporting RDG. For example, organizations with an intact Research Data Management

Unit/Centre/Department were identified as a central administration entity by 23.5 percent (n=8) of the panelists. This underscores the necessity for specialized units or centers dedicated to overseeing and managing research data across various disciplines.

Following closely, the Office of Research Integrity and Ethics was mentioned by 20.6 percent (n=7) of the panelists, emphasizing its central role in administration. This highlights the crucial role of ethical considerations in the overarching administration of research data. These results align with the responses for Question 4 on the Research Data Governance Office, which received most votes.

Additionally, Faculties/Kulliyyahs/Schools were recognized as central administration entities by 17.6 percent (n=6) of the panelists, indicating a decentralized approach where individual academic units contribute to the central administration of research data. The Office of Deputy Vice Chancellor (Research) was involved in central administration in 17.6 percent (n=6) of instances, signifying its significance in RDG as a top-level institutional approach aligning research data administration with broader research leadership.

The Human Resources Department was identified by 17.6 percent (n=6) of the panelists, suggesting a role in central administration related to workforce development, training, incentives, and compliance aspects of RDM. The Office of Deputy Dean (Research) and Project Management Office each play a role in central administration, as specified by 14.7 percent (n=5) of the panelists, indicating a shared responsibility for overseeing and coordinating RDM efforts.

The Corporate Strategy and Communications Office was chosen by 11.8 percent (n=4) of them, suggesting a role in central administration related to strategic planning and communication of research data initiatives. The Quality and Accreditation Unit/Office/Department/Division, Legal Unit/Office/Department/Division, and Incubation Department/Centre for Innovation/Research and Development (R&D) Department each play a role in central administration, as specified by 5.9 percent (n=2) of the panelists. Although they have fewer votes, this signifies specialized units contributing to the central administration of research data with a focus on quality, legal aspects, and innovation, respectively. However, the Academic Administration and Services Department and Centre of Excellence/Centre for Development of Academic Excellence did not appear in the identified role for Central Administration supporting RDG, suggesting a specific distribution of administrative responsibilities in this context.

To sum up, the distribution of Central Administration role reflects a blend of specialized units, top-level leadership, faculties, and administrative offices. The existence of dedicated RDM units and the involvement of various departments underscore the importance of a multidimensional approach to central administration in the context of RDG. Thus, the findings emphasize that other than the Library, Information Technology Office, and Research Management Office, each office identified in Table 6.9 as contributing to RDG is perceived as crucial and relevant to the initiatives.

Table 6.9: Panelist responses on Administrative Offices supporting Research Data Governance in Round II

Administrative Offices Supporting RDG	Count	Percentage
Research Data Management Unit/Centre/Department	8	23.5%
Office of Research Integrity and Ethics	7	20.6%
Faculty/Kulliyyah/School	6	17.6%
Office of Deputy Vice Chancellor (Research)	6	17.6%
Human Resources Department	6	17.6%
Office of Deputy Dean (Research)	5	14.7%
Project Management Office	5	14.7%
Corporate Strategy and Communications Office	4	11.8%
Quality and Accreditation Unit/Office/Department/Division	2	5.9%
Legal Unit/Office/Department/Division	2	5.9%
Incubation Department/Centre for Innovation/Research and Development (R&D) Department	2	5.9%
Academic Administration and Services Department	0	0.0%
Centre of Excellence/Centre for Development of Academic Excellence	0	0.0%

6.3 Identifying Structural Positions

How do data practitioners identify the structural positions of research data governance roles within research performing organizations? RQ4 was addressed in two rounds, Round III and Round IV. The same questions that were posed in Round III related to addressing this RQ4, were presented again to the panelists who participated in Round IV. They were asked to either retain or revise their previous responses based on the findings from Round III. Apart from these questions on structural positions, the panelists were also tasked with formulating research-based descriptions of roles. Additionally, they were required to provide feedback on the appropriate nomenclatures for the roles, tailored to the RDG context. The adoption of research-based descriptions and nomenclatures is crucial for the development of the RDG framework. All the questions aimed to arrive at a consensus regarding the most effective structural positions, descriptions, and nomenclatures.

6.3.1 Executive Sponsor

In Round III, responses to the first question regarding the appropriate level of governance for an Executive Sponsor within RPOs revealed that 39.4 percent (n=13) of the panelists advocate for the "Strategic" level. Notably, 36.4 percent (n=12) proposed a dual role at both "Strategic & Tactical" levels, while 15.2 percent (n=5) argued for engagement across all three levels: "Strategic, Tactical, & Operational". A smaller portion, 6.1 percent (n=2), suggested involvement of the Executive Sponsor at the "Tactical & Operational" levels, a minimal 3 percent (n=1) recommended sole positioning at the "Tactical" level, and all panelists agreed that the role should not be positioned at the "Operational" level since there was none of them selected this option.

Meanwhile, in Round IV, the majority of the panelists, 61.5 percent (n=16), advocated for positioning the Executive Sponsor at the "Strategic" level. This marks an increase from Round III, where only 39.4 percent preferred the "Strategic" level. The preference for "Strategic & Tactical" decreased from 36.4 percent in the Round III to 23.1 percent (n=6) in Round IV. However, it is noteworthy that this still indicates a belief that the Executive Sponsor's role involves not only high-level strategic decisions but also some involvement in the tactical aspects of governance. The choice of "Strategic, Tactical & Operational" remained consistent at 15.4 percent (n=4), implying a more comprehensive involvement across different governance layers. It is notable that there were no responses for the last three options in Round IV, indicating a clear alignment toward strategic involvement for the Executive Sponsor and a diminished focus on operational considerations. The shift towards a higher preference for the "Strategic" level suggests a potential evolving perspective among the panelists regarding the role of the Executive Sponsor. This shift may underscore a perceived need for more strategic direction and oversight in the governance structure. Table 6.10 provides a detailed distribution of

responses among panelists on the position of an Executive Sponsor in the governance structure.

Table 6.10: Panelist responses on level of governance for an Executive Sponsor in Round III and IV

Governance Level for Executive Sponsor	Round III			Round IV		
	Rank	Count	Frequency	Rank	Count	Frequency
Strategic	1	13	39.4%	1	16	61.5%
Strategic & Tactical	2	12	36.4%	2	6	23.1%
Strategic, Tactical & Operational	3	5	15.2%	3	4	15.4%
Tactical & Operational	4	2	6.1%	4	0	0.0%
Tactical	5	1	3.0%	4	0	0.0%
Operational	6	0	0.0%	4	0	0.0%
TOTAL	**-**	**33**	**[a]100%**	**-**	**26**	**[a]100%**

Notes.
[a]Percentages were rounded to the nearest whole number

6.3.2 Data Governance Leader

In Round III, the panelists were then asked the next question to specify the ideal governance position for the Data Governance Leader role. The findings show that the majority of the panelists, 42.4 percent (n=14), believed that the Data Governance Leader should be positioned at both the "Strategic & Tactical" levels of governance. A substantial portion, 27.3 percent (n=9), suggested that the Data Governance Leader should be involved at all three levels of governance: "Strategic, Tactical, & Operational". Six (18.2%) of the panelists believed that the Data Governance Leader should be positioned primarily at the "Strategic" level, focusing on high-level decision-making and overarching governance strategy. Two (6.1%) suggested that the Data Governance Leader should be positioned at the "Tactical" level. Another 6.1 percent (n=2) believed that the Data Governance Leader should be involved at both the "Tactical & Operational" levels, focusing on more immediate and specific operational aspects of governance. No responses received for the "Operational" level of governance. Overall, the data indicates a diversity of perspectives on the appropriate level of governance for the Data Governance

Leader. The most common belief is a dual focus on “Strategic & Tactical” levels, emphasizing the need for a balance between high and middle level strategy involvement.

Comparing this data to the previous Round III, it is interesting to note that the percentage of the panelists favoring “Strategic & Tactical” has increased from 42.4 percent to 61.5 percent (n=16), indicating a shift in preferences, thus, making it the predominant choice in Round IV. This indicates a perspective that emphasizes a dual focus on high and middle level strategy. The percentage for “Strategic, Tactical & Operational” has decreased from 27.3 percent to 23.1 percent (n=6). The option “Strategic” at 18.2 percent in Round III and decreased to 11.5 percent (n=3) in this round. While, “Tactical & Operational” remained only 1 response (3.8%) if compared to Round III with 2 responses (6.1%). Notably, the option “Tactical” and “Operational” had no responses in this round. The results suggest a clear preference for situating the role of Data Governance Leader at the strategic level, with active engagement in the tactical level. However, it is considered inappropriate for placement the role at the operational level. Overall, Round IV data suggests a stronger preference for positioning the Data Governance Leader at the “Strategic & Tactical” levels. Table 6.11 presents a comprehensive breakdown of responses from panelists regarding the Data Governance Leader role in the governance structure.

Table 6.11: Panelist responses on level of governance for a Data Governance Leader in Round III and IV

Governance Level for Data Governance Leader	Round III			Round IV		
	Rank	Count	Frequency	Rank	Count	Frequency
Strategic & Tactical	1	14	42.4%	1	16	61.5%
Strategic, Tactical & Operational	2	9	27.3%	2	6	23.1%
Strategic	3	6	18.2%	3	3	11.5%
Tactical	4	2	6.1%	5	0	0.0%
Tactical & Operational	4	2	6.1%	4	1	3.8%
Operational	5	0	0.0%	5	0	0.0%
TOTAL	**-**	**33**	**[a]100%**	**-**	**26**	**[a]100%**

Notes.
[a]Percentages were rounded to the nearest whole number

6.3.3 Research Data Governance Committee

In the next question of Round III, which inquiries about the appropriate structural position of the Research Data Governance Committee, the majority preference among the panelists is to position the committee at the “Tactical & Operational” level, as shown in Table 6.12. This choice received 36.4 percent of the responses (n=12). The second most favoured choice is “Strategic, Tactical & Operational”, chosen by 27.3 percent (n=9). The remaining options include “Strategic & Tactical”, chosen by 18.2 percent (n=6), the “Strategic” level with 12.1 percent (n=4) of the panelists, and “Tactical” which obtained 6.1 percent (n=2) of the panelists. While none of them opted to position the role at the “Operational” level. Overall, the findings indicate diverse perspectives on the appropriate level of governance for the Research Data Governance Committee. The prevalence of responses at multiple levels suggests that panelists view the Committee’s role as multi-faceted, potentially requiring engagement across different governance layers.

In the analysis of responses from Round IV for the same question, the findings reveal a notable shift in the panelists’ preferences for the positioning of the Research Data Governance Committee role. The most favored choice remained at the “Tactical & Operational” level, experiencing a substantial increase from 36.4 percent to 69.2 percent (n=18). The second most popular option is “Strategic, Tactical & Operational”, selected by 19.2 percent (n=5). Conversely, the option “Strategic & Tactical” has decreased from 18.2 percent to 3.8 percent (n=1), and “Strategic” has also decreased from 12.1 percent to 7.7 percent (n=2). Interestingly, there were no responses indicating a preference for the “Tactical” and “Operational” levels in Round IV. It suggests that the position should be situated at a joint "Tactical & Operational” levels, rather than being separated into individual tiers.

The provided comment by P30 believes that the Committee should be involved in strategic decisions, emphasizing a broader role beyond tactical and operational aspects. However, both findings from Rounds III and IV underscore a consistent and strengthened preference for the "Tactical & Operational" level in positioning the Research Data Governance Committee. This preference implies a focus on practical and specific operational considerations.

Table 6.12: Panelist responses on level of governance for Research Data Governance Committee in Round III and IV

Governance Level for RDG Committee	Round III			Round IV		
	Rank	Count	Frequency	Rank	Count	Frequency
Tactical & Operational	1	12	36.4%	1	18	69.2%
Strategic, Tactical & Operational	2	9	27.3%	2	5	19.2%
Strategic & Tactical	3	6	18.2%	4	1	3.8%
Strategic	4	4	12.1%	3	2	7.7%
Tactical	5	2	6.1%	5	0	0.0%
Operational	6	0	0.0%	5	0	0.0%
TOTAL	**-**	**33**	**[a]100%**	**-**	**26**	**[a]100%**

Comment from Round IV

1. P30-I think the committee should also be involved in strategic, not only tactical and operations.

Notes.

[a]Percentages were rounded to the nearest whole number

6.3.4 Office of Research Data Governance

Proceeding to the next question on positioning the Office of Research Data Governance, the findings from Round III (as in Table 6.13) indicate that the highest percentage of panelists, 39.4 percent (n=13), believed that the Office of Research Data Governance should be positioned at both the "Tactical & Operational" levels of governance. The second most favored option is "Strategic, Tactical & Operational" at 27.3 percent (n=9). Meanwhile, 15.2 percent (n=5) believed that the Office of Research Data Governance should primarily operate at the "Operational" level, focusing on specific processes and activities. Three (9.1%) of the panelists believed that the Office of Research Data Governance should be positioned at both the "Strategic & Tactical" levels. Two (6.1%) suggested that the Office of Research Data Governance should be positioned

primarily at the “Tactical” level. The remaining 3 percent (n=1) believed that the Office of Research Data Governance should primarily operate at the “Strategic” level.

Meanwhile, in Round IV, the majority of the panelists hold the view that the structural position of the Office of Research Data Governance still remains at the “Tactical & Operational” levels, showing a significant increase from 39 percent to 65.4 percent (n=17). The second most favoured option is “Strategic, Tactical & Operational” at 19.2 percent (n=5). The option “Operational” has decreased from 15.2 percent to 11.5 percent (n=3). The options “Strategic & Tactical” has 1 response (3.8%), while “Tactical”, and “Strategic” have no responses in this round. It indicates that positioning the Office of Research Data Governance at the strategic level is not the preferred choice. Additionally, it cannot solely operate at the tactical level without including the operational level. The findings highlight a noticeable shift towards a stronger preference for the “Tactical & Operational” levels for the Office of Research Data Governance in Round IV.

Table 6.13: Panelist responses on level of governance for Office of Research Data Governance in Round III and IV

Governance Level for RDG Office	Round III			Round IV		
	Rank	**Count**	**Frequency**	**Rank**	**Count**	**Frequency**
Tactical & Operational	1	13	39.4%	1	17	65.4%
Strategic, Tactical & Operational	2	9	27.3%	2	5	19.2%
Operational	3	5	15.2%	3	3	11.5%
Strategic & Tactical	4	3	9.1%	4	1	3.8%
Tactical	5	2	6.1%	5	0	0.0%
Strategic	6	1	3.0%	5	0	0.0%
TOTAL	**-**	**33**	**[a]100%**	**-**	**26**	**[a]100%**

Notes.
[a]Percentages were rounded to the nearest whole number

6.3.5 Research Data Owner

In Round III, the panelists were also asked about the suitable governance structure position for the Research Data Owner within RPOs. The highest responses are for positioning the Research Data Owner at the "Operational" level (39.4%, n=13), followed by the "Tactical & Operational" levels (33.3%, n=11). Additionally, a significant portion of panelists (24.2%, n=8) believed the Research Data Owner should be positioned at the "Strategic, Tactical & Operational" levels. While the "Tactical" level had a lower percentage at 3 percent (n=1). Overall, the findings indicate diverse perspectives on the appropriate level of governance for the Research Data Owner, with a notable emphasis on the operational and tactical levels of governance. There is only one comment from one of the panelists (P05), and it is solely for clarification.

While, in Round IV, the findings indicate an increase in positioning the Research Data Owner at the "Operational" level, with 65.4 percent (n=17) of the panelists selecting this option. This is followed by the "Tactical & Operational" levels, chosen by 30.8 percent (n=8) of panelists. Only 3.8 percent (n=1) of them opted for "Strategic, Tactical & Operational". Notably, there were no responses for the "Strategic", "Tactical", and "Strategic & Tactical" levels.

Furthermore, there is one comment made by P32 offering a perspective on the division of responsibilities. It suggests that the Research Data Owner should primarily focus on creating the data and knowledge, thus, more suitable to be positioned at the "Operational" level. While those at the strategic level should manage data information and plan for future data-related activities. This comment adds valuable context to the perspectives of the panelists regarding the governance level of the Research Data Owner, underscoring the importance of delineating roles and responsibilities between the operational and strategic levels. A comprehensive overview is presented in Table 6.14.

Table 6.14: Panelist responses on level of governance for Research Data Owner in Round III and IV

Governance Level for Research Data Owner	Round III			Round IV		
	Rank	Count	Frequency	Rank	Count	Frequency
Operational	1	13	39.4%	1	17	65.4%
Tactical & Operational	2	11	33.3%	2	8	30.8%
Strategic, Tactical & Operational	3	8	24.2%	3	1	3.8%
Tactical	4	1	3.0%	4	0	0.0%
Strategic	5	0	0.0%	4	0	0.0%
Strategic & Tactical	5	0	0.0%	4	0	0.0%
TOTAL	**-**	**33**	**[a]100%**	**-**	**26**	**[a]100%**
Comment, feedback or rewording for clarification:	-	1	3.0%	-	1	3.8%

From Round III:

1. P05-This is crucial to ensure ability of all levels of the ministry to conduct study and becoming primary investigator.

From Round IV:

1. P32-Research Data Owner will be focusing on the creating the data and know-how, while strategic should be the one who know the data information and arrange the future planning for the data.

Notes.
[a]Percentages were rounded to the nearest whole number

6.3.6 Research Data Governor

In Round III, majority of the panelists reached a consensus on positioning the Research Data Governor at the "Tactical & Operational" levels, with 36.4 percent (n=12) choosing this option. Following closely, the second highest number of panelists, at 21.2 percent (n=7), believed that this role should be positioned under "Strategic, Tactical & Operational". Other groups expressed varied opinions, suggesting that the Research Data Governor could be positioned under "Strategic & Tactical", "Tactical", "Operational", and "Strategic", each with different percentages. Examining the Round III survey data reveals a diversity of opinions regarding the appropriate level of governance for the Research Data Governor. While a significant number of panelists endorse its involvement at the "Tactical & Operational" levels, there is also substantial support for strategic responsibilities.

Meanwhile, in Round IV, the overwhelming majority of panelists (80.8%, n=21) believed that the Research Data Governor should be positioned at the "Tactical & Operational" level. The other options, including "Strategic, Tactical & Operational", "Strategic & Tactical", and "Tactical", have much lower percentages. Notably, there were no panelists who chose "Operational" or "Strategic" in this round. All panelists reached a consensus that the role should not be situated at the strategic level. However, it is also acknowledged that the role cannot solely exist in a distinct operational level; rather, it should be positioned at both the tactical and operational levels. Therefore, the findings indicate a strong preference for positioning the Research Data Governor at the "Tactical & Operational" level in Round IV of the survey. The following Table 6.15 provides an analysis of the distribution of responses among panelists regarding the structural position of the Research Data Governor.

Table 6.15: Panelist responses on level of governance for Research Data Governor in Round III and IV

Governance Level for Research Data Governor	Round III			Round IV		
	Rank	Count	Frequency	Rank	Count	Frequency
Tactical & Operational	1	12	36.4%	1	21	80.8%
Strategic, Tactical & Operational	2	7	21.2%	2	2	7.7%
Strategic & Tactical	3	6	18.2%	3	1	3.8%
Tactical	4	4	12.1%	2	2	7.7%
Operational	5	2	6.1%	4	0	0.0%
Strategic	5	2	6.1%	4	0	0.0%
TOTAL	**-**	**33**	**[a]100%**	**-**	**26**	**[a]100%**

Notes.
[a]Percentages were rounded to the nearest whole number

6.3.7 Research Data Steward

For the Research Data Steward role, questions were designed to investigate whether there is a need to place the Chief Research Data Steward and the Research Data Steward in different positions. The discussion initiates with the findings related to the Chief Research Data Steward as illustrated in Table 6.16, followed by insights into the Research Data Steward, as in Table 6.17.

In Round III, the majority of panelists (33.3%, n=11) expressed the belief that the Chief Research Data Steward should be positioned at the “Strategic, Tactical & Operational” level of governance. Meanwhile, the “Tactical & Operational” level is the second highest choice with 27.3 percent (n=9) of the panelists selecting this position. Combining “Strategic & Tactical” levels for the Chief Research Data Steward received 15.2 percent (n=5) responses, and separately positioning the role at “Strategic” and “Tactical” levels both had equal responses at 9.1 percent (n=3) each. The “Operational” level had the least responses, with 6.1 percent (n=2).

In Round IV, an increasing majority of panelists (69.2%, n=18) believed that the Chief Research Data Steward should be positioned at the “Strategic, Tactical & Operational” level. The provided comment by P32 suggests that the Chief Research Data Steward should be someone capable of understanding all aspects of data and its management. Some panelists maintained the responses as in the previous round, with 19.2 percent (n=5) advocating for “Tactical & Operational”, and 11.5 percent (n=3) for “Strategic & Tactical”. Notably, there were no responses supporting “Strategic”, “Tactical”, or “Operational” positions in this round. The panelists unanimously agreed that this role cannot be confined to a single, distinct position.

In contrast, the results for the Research Data Steward in Round III show that the majority of panelists preferred positioning the Research Data Steward at the “Tactical & Operational” level, with 45.5 percent (n=15) choosing this option. The second most favoured option is “Operational” at 30.3 percent (n=10). Other options, including “Strategic, Tactical & Operational”, “Tactical”, and “Strategic”, received varying percentages. Interestingly, no panelist opted for “Strategic & Tactical” in this round.

Moving to Round IV, there is a notable increase in the number of panelists who agreed to position the Research Data Steward at the "Tactical & Operational" level (76.9%, n=20). The option "Operational" maintains a presence at 23.1 percent (n=6). However, in this round, no panelists selected "Strategic, Tactical & Operational", "Tactical", "Strategic", or "Strategic & Tactical". They unanimously agreed that this role should not be at the strategic level. The data from Round IV indicates a strong and growing preference for positioning the Research Data Steward at the "Tactical & Operational" level. The comprehensive analysis of responses for this question is provided in Table 6.17.

Table 6.16: Panelist responses on level of governance for Chief Research Data Steward in Round III and IV

Governance Level for Chief Research Data Steward	Round III			Round IV		
	Rank	Count	Frequency	Rank	Count	Frequency
Strategic, Tactical & Operational	1	11	33.3%	1	18	69.2%
Tactical & Operational	2	9	27.3%	2	5	19.2%
Strategic & Tactical	3	5	15.2%	3	3	11.5%
Strategic	4	3	9.1%	4	0	0.0%
Tactical	4	3	9.1%	4	0	0.0%
Operational	5	2	6.1%	4	0	0.0%
TOTAL	**-**	**33**	**[a]100%**	**-**	**26**	**[a]100%**
Comment, feedback or rewording for clarification:	-	0	0.0%	-	1	3.8%

From Round IV:

1. P32-Person who can know all the data and its management.

Notes.
[a]Percentages were rounded to the nearest whole number

Table 6.17: Panelist responses on level of governance for Research Data Steward in Round III and IV

Governance Level for Research Data Steward	Round III			Round IV		
	Rank	Count	Frequency	Rank	Count	Frequency
Tactical & Operational	1	15	45.5%	1	20	76.9%
Operational	2	10	30.3%	2	6	23.1%
Strategic, Tactical & Operational	3	5	15.2%	3	0	0.0%
Tactical	4	2	6.1%	3	0	0.0%
Strategic	5	1	3.0%	3	0	0.0%
Strategic & Tactical	6	0	0.0%	3	0	0.0%
TOTAL	**-**	**33**	**[a]100%**	**-**	**26**	**[a]100%**

Notes.
[a]Percentages were rounded to the nearest whole number

6.3.8 Research Data Custodian

During Round III, one participant did not provide a response, resulting in an analysis based on 32 received responses. Findings from this round suggest that most of the panelists believed the Research Data Custodian is best positioned at the “Tactical & Operational” level, with 43.8 percent (n=14) choosing this option. This indicates that a significant number of panelists see the custodian role encompassing both tactical and operational responsibilities. Following closely are the options “Operational” and “Strategic, Tactical & Operational”, each selected by 18.8 percent (n=6) of the panelists. Additionally, 9.4 percent (n=3) believed the role is best suited at the “Tactical” level, while the remaining panelists chose “Strategic & Tactical” and “Strategic” levels, with 6.3 percent (n=2) and 3.1 percent (n=1), respectively.

While in Round IV, it is evident that some panelists changed their preference, as there is an increase in the number of panelists (76.9%, n=20) choosing the option that the Research Data Custodian should be positioned at the “Tactical & Operational” level. Other options, including “Operational”, “Strategic, Tactical & Operational”, and “Tactical”, have lesser percentages, with 11.5 percent (n=3), 7.7 percent (n=2), and 3.8 percent (n=1), respectively. Meanwhile, “Strategic & Tactical” and “Strategic” levels of governance had no responses. Overall, the data indicates a strong and consistent preference for positioning the Research Data Custodian at the “Tactical & Operational” level in the following round of the survey, emphasizing a focus on operational and tactical aspects rather than strategic decision-making. As supported by a comment made by P29 suggests that the Research Data Custodian does not necessarily need to be involved in the strategic decision-making. Refer to Table 6.18 for a detailed breakdown of panelist responses on the structural position of a Research Data Custodian.

Table 6.18: Panelist responses on level of governance for Research Data Custodian in Round III and IV

Governance Level for Research Data Custodian	Round III			Round IV		
	Rank	Count	Frequency	Rank	Count	Frequency
Tactical & Operational	1	14	43.8%	1	20	76.9%
Operational	2	6	18.8%	2	3	11.5%
Strategic, Tactical & Operational	2	6	18.8%	3	2	7.7%
Tactical	3	3	9.4%	4	1	3.8%
Strategic & Tactical	4	2	6.3%	5	0	0.0%
Strategic	5	1	3.1%	5	0	0.0%
TOTAL	**-**	**32***	**[a]100%**	**-**	**26**	**[a]100%**
Comment, feedback or rewording for clarification:	-	0	0%	-	1	3%

From Round IV:

1. P29-Do not have to involve in strategic decision.

Notes.
*One missing value.
[a]Percentages were rounded to the nearest whole number

6.3.9 Researcher

Table 6.19 has a detailed summary of what panelists think about positioning the Researcher role in the governance structure. In Round III, most of the panelists (60.6%, n=20) believed that Researcher should be positioned at the "Operational" level of governance. Nine panelists (27.3%) have selected the "Strategic, Tactical & Operational" level, indicating the belief that Researcher should have responsibilities across all three levels of governance. Additionally, four of them (12.1%) believed the role should be positioned at the "Tactical & Operational". Notably, no responses were received for other options, indicating a unanimous belief among panelists that Researcher should be involved in the operational aspects of activities.

Moving to Round IV, the number of panelists maintaining the belief that Researcher should be positioned at the "Operational" level remains the same at 20 (76.9%) panelists. The choice of "Strategic, Tactical & Operational" followed at 15.4 percent (n=4), and "Tactical & Operational" at 7.7 percent (n=2). Like Round III, the rest of the options received no responses. When comparing the data between both rounds, there is consistency in the preference for positioning Researcher role at the "Operational" level,

with a similar number of panelists (n=20) expressing this preference in both rounds. The data suggests a prevailing belief that Researcher primarily operate at the "Operational" level of governance.

Table 6.19: Panelist responses on level of governance for Researcher in Round III and IV

Governance Level for Researcher	Round III			Round IV		
	Rank	Count	Frequency	Rank	Count	Frequency
Operational	1	20	60.6%	1	20	76.9%
Strategic, Tactical & Operational	2	9	27.3%	2	4	15.4%
Tactical & Operational	3	4	12.1%	3	2	7.7%
Strategic	4	0	0.0%	4	0	0.0%
Tactical	4	0	0.0%	4	0	0.0%
Strategic & Tactical	4	0	0.0%	4	0	0.0%
TOTAL	**-**	**33**	**[a]100%**	**-**	**26**	**[a]100%**

Notes.
[a]Percentages were rounded to the nearest whole number

6.3.10 Research Data Consumer

In Round III, most of the panelists (63.6%, n=21) believed that the Research Data Consumer should be positioned at the "Operational" level. Another 12 panelists, constituting 21.2 percent (n=7) and 15.2 percent (n=5), respectively, chose either "Tactical & Operational" or "Strategic, Tactical & Operational" as the appropriate level. Notably, no responses were received for other options.

Moving to Round IV, only two options were selected: "Operational" and "Tactical & Operational". The overwhelming majority (92.3%, n=24) believed that the Research Data Consumer should be positioned at the "Operational" level of governance. A small percentage of panelists (7.7%, n=2) chose the "Tactical & Operational" level for Research Data Consumer. Overall, in Round IV, the predominant view is that Research Data Consumer should primarily have operational responsibilities. The belief in the "Tactical & Operational" level, although smaller, suggests that some panelists still see a combined tactical and operational role for Research Data Consumer. However, the absence of their

selection of options involving a strategic or exclusion of tactical governance levels indicates a shift in emphasis toward greater operational involvement for Research Data Consumer. Table 6.20 provides a detailed distribution of panelist responses to the question.

Table 6.20: Panelist responses on level of governance for Research Data Consumer in Round III and IV

Governance Level for Research Data Consumer	Round III			Round IV		
	Rank	Count	Frequency	Rank	Count	Frequency
Operational	1	21	63.6%	1	24	92.3%
Tactical & Operational	2	7	21.2%	2	2	7.7%
Strategic, Tactical & Operational	3	5	15.2%	3	0	0.0%
Strategic	4	0	0.0%	3	0	0.0%
Tactical	4	0	0.0%	3	0	0.0%
Strategic & Tactical	4	0	0.0%	3	0	0.0%
TOTAL	**-**	**33**	**[a]100%**	**-**	**26**	**[a]100%**

Notes.
[a]Percentages were rounded to the nearest whole number

6.3.11 Administrative Offices

Table 6.21 shows a comprehensive breakdown of responses among panelists regarding the governance structure of Administrative Offices. In Round III, the panelists favored positioning the Administrative Offices at the “Operational” level, with 60.6 percent (n=20) of them choosing this option. Additionally, 21.2 percent (n=7) supported the “Strategic, Tactical & Operational” level, indicating a belief that administrative roles should encompass a combination of strategic, tactical, and operational responsibilities. Meanwhile, the “Tactical & Operational” level received 12.1 percent (n=4) support. Both “Strategic” and “Tactical” levels had one panelist each (3%) selecting those options, and none advocated for the “Strategic & Tactical” level.

In Round IV, a consistent number of panelists, matching the previous round at 76.9% (n=20), continued to assert that the appropriate governance level for the Administrative Offices should be at the "Operational" level. This perspective was supported by a comment by P32 stating, "*Administrative Offices should focus on the operational part*". The other options, including "Strategic, Tactical & Operational", "Tactical & Operational", "Strategic", and "Tactical" had lower percentages, and notably, "Strategic", "Tactical", and "Strategic & Tactical" received no responses. Therefore, it is concluded that the findings strongly indicate a consensus among respondents that Administrative Offices should primarily positioned at the "Operational" level of governance.

Table 6.21: Panelist responses on level of governance for Administrative Offices in Round III and IV

Governance Level for Administrative Offices	Round III			Round IV		
	Rank	**Count**	**Frequency**	**Rank**	**Count**	**Frequency**
Operational	1	20	60.6%	1	20	76.9%
Strategic, Tactical & Operational	2	7	21.2%	2	3	11.5%
Tactical & Operational	3	4	12.1%	2	3	11.5%
Strategic	4	1	3.0%	3	0	0.0%
Tactical	4	1	3.0%	3	0	0.0%
Strategic & Tactical	5	0	0.0%	3	0	0.0%
TOTAL	**-**	**33**	**[a]100%**	**-**	**26**	**[a]100%**
Comment, feedback or rewording for clarification:	-	0	0.0%	-	1	3.8%

From Round IV:

1. P32-Administrative Offices should focus on operational part.

Notes.

[a]Percentages were rounded to the nearest whole number

6.3.12 External Bodies

In Round III, the panelists expressed varied preferences regarding the level of governance for External Bodies. The most popular choice is the "Strategic" level, with 33.3 percent (n=11) of them selecting this option. Other choices, including "Operational", "Strategic & Tactical", "Tactical", "Strategic, Tactical & Operational", and "Tactical & Operational", received varying percentages.

Moving to Round IV, a substantial majority of panelists (65.4%, n=17) held the belief that External Bodies, including National Funders and National/Institutional Repositories, should be positioned at the "Strategic" level of governance. Support for "Strategic & Tactical" was at 15.4 percent (n=4), while both "Operational" and "Strategic, Tactical & Operational" had 7.7 percent (n=2) support each. The "Tactical" level received 3.8 percent (n=1) support, and no panelists indicated support for the "Tactical & Operational" level. A comment by P29 suggests that this panelist views the governance of External Bodies as more appropriate at the strategic level, particularly on a national level, which aligns with the group's responses. Overall, in Round IV, the prevailing view is that External Bodies should be positioned at the "Strategic" level of governance, aligning with the comment. Table 6.22 provides a comprehensive view of the distribution of panelist responses.

Table 6.22: Panelist responses on level of governance for External Bodies in Round III and IV

Governance Level for External Bodies	Round III			Round IV		
	Rank	Count	Frequency	Rank	Count	Frequency
Strategic	1	11	33.3%	1	17	65.4%
Operational	2	6	18.2%	2	2	7.7%
Strategic & Tactical	2	6	18.2%	3	4	15.4%
Tactical	3	4	12.1%	5	1	3.8%
Strategic, Tactical & Operational	4	3	9.1%	4	2	7.7%
Tactical & Operational	4	3	9.1%	6	0	0.0%
TOTAL	**-**	**33**	**[a]100%**	**-**	**26**	**[a]100%**
Comment, feedback or rewording for clarification:	-	0	0.0%	-	1	3.8%

From Round IV:

1. P29-Strategic level is more appropriate for a national level.

Notes.

[a]Percentages were rounded to the nearest whole number

6.4 Research-Derived Descriptions for Research Data Governance Roles

In Rounds III and IV, in addition to asking the panelists about the descriptions of RDG areas and decision domains, as deliberately discussed in Chapter 5, inquiries were also made regarding the descriptions of RDG roles.

The outcomes of Round III indicate the panelists' overall acceptance and agreement with the majority of the initial RDG role descriptions. Among the 11 descriptions, three of RDG roles received proposed new descriptions, which were then incorporated into Round IV for reassessment. During this phase, the panel had the opportunity to reconsider and make decisions regarding the acceptance or rejection of the revised descriptions. Whereas the remaining descriptions were confirmed as standard research descriptions through the Delphi study. The panelists evaluated these initial RDG role descriptions, as detailed in Table 6.23.

Only these three items were included in Round IV for re-evaluation. According to the results, the panelists reached a majority agreement on the revised descriptions for all RDG roles in this round. Consequently, these descriptions, along with those that secured unanimous agreement in the previous round, were officially endorsed as standard descriptions. The results are presented in Table 6.24.

Table 6.23: Assessment of descriptions of RDG roles in Round III

ID	ITEM	INITIAL DESCRIPTION	ACCEPTED		REVISED		REVISED DESCRIPTION
			Count	%	Count	%	
R1	Executive Sponsor	A top senior executive in charge of coordinating research data governance activities and programs.	32	97%	1	3%	A top senior executive in charge of overseeing research data governance policy and funding, and strategizing activities and programs.
R2	Data Governance Leader	Accountable for the overall management of the organization's data and/or information governance.	33	100%	-	-	-
R3	Research Data Governance Committee	The Research Data Governance Committee is designed to comprise representatives from the major strategic stakeholders across the institution. The work of the committee is coordinated by the Research Data Governance Leader.	33	100%	-	-	-
R4	The Office of Research Data Governance	Responsible for monitoring Key Risk Indicators relating to incidences of data misconduct.	32	97%	1	3%	Responsible for ensuring the application of data governance principles by promoting them, facilitating best practices, and monitoring instances of data misconduct.
R5	Research Data Owner	Accountable for research data and materials including, but not limited to the collection, development, maintenance, distribution and security of research data and materials.	32	97%	1	3%	Accountable for research data and materials including, but not limited to the collection, development, maintenance, distribution, preservation, reuse, and security of research data and materials.
R6	Research Data Governor	Accountable for the implementation of research data governance policy in their colleges or schools or departments.	33	100%	-	-	-

Table 6.23, continued

ID	ITEM	INITIAL DESCRIPTION	ACCEPTED		REVISED		REVISED DESCRIPTION
			Count	%	Count	%	
R7	Research Data Steward	Accountable for the quality and integrity, implementation, and enforcement of research data management within their research project. Every research data set may have one or more Research Data Stewards.	33	100%	-	-	-
R8	Research Data Custodians	Established to ensure that important datasets are developed, maintained, and are accessible within their defined specifications. Designating a person or agency as being in charge of overseeing these aspects of data management helps to ensure that datasets do not become compromised.	33	100%	-	-	-
R9	Researcher	Any individuals including staff members, collaborators, adjuncts, academic title holders and students who are involved in conducting research.	33	100%	-	-	-
R10	Research Data Consumer	Any staff member, contractor, consultant, third party, or authorized agent who accesses, inputs, amends, deletes, extracts or analyses research data. They are not generally involved in the governance process but are responsible for the quality assurance of data.	33	100%	-	-	-
R11	Administrative Offices	Responsible for supporting researchers and schools/departments/offices in implementing research data governance and associated procedures within their areas of responsibility.	33	100%	-	-	-

Table 6.24: Re-assessment of descriptions of RDG roles in Round IV

NO.	ITEM ID	REVISED DESCRIPTION	USED REVISED DESCRIPTION		RETAINED INITIAL DESCRIPTION	
			Count	Percentage	Count	Percentage
1	R1	A top senior executive in charge of overseeing research data governance policy and funding, and strategizing activities and programs.	17	65.4%	9	34.6%
2	R4	Responsible for ensuring the application of data governance principles by promoting them, facilitating best practices, and monitoring instances of data misconduct.	20	76.9%	6	23.1%
3	R5	Accountable for research data and materials including, but not limited to the collection, development, maintenance, distribution, preservation, reuse, and security of research data and materials.	21	80.8%	5	19.2%

6.5 Nomenclature for Research Data Governance Roles

In addition to the questions about the descriptions of RDG roles, the panelists responded to inquiries regarding the appropriate nomenclature of these roles. Seven questions related to the RDG nomenclatures were initially asked in Round III and later reiterated in Round IV to determine the most ideal nomenclature for each RDG role to be integrated into the RDG framework.

First, the focus was on identifying the best nomenclature for the Executive Sponsor role. In Round III, “Data Governance Executive” received 60.6 percent (n=20) responses, and this consensus strengthened to 84.6 percent (n=22) in Round IV, indicating a clear preference for this title. The shift in preference suggests a refined consensus among the panelists. On the contrary, “Executive Sponsor” witnessed a significant decrease from 15.2 percent (n=5) agreement in Round III to 3.8 percent (n=1) in Round IV, which contradicts the viewpoint expressed by P30, who prefers “Executive Sponsor” above other options. Other nomenclatures, including “Data Compliance Advocate”, “Data Governance Champion”, and “Data Oversight Director”, were chosen by 6.1 percent (n=2) of the panelists in Round III. However, preferences changed in Round IV, with a

decrease to 3.8 percent (n=1) for the last two options and 0 percent for "Data Compliance Advocate".

Similarly, the suggested addition by panelists, "Deputy Vice Chancellor (Research)", experienced a decrease from 3 percent (n=1) in Round III to 0 percent in Round IV. Meanwhile, the title "Executive Board / Research Board / Top University Management Committee" maintained a consistent one agreement in both rounds. In conclusion, the majority consensus points towards "Data Governance Executive" as the most fitting nomenclature to represent the Executive Sponsor role within the context of Malaysia RPOs. Table 6.25 provides the detailed breakdown of panelist responses.

Table 6.25: Panelist responses on appropriate nomenclature for the Executive Sponsor role in Round III and IV

Nomenclature for Executive Sponsor Role	Round III		Round IV	
	Count	Frequency	Count	Frequency
Data Governance Executive	20	60.6%	22	84.6%
Executive Sponsor	5	15.2%	1	3.8%
Data Compliance Advocate	2	6.1%	0	0.0%
Data Governance Champion	2	6.1%	1	3.8%
Data Oversight Director	2	6.1%	1	3.8%
[*]Deputy Vice Chancellor (Research)	1	3.0%	0	0.0%
[*]Executive Board / Research Board / Top University Management Committee	1	3.0%	1	3.8%
TOTAL	**33**	**[a]100%**	**26**	**[a]100%**
Comment, feedback or rewording for clarification:	1	3.0%	0	0.0%

From Round III

1. P30-In our situation, Data Governance Champion can be considered because in our University's Roadmap, all strategic initiatives led by a Champion. But I think Executive sponsor is more suitable.

Notes.
[*]Nomenclature identified by panelist in Round I was included as new option for group reassessment in Round II
[a]Percentages were rounded to the nearest whole number

Secondly, the focus shifted to identifying the best nomenclature for the Data Governance Leader role. The data comparing Rounds III and IV (as in Table 6.26) provides insights into the shifting preferences for the nomenclature of the Data Governance Leader role. In Round III, "Data Governance Director" held 45.5 percent

(n=15) responses, and these responses increased to 76.9 percent (n=20) in Round IV, suggesting a notable shift in favour of this title. Conversely, “Data Governance Leader” experienced a decrease from 21.2 percent (n=7) in Round III to 7.7 percent (n=2) in Round IV.

The terms “Chief Data Officer”, “Chief Information Officer”, and “Chief Digital Officer” maintained varying levels of agreement across both rounds, with “Chief Data Officer” experienced a slight decrease from 18.2 percent (n=6) in Round III to 11.5 percent (n=3) in Round IV. The term “Chief Information Officer” declined from 12.1 percent (n=4) to 3.8 percent (n=1), while “Chief Digital Officer” received 3 percent agreement (n=1) in Round III and received no agreement in Round IV.

The comments from both rounds add valuable context. In Round III, P30 indicated that, in his institution, “Chief Information Officer” is considered the most appropriate. In Round IV, comment by P32 suggests that “Data Governance Director” is perceived as providing a better understanding compared to “Chief Data Officer”. Additionally, there is a recognition made by P30 that the current nomenclature might differ across organizations, emphasizing the importance of considering contextual variations. However, in general, the prevailing consensus indicates that “Data Governance Director” is the most appropriate term to denote the role of Data Governance Leader within the framework of Malaysia RPOs.

Table 6.26: Panelist responses on appropriate nomenclature for the Data Governance Leader role in Round III and IV

Nomenclature for Data Governance Leader Role	Round III		Round IV	
	Count	Frequency	Count	Frequency
Data Governance Director	15	45.5%	20	76.9%
Data Governance Leader	7	21.2%	2	7.7%
Chief Data Officer	6	18.2%	3	11.5%
Chief Information Officer	4	12.1%	1	3.8%
Chief Digital Officer	1	3.0%	0	0.0%
TOTAL	**33**	**[a]100%**	**26**	**[a]100%**
Comment, feedback or rewording for clarification:	1	3.0%	2	7.7%

From Round III

1. P30-In our institution, CIO is the most appropriate, I guess.

From Round IV

1. P32-Data Governance Director will have a better picture compared to Chief Data Officer.
2. P30-Maybe the current nomenclature used in certain organizations differs.

Notes.
[a]Percentages were rounded to the nearest whole number

Moving on to the third role, the Research Data Governance Committee, consensus among panelists in Round III pointed towards “Research Data Governance Committee” as the most suitable nomenclature, with the highest agreement at 45.5 percent (n=15). This agreement increased to 57.7 percent, maintaining the same number of panelists (n=15) in Round IV. In contrast, “Research Data Governance Steering Committee” experienced a decrease from 27.3 percent (n=9) in Round III to 19.2 percent (n=5) in Round IV. The terms “Data Governance and Research Committee”, “Research Data Stewardship Committee”, and “Data Governance and Research Ethics Committee” maintained varying levels of agreement across both rounds, with relatively minor changes.

Thus, the consensus leans towards the title “Research Data Governance Committee”, as preferred by the panelists. Table 6.27 illustrates a preference for the term “Research Data Governance Committee” role across both rounds.

Table 6.27: Panelist responses on appropriate nomenclature for the Research Data Governance Committee role in Round III and IV

Nomenclature for Research Data Governance Committee	Round III		Round IV	
	Count	Frequency	Count	Frequency
Research Data Governance Committee	15	45.5%	15	57.7%
Research Data Governance Steering Committee	9	27.3%	5	19.2%
Data Governance and Research Committee	4	12.1%	3	11.5%
Research Data Stewardship Committee	4	12.1%	2	7.7%
Data Governance and Research Ethics Committee	1	3.0%	1	3.8%
TOTAL	**33**	**[a]100%**	**26**	**[a]100%**

Notes.
[a]Percentages were rounded to the nearest whole number

The following Table 6.28 illustrates the results obtained from Rounds III and IV regarding preferences for the nomenclature of the next role related to Research Data Owner. In Round III, “Research Data Owner” was strongly favoured, achieving 75.8 percent agreement (n=25). The percentage increased to 88.5 percent in Round IV, even with a slightly lower number of panelists agreeing (n=23). In contrast, the term “Research Data Custodian” obtained an agreement of 4 (12.1%) and 3 (11.5%) panelists in Rounds III and IV, respectively. However, both “Research Data Steward” and “Research Data Guardian” witnessed a decrease from 9.1 percent agreement (n=3) in Round III to 0 percent in Round IV and from 3 percent in Round III (n=1) to 0 percent in Round IV, respectively. Meanwhile, “Research Data Caretaker” had no agreement in both rounds. The absence of agreement for “Research Data Steward”, “Research Data Guardian”, and “Research Data Caretaker” in Round IV suggests a refined preference for the term “Research Data Owner” as the primary nomenclature for this role within the context of RDG in Malaysia RPOs.

Table 6.28: Panelist responses on appropriate nomenclature for the Research Data Owner role in Round III and IV

Nomenclature for Research data Owner	Round III		Round IV	
	Count	Frequency	Count	Frequency
Research Data Owner	25	75.8%	23	88.5%
Research Data Custodian	4	12.1%	3	11.5%
Research Data Steward	3	9.1%	0	0.0%
Research Data Guardian	1	3.0%	0	0.0%
Research Data Caretaker	0	0.0%	0	0.0%
TOTAL	**33**	**[a]100%**	**26**	**[a]100%**

Notes.
[a]Percentages were rounded to the nearest whole number

Table 6.29 presents results from Rounds III and IV, providing insights into the favoured nomenclature for the role associated with Research Data Governor. In Round III, "Research Data Steward" received the highest favour, with 36.4 percent (n=12) agreement, and this consensus increased to 53.8 percent (n=14) in Round IV. In contrast, the term "Research Data Control Manager" saw a slight decrease from 33.3 percent (n=11) agreement in Round III to 34.6 percent (n=9) in Round IV. Similarly, the title "Research Data Management Custodian" also experienced a decline from 12.1 percent agreement (n=4) in Round III to 7.7 percent (n=2) in Round IV. Other terms, such as "Research Data Governor", "Research Data Integrity Supervisor", and "Research Data Oversight Officer", all obtained 6.1 percent agreement (n=2) each in Round III. However, in Round IV, these terms showed varying levels of agreement, all showing a decrease compared to Round III. Thus, the data suggests that the panelists' preference for the Research Data Governor role is "Research Data Steward".

Table 6.29: Panelist responses on appropriate nomenclature for the Research Data Governor role in Round III and IV

Nomenclature for Research Data Governor	Round III		Round IV	
	Count	Frequency	Count	Frequency
Research Data Steward	12	36.4%	14	53.8%
Research Data Control Manager	11	33.3%	9	34.6%
Research Data Management Custodian	4	12.1%	2	7.7%
Research Data Governor	2	6.1%	1	3.8%
Research Data Integrity Supervisor	2	6.1%	0	0.0%
Research Data Oversight Officer	2	6.1%	0	0.0%
TOTAL	**33**	**[a]100%**	**26**	**[a]100%**

Notes.
[a]Percentages were rounded to the nearest whole number

The determination of the best nomenclature for the Research Data Steward role reveals interesting trends across Rounds III and IV (Table 6.30). Initially, “Business Research Data Steward” and “Research Data Curator” were equally favoured, each receiving 27.3 percent (n=9) agreement. In Round IV, however, “Research Data Curator” increased to 34.6 percent with the same number of agreement (n=9), while “Business Research Data Steward” decreased to 26.9 percent agreement (n=7). On the other hand, the term “Research Data Governance Manager” witnessed a decrease from 15.2 percent (n=5) in Round III to 11.5 percent (n=3) in Round IV. “Research Data Compliance Coordinator” and “Research Data Custodian” experienced notable decreases from 9.1 percent (n=3) in Round III to 0 percent in Round IV, and from 9.1 percent (n=3) to 7.7 percent (n=2), respectively. The absence of agreement for “Research Data Compliance Coordinator” in Round IV suggests a refined preference for other role nomenclatures. The term “Research Data Steward” showed an increment from 9.1 percent (n=3) to 19.2 percent (n=5). Furthermore, one panelist introduced the term “Data Steward” in Round III, but it received no agreement in Round IV. In summary, while there was a slight variance in preferences between the first two nomenclatures, the Research Data Steward role is denoted by the term "Research Data Curator" due to its greater preference among the panelists.

Table 6.30: Panelist responses on appropriate nomenclature for the Research Data Steward role in Round III and IV

Nomenclature for Research Data Steward Role	Round III		Round IV	
	Count	Frequency	Count	Frequency
Business Research Data Steward	9	27.3%	7	26.9%
Research Data Curator	9	27.3%	9	34.6%
Research Data Governance Manager	5	15.2%	3	11.5%
Research Data Compliance Coordinator	3	9.1%	0	0.0%
Research Data Custodian	3	9.1%	2	7.7%
Research Data Steward	3	9.1%	5	19.2%
[*]Data Steward	1	3.0%	0	0.0%
TOTAL	**33**	**[a]100%**	**26**	**[a]100%**

Notes.
[*]Nomenclature identified by panelist in Round I was included as new option for group reassessment in Round II
[a]Percentages were rounded to the nearest whole number

The final nomenclature is related to Research Data Custodian role. The findings from both Rounds III and IV (as shown in Table 6.31) reveal the preferences for the nomenclature. In Round III, there was a strong preference for the term “Research Data Custodian”, with 69.7 percent agreement (n=23), and this consensus increased to 92.3 percent (n=24) in Round IV. In contrast, the term “Technical Research Data Steward” experienced a decrease from 15.2 percent (n=5) in Round III to 3.8 percent (n=1) in Round IV. Other terms such as “Research Data Custodial Officer”, “Research Data Caretaker”, and “Research Data Guardian” maintained varying levels of agreement across both rounds. Specifically, “Research Data Custodial Officer” decreased from 12.1 percent (n=4) in Round III to 3.8 percent (n=1) in Round IV. “Research Data Caretaker” received 3 percent (n=1) agreement in Round III and received no responses in Round IV. Meanwhile, “Research Data Guardian” did not receive any agreement in both Rounds. The strong and increasing agreement for the term “Research Data Custodian” suggests a clear preference for this nomenclature in describing role related to the custodianship of research data within Malaysia RPOs.

Table 6.31: Panelist responses on appropriate nomenclature for the Research Data Custodian role in Round III and IV

Nomenclature for Research Data Custodian Role	Round III		Round IV	
	Count	Frequency	Count	Frequency
Research Data Custodian	23	69.7%	24	92.3%
Technical Research Data Steward	5	15.2%	1	3.8%
Research Data Custodial Officer	4	12.1%	1	3.8%
Research Data Caretaker	1	3.0%	0	0.0%
Research Data Guardian	0	0.0%	0	0.0%
TOTAL	**33**	**[a]100%**	**26**	**[a]100%**

Notes.
[a]Percentages were rounded to the nearest whole number

6.6 Summary of Chapter Six

The main objectives of this chapter are to identify the stakeholders who should assume the identified governance roles and to determine an individual's governance role along with their structural position within an RPO. Through a comprehensive analysis, valuable insights were acquired that contribute to a better understanding of the governance structural dimension of RDG. The study refers to the Data Governance Instantiation of Agile Governance Model (AGM) presented by Korhonen et al. (2013) for structural guidance. Figure 6.1 below presents a comprehensive view of the RDG structural dimension, drawing on the findings discussed in this chapter. It illustrates the stakeholders and demonstrates how governance roles interconnect and align within the governance structure. In Chapter Seven, the focus shifts to the elaboration of the proposed RDG framework, which emerges from the synthesis of insights gathered across all phases of the Delphi study. This chapter delves into the breakdown of each framework component, providing a comprehensive understanding of its structure and function.

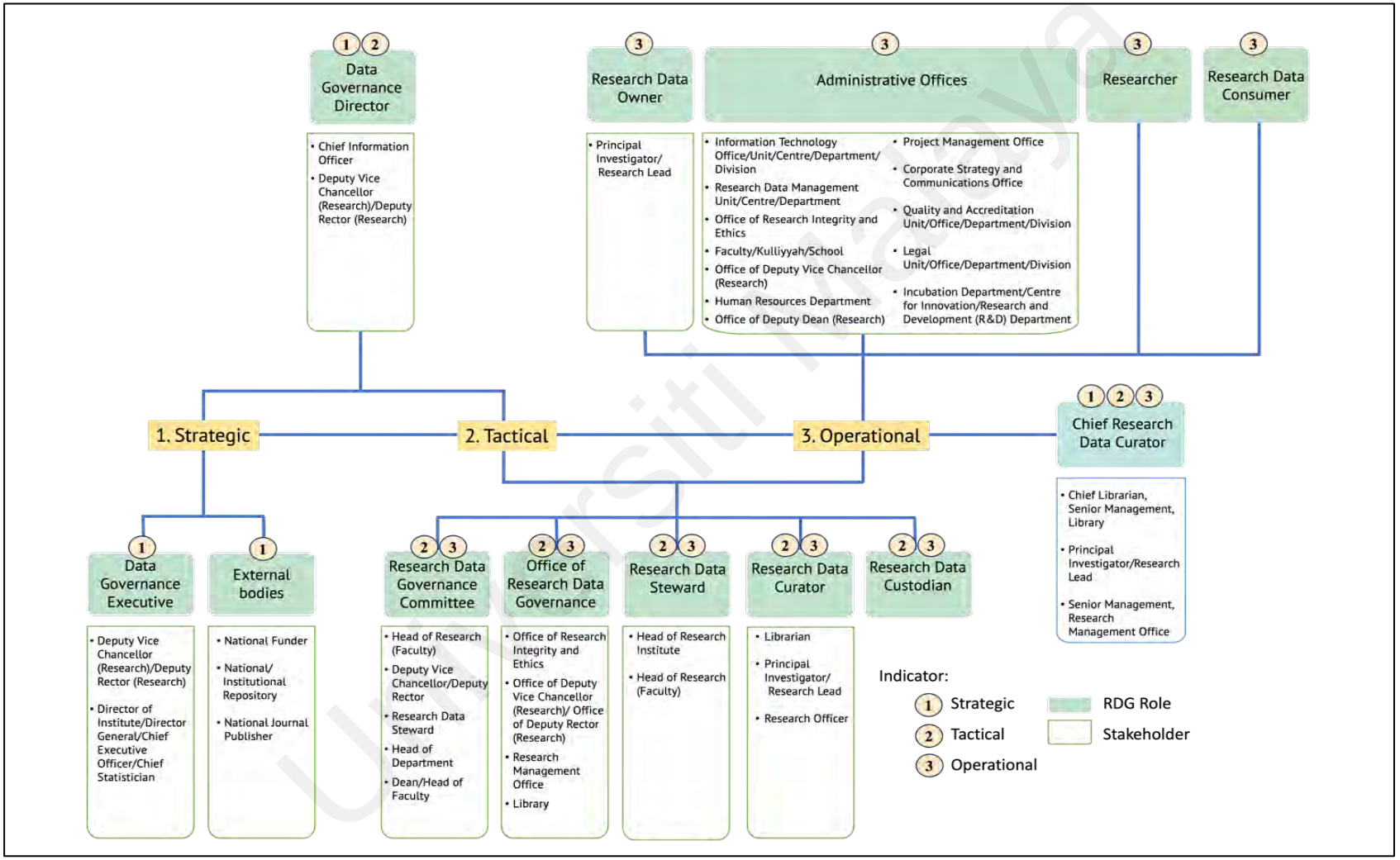


Figure 6.1: A diagram illustrating the connection between RDG role and its structural position

CHAPTER 7: RESEARCH DATA GOVERNANCE FRAMEWORK

"This is the Book about which there is no doubt, a guidance for those conscious of Allah".
(Al-Baqarah, 2:2)

7.1 Introduction

This chapter addresses RO3: *To develop a research data governance framework for research performing organizations.* Given the government's growing focus on driving innovation and cultivating a knowledge-based economy, as well as initiatives aimed at promoting open science in Malaysia, there is a pressing need for RDM among RPOs. Considering the dynamic nature of the research landscape, it has become essential to develop an RDG framework. Having proper management and governance in place for research data is vital in maintaining its quality, reliability, confidentiality, and availability. This framework includes crucial structural elements, such as the roles (people) involved, including stakeholders and their positions within the organization. Additionally, it addresses functional aspects related to processes, covering various tasks, areas, and decision domains.

7.2 Creation of Research Data Governance Framework

The main goal of the proposed framework is to align RDG with existing RPO's governance structures, rather than creating a separate, parallel system. The framework can be integrated with existing research governance and IT governance, making it easier to implement without disrupting existing processes. It encourages RPOs to build upon what is already in place and refine their governance practices for research data in a way that complements their broader organizational goals.

To comprehensively understand this RDG framework, the study introduces a two-layer dimensional framework. The first layer, the structural dimension, addresses the question of "WHO" by detailing RDG roles and governance structure. The second layer, the functional dimension, explores various aspects of data governance: "WHAT?" refers to

data governance actions, “WHICH?” identifies the RDG areas, “HOW?” describes the RDG decision domains, and “WHY?” explains the vision, mission, and drivers behind the framework. Figure 7.1 illustrates the relationships among these various elements of the RDG framework in the form of a building structure.

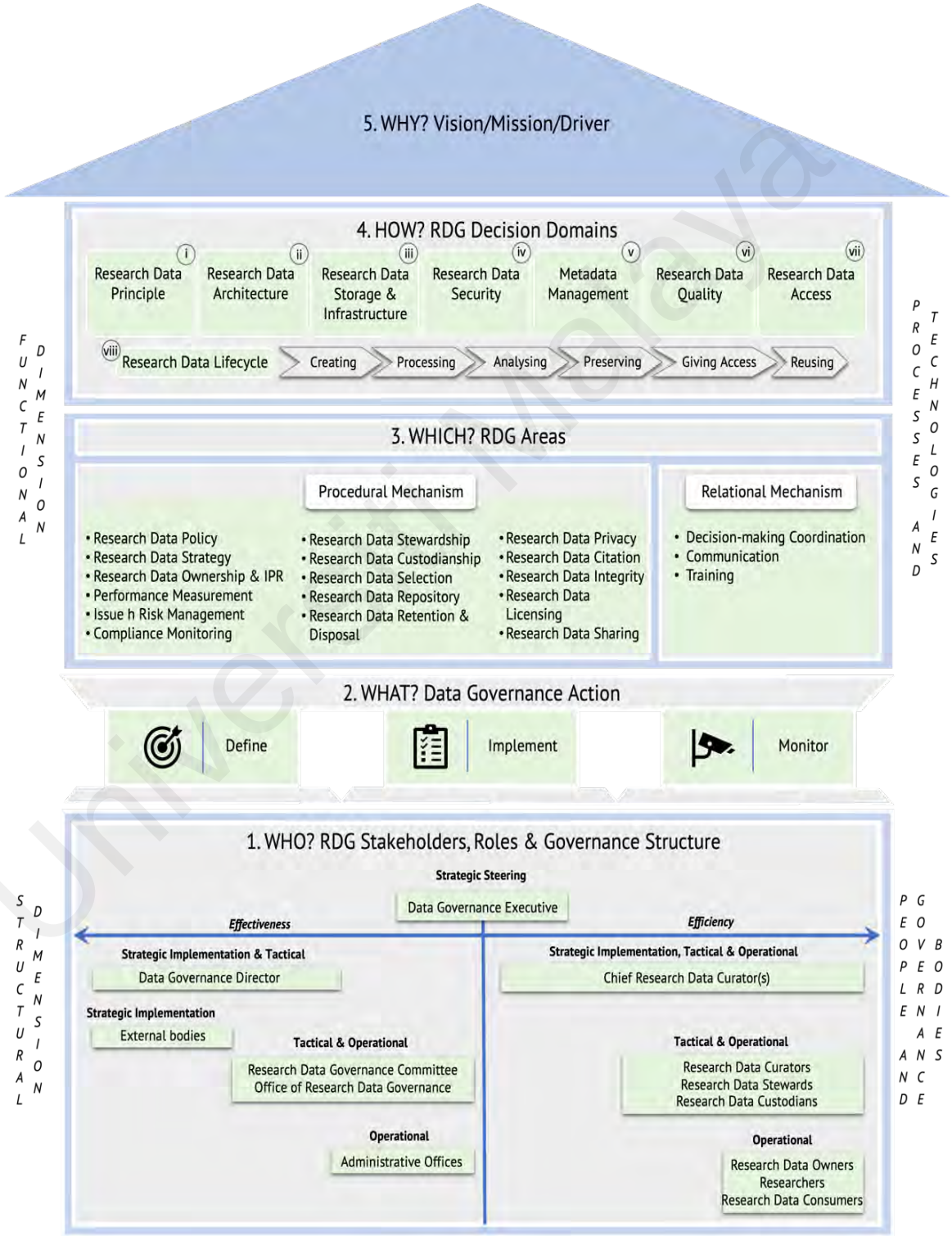


Figure 7.1: Research Data Governance Framework of this Study

The RDG framework outlines roles and its governance structure, actions (responsibilities), as well as areas and decision domains (processes) for effective RDG practices and implementation. The framework is best compared to a building and its structure, serving as a metaphor to illustrate RDG practices within an RPO. Ensuring a safe living environment in a building involves several stages, each requiring careful planning, implementation, coordination, and the skills of proficient individuals. For a building, it involves constructing the basic structure, including walls, floors, and the roof, with each floor representing distinct facets of RDG structural and functional dimensions. In addition to the two floors, the building is illustrated by three primary staircases and a roof, each corresponding to the 4W1H (Who, What, Which, Why and How) questions.

7.2.1 Structural Dimensions

This section details the RDG roles, and their governance structure as defined in the foundation of the framework.

7.2.1.1 RDG roles and governance structure

The first floor of the RDG framework, representing “WHO”, embeds foundational values where individuals contribute to the efficacy of RDG implementation. Assigning individuals or data practitioners and governance bodies for RDG is crucial in forming an effective framework that preserves the credibility, confidentiality, and accessibility of research data. Clearly defining roles and responsibilities is vital for accountability and promoting seamless collaboration (Alhassan et al., 2018; Alhassan et al., 2019a; Omar & Almaghthawi, 2020). The participation of individuals and governance bodies is essential to guarantee the smooth execution of the program, particularly during the initiation of RDG within RPOs. Processes cannot move forward, and technologies hold little significance without the identification of committed individuals, departments, and offices.

Engaging key stakeholders, such as researchers, administrators, IT personnel, and external collaborators, is pivotal for a thorough implementation of RDG within the organization. It is crucial for top leaders to grasp the distinctive needs and perspectives of each stakeholder group to customize the governance framework appropriately. A well-defined structure establishes effective communication channels, reporting mechanisms, and decision-making processes (Abraham et al., 2019). It fosters collaboration among stakeholders and ensures a shared understanding of everyone's role in governing research data.

The individuals or data practitioners and governance bodies are organized into four main levels of RDG governance: (a) strategic steering, (b) strategic implementation, (c) tactical, and (d) operational, based on effectiveness and efficiency aspects (Korhonen et al., 2013). Individuals at the strategic level are responsible for designating and executing strategic plans and long-term goals for research data. Those at the tactical level support these strategic plans by translating them into specific plans relevant to lower-level departments, units, or faculties. Entities at the operational level focus on the specific procedures and processes for research data that occur within the lowest levels of the organization (Education Portal, 2013).

The effectiveness aspect, concentrating on achieving goals, involves activities like design and planning, fostering adaptability, and innovation. Specific individuals and governance bodies associated with the effectiveness aspect play a crucial role in ensuring the planning and adaptability of RDG within the organization. These entities encompass: (a) Data Governance Director, (b) External Bodies, (c) Research Data Governance Committee, (d) Office of Research Data Governance, and (e) Administrative Offices.

In contrast, the efficiency aspect, optimizing resource use, includes activities like development and operational work, ensuring predictability and accountability. On the efficiency aspect, entities responsible for ensuring accountability, development, and operational work include the Chief Research Data Curator(s), Research Data Curators, Research Data Stewards, Research Data Custodians, Research Data Owners, Researchers, and Research Data Consumers.

It is important to note that the "WHO" section in Figure 7.1 does not represent an organizational chart. Instead, this section explains how various bodies collaborate to carry out the RDG function. The entities are positioned differently within the structure, with some holding responsibilities across more than one level. The specific responsibilities and contributions of each entity are discussed below:

a) ***Data Governance Executive (Strategic Steering: Effectiveness & Efficiency)*** - At the strategic steering level, the top tier of "WHO", the Data Governance Executive, who oversees both effectiveness and efficiency at the strategic steering level, leads the oversight of critical aspects, such as RDG policy and funding. This role involves streamlining strategies, setting the overall direction and establishing long-term aspirations, ensuring both effectiveness and efficiency in data governance. This forward-looking phase involves leaders making high level determinations to safeguard and make use of the organization's assets, such as research data, in alignment with its vision and mission, effectively navigating towards successful RDM. In Malaysia academic settings and higher education institutions, the role of the Data Governance Executive is ideally fulfilled by the Deputy Vice Chancellor (Research) or Deputy Rector (Research). However, in cases where this position is not in place, such as in research organizations or institutions

without a designated role, the Director of the Institute, Director General, Chief Executive Officer, or another top executive from the Directorship and C-Levels is the most fitting choice. These individuals hold high credibility within the organization and have a deep understanding of internal challenges. Moreover, they possess the necessary authority to engage in strategic management and make lasting decisions for the entire organization, including those related to RDG.

b) ***Data Governance Director (Strategic Implementation & Tactical: Effectiveness)*** - During this strategic implementation phase, effective management plays a pivotal role in ensuring the seamless execution of the organization's vision towards RDG. The implementation of strategic initiatives involves translating overarching objectives into practical plans and actions at the organizational level. Data Governance Director is accountable to coordinate the overall management of the organization's data governance. Ideally, this role is filled by either the Chief Information Officer or the Deputy Vice Chancellor (Research)/Deputy Rector (Research), depending on post availability, organization size, and its specific needs. Additionally, the Data Governance Director actively participates in meetings with the Research Data Governance Committee, leading the committee in the overall management of the organization's data governance. Therefore, the role of the Data Governance Director extends beyond strategic implementation to also include involvement at the tactical level, ensuring effective communication for information dissemination and monitoring.

c) ***External Bodies (Strategic Implementation: Effectiveness)*** - External bodies, including funders, national or institutional repositories, and national journal publishers, also play a strategic role in implementation, particularly in the realm of effectiveness perspective. While these entities originate from the external environment, they play a crucial role in shaping policies related to RDG, establishing standards that researchers and institutions must follow to uphold ethical and legal practices when handling research data. Additionally, they contribute to the development of standards and interoperability frameworks for research data, ensuring seamless sharing and utilization across various platforms and systems.

d) ***Research Data Governance Committee (Tactical & Operational: Effectiveness)*** - The Research Data Governance Committee operates at the tactical and operational levels under the effectiveness aspect. Among its primary responsibilities is to articulate strategic research data requirements that align with organizational and management needs. Comprising representatives from major strategic stakeholders across the organization, the ideal composition includes Deputy Vice Chancellors/Deputy Rectors, Heads of Research (Faculty), Research Data Stewards, Heads of Departments, and Deans/Heads of Faculties. The coordination of the Committee's work is done by the Data Governance Director.

e) ***Office of Research Data Governance (Tactical & Operational: Effectiveness)*** -Another entity operating at the tactical and operational levels under the effectiveness aspect, in addition to the Research Data Governance Committee, is the Office of Research Data Governance. The arrangement may vary across organizations; some may delegate the responsibilities to office such as the Office of Research Integrity and Ethics, Office of Deputy

Vice Chancellor (Research)/Office of Deputy Rector (Research), Research Management Office, or Library, depending on the availability of the respective office, suitability, and the organization's needs. The primary responsibility of this Office is to ensure the application of data governance principles by promoting them, facilitating best practices, and monitoring instances of data misconduct. It monitors compliance with RDG policies and supporting processes, provides advice and guidance, and conducts reviews related to RDM, particularly concerning research integrity and ethics. Additionally, it plays a role in approving and implementing processes that support RDG.

f) ***Administrative Offices (Operational: Effectiveness)*** - To enable the implementation of RDG, administrative offices, including centralized or decentralized departments and/or offices, collaborate to support RDG practices and implementation. These administrative offices are situated at the operational level from an effectiveness perspective. When integrating the roles of administrative offices to support the Office of RDG and complement one another, it is ideally structured so that the Office of RDG manages tactical operations, while other administrative offices handle operational matters. Their primary responsibility is to support researchers, schools, departments, and/or offices in implementing RDG and associated procedures within their respective areas. The administrative offices involved encompass a diverse range, including (but not limited to):

    (i) Office of Research Integrity and Ethics

    (ii) Office of Deputy Vice Chancellor (Research)

    (iii) Research Management Office

    (iv) Library

(v) Information Technology Unit/Centre/Department/Division

(vi) Research Data Management Unit/Centre/Department

(vii) Faculty/Kulliyyah/School

(viii) Human Resources Department

(ix) Office of Deputy Dean (Research)

(x) Project Management Office

(xi) Corporate Strategy and Communications Office

(xii) Quality and Accreditation Unit/Office/Department/Division

(xiii) Legal Unit/Office/Department/Division

(xiv) Incubation Department/Centre for Innovation/Research and Development (R&D) Department

g) ***Chief Research Data Curator (Strategic Implementation, Tactical & Operational: Efficiency)*** - Chief Research Data Curator is tasked with enhancing the efficiency aspect across strategic implementation, tactical, and operational levels. This multifaceted role includes ensuring the quality and integrity of data, overseeing the implementation and enforcement of RDM within the organization, department, and research projects. The Chief Research Data Curator is also responsible for making tactical decisions, utilizing their expertise to make informed choices that contribute to the success of RDG practices. Additionally, they oversee the specific operational activities and processes, supervising and monitoring the correct execution and updates of RDM regulations and procedures. This position may be assigned to an individual or a team, with suitable candidates potentially including heads of the library, principal investigators or research leads, and senior officials from the Research Management Office.

h) ***Research Data Curators (Tactical & Operational: Efficiency)*** - Research Data Curators are accountable in establishing clear data management responsibilities within a smaller group. They play a key role in organizing resources and providing support for RDM initiatives within the group. Additionally, they play a pivotal role in managing custodial responsibilities when researchers depart, maintaining the integrity and accessibility of valuable datasets. Additionally, Research Data Curators are responsible in providing training and ongoing support for effective data management practices, contributing to the development of a research community that prioritizes responsible and sustainable data stewardship. Research Data Curators, comprising individuals like librarians, principal investigators, and research officers. Each research data set may have one or more assigned curators.

i) ***Research Data Stewards (Tactical & Operational: Efficiency)*** - Research Data Stewards are accountable for implementing RDG policies within their colleges, schools, or departments. Among their various responsibilities, Research Data Stewards actively model responsible data management behaviour. Additionally, they take on the responsibility of ensuring effective communication regarding data management practices. In addition, these Research Data Stewards play a pivotal role in guaranteeing the integration of comprehensive data management plans into grant applications, thereby aligning research initiatives with data governance standards, and contributing to the overall integrity and reliability of the research process. This group may include Heads of Research Institute and/or Heads of Research (Faculty).

j) ***Research Data Custodians (Tactical & Operational: Efficiency)*** - Research Data Custodians, often deal with technical aspects, are established to ensure that important datasets are developed, maintained, and accessible within defined specifications. Designating a person or body to oversee these aspects of data management helps prevent the compromise of datasets.

k) ***Research Data Owner (Operational: Efficiency)*** - At the core of organizational efficiency lies operational governance, which encompasses the specific tasks and decision-making processes of front-line individuals. This critical aspect relies on frontline employees to execute tactical plans and uphold the RDG strategic objectives, making their contribution paramount to the organization's overall success. The term 'Research Data Owner' in this context does not imply ownership of the research data by the organization; instead, it represents individuals accountable for research data and materials, covering tasks such as collection, development, maintenance, distribution, preservation, reuse, and security. Ideally, Research Data Owners are the Principal Investigators or Research Leads of research projects.

l) ***Researchers (Operational: Efficiency)*** - Researchers encompass individuals such as staff members, collaborators, adjuncts, academic title holders, and students engaged in conducting research. They are the creators responsible for conducting research and producing research data. Their operational tasks include publishing research data to disciplinary, institutional, or other established repositories to maximize research value, managing research data throughout the research data lifecycle in accordance with organizational policies, guidelines, standards, and applicable funder, legislative, and ethical requirements. Additionally, researchers protect confidential, personal, and sensitive research data, following legal and ethical requirements related to the

research conducted. They also regularly back up research data and records according to best practices in their respective research fields.

m) ***Research Data Consumers (Operational: Efficiency)*** - Research Data Consumers acknowledge the sources of research data and abide by the terms and conditions under which the research data are accessed.

7.2.2 Functional Dimensions

This section explores the functional dimensions that focus on data governance actions, areas, decision domains, and vision/mission/driver that an RPO may adopt.

7.2.2.1 Data Governance Actions

Beyond the foundation, three primary staircases connect the first and second floors, denoting “WHAT” elements. They signify the imperative for individuals within the RPO to comprehend their roles and responsibilities for effective RDG implementation. This entails three governance actions: Define, Implement, and Monitor, which establish connections to the subsequent RDG elements.

The RDG roles, as discussed earlier, have the authority over the management of research data. It is crucial for them to articulate their responsibilities in effectively dealing with research data. Each role entails specific duties aimed at ensuring the proper handling of research data according to established policies, standards, guidelines, rules, and regulations. Figure 7.2 detailed out the main responsibilities given to each RDG role, providing a detailed breakdown of the essential tasks involved in each role.

Organization

1. Ensures research data availability for the research community.
2. Ensures support for grant holders in policy compliance.
3. Manages an institutional metadata catalogue, especially for publicly funded research.
4. Safeguards researchers' rights, ensuring access to their data.
5. Establishes a governance committee overseeing data governance implementation.
6. Maintains institutional-level research data governance policies.
7. Enables planning and execution of good research data management practices.
8. Collaborates with funders, policymakers, and stakeholders to align governance with sector requirements.
9. Ensures ethical use of research elements like animals, human subjects, and materials.
10. Facilitates investigations into scientific misconduct or conflict of interest.
11. Takes custody of data when necessary for appropriate access.
12. Owns all research data and associated intellectual property.
13. Develops best practices for data sharing in different fields
14. Provides researchers with training on effective research data management.
15. Ensures long-term stewardship based on institutional/national data infrastructure.
16. Offers facilities, advisory services, and resources for secure data storage and management.
17. Ensures storage facilities comply with legal and regulatory requirements.
18. Acknowledges contributions of researchers who generate, preserve, and share key datasets.

Data Governance Executive

1. Provides necessary facilities and support for research data management.
2. Oversees the implementation of research data governance as a framework for good research data management.

Data Governance Director

1. Ensures regular review and maintenance of research data governance policies in line with sector requirements.
2. Develops and maintains a central repository for governance policies, guiding principles, and decisions.
3. Maintains the Research Data Governance Committee agenda and convenes meetings.
4. Communicates outcomes of the Research Data Governance Committee.
5. Serves as an expert on research data governance and recommends solutions.

Research Data Governance Committee

1. Provides a common vocabulary for primary research data entities and types, essential to the organization.
2. Oversees initiatives to enhance effective and efficient research data utilization.
3. Supports research data management for institutional initiatives.
4. Cultivates a research data management culture that adds value to the institution.

Office of Research Data Governance

1. Signs off policies, supports cultural and behavioral changes, and allocates resources to research data governance activities.
2. Provides advice, guidance, and reviews on research data management, emphasizing research integrity and ethics.

External Bodies

National Funder

1. Defines research data governance principles aligned with funding principles.
2. Reviews implementation of research data management plans.
3. Specifies retention periods for significant research data.

National Repository

1. Specifies required contextual information and metadata for deposited data.

Administrative Offices

Library

1. Coordinates a network of data curators.
2. Manages research data metadata records and publishes them on a public catalog.
3. Maintains a research data repository for registering data and other outputs.
4. Defines the institution's support for research data management.
5. Defines high-quality infrastructure for data collection, storage, and sharing.
6. Reviews research data management policies with governance bodies.
7. Offers training on metadata standards, data management plans, and legal/ethical issues.
8. Advises on research data costs throughout the project lifecycle.

Information Technology Office

1. Provides a standardized, robust infrastructure for effective research data management.
2. Offers technical support for data management, including storage, backup, and archiving.
3. Provides secure access management following ICT security guidelines.

Research Management Office

1. Advocates organizational awareness of research data management issues.
2. Ensures governance policies are updated based on the latest funder requirements and national directives.
3. Organizes training events on research data management.
4. Provides advice, guidance, and assistance to researchers in preparing data management plans.
5. Provides templates for incoming and outgoing research data agreements.

Research Data Steward

1. Facilitates resources and support for research data management.
2. Approves storage, disposal, and publication of research data.
3. Model responsible data management behavior.
4. Ensures effective communication on data management.
5. Ensures the requirement of data management plans for grant applications.
6. Ensures compliance by principal investigators.
7. Reports IT security incidents and data breaches.
8. Supervises adherence to regulations and procedures.
9. Approves attestations for researcher awareness.
10. Promotes a culture of data awareness and training.
11. Makes decisions in ambiguous data guidelines.
12. Sets expectations for data classification and retention.
13. Sponsors, secures, and/or influences resources for research data management.

Research Data Curator

1. Securely store and protect active research data to meet compliance requirements.
2. Assists in external audits and provides data access.
3. Ensures appropriate classification and management of research data based on their sensitivity.
4. Ensures the security of confidential data in accordance with privacy laws.
5. Ensures the archival of data for long-term preservation.
6. Ensures data licensing for reuse and dissemination.
7. Ensures management of custodial responsibilities for departing researchers.
8. Provides training and support for data management.
9. Ensures data retention as per organizational guidelines.
10. Reviews and recommends improvements to data management plans.
11. Develops processes for data selection, storage, and protection.
12. Monitors group's data management regulations.
13. Ensures the availability of data management plans for research projects.
14. Ensures the inclusion of data management costs in research proposals.
15. Ensures appropriate data access to project team members.
16. Establishes clear data management responsibilities in the group.
17. Arranges resources and support for research data management in the group.

Researcher

1. Selects research data for long-term preservation based on verification/replication and reuse needs.
2. Retains research data securely for a period determined by the organization.
3. Ensures written agreements cover ownership, sharing, storage, and disposal of research data.
4. Writes a comprehensive research data management plan.
5. Protects intellectual property according to funder or contractual obligations.
6. Provides sufficient metadata for discoverability and reusability of research data.
7. Publishes data to established repositories, maximizing research value.
8. Guarantees integrity, quality, security, and persistent availability of research data.
9. Disposes of data and materials securely.
10. Includes a data access statement in publications.
11. Reports security breaches affecting research data.
12. Hands over data and materials after projects or leaving the institution.
13. Manages data throughout the lifecycle in line with policies, guidelines, and requirements.
14. Arranges safe and secure storage of research materials.
15. Understands and addresses risks of third-party storage solutions.
16. Protects confidential and sensitive data in line with legal and ethical requirements.
17. Participates in training on research data management and contractual obligations.
18. Upholds open-source file formats and types recommended for preservation.
19. Regularly backs up research data in accordance with best practices.
20. Develops operating procedures to comply with research data governance.
21. Works with the information security team to ensure system controls.
22. Identifies and addresses research data integrity and quality issues.
23. Periodically reviews data access and usage agreements.

Research Data Consumer

1. Acknowledges data sources and adhere to access terms and conditions.
2. Complies with controls specified in the data management plan.

Figure 7.2: RDG roles and responsibilities

7.2.2.2 RDG Areas

RDG is a multifaceted approach that integrates various processes and technologies to guarantee the effective management, quality enhancement, security enforcement, and enhanced usability of research data within an RPO. It encompasses various areas and decision domains to ensure that research data is effectively managed, protected, used, and re-used for the advancement of research and innovation. The second floor, enriched with "WHICH" questions, encompasses comprehensive RDG areas covering procedural and relational mechanisms. Table 7.1 and Table 7.2 list the areas within RDG, under the procedural and relational mechanisms, respectively.

Table 7.1: The "WHICH" RDG areas categorized under Procedural Mechanism

No	Procedural mechanism	Description
1	Research Data Policy	All policies, procedures, guidelines, standards, and requirements that have been developed within an institution or partnership for the management, archiving and sharing of research data.
2	Research Data Strategy	How an organization improves specific business objectives by strategically using its research data as assets.
3	Research Data Ownership & IPR	The legal or moral rights that gives individuals, groups, or organizations the authority to determine storage, retention, disposal, publication, or licensing arrangements of ideas, inventions, and creative expressions.
4	Performance Measurement	Evaluating how well research data are managed and the value they deliver for users and other stakeholders.
5	Issue & Risk Management	The ability to identify and monitor risks (threats and opportunities), to plan and implement responses to those risks, and respond to other issues that affect the change initiative.
6	Compliance Monitoring	The observation of the environment to identify gaps between the actual operations, the internal policies and standards, and the requirements as they derive from external regulations, laws, and orders.
7	Research Data Stewardship	The oversight of the entire research data lifecycle, aiming to ensure that the right processes are put in place and that appropriate decisions are made to make research datasets Findable, Accessible, Interoperable and Reusable (FAIR).
8	Research Data Custodianship	A set of activities required to ensure that important datasets are developed, maintained, and are accessible within their defined specifications.
9	Research Data Selection	The process of choosing research datasets, which are considered worth long-term preservation by a data archive.
10	Research Data Repository	Digital platform where research data may be stored and made available. This includes but is not limited to an institutional data repository, a discipline specific repository, a funder repository, or a commercial system.
11	Research Data Retention & Disposal	The long-term storage of research data and records after the completion of a research activity/project, for the purposes of meeting legal obligations or other purposes. Meanwhile, research data disposal describes how research data is destroyed.
12	Research Data Privacy	The rights and obligations of individuals and the organization with respect to the collection, use, retention, and disclosure of personal information.
13	Research Data Citation	A reference to research data for the purpose of credit attribution and facilitation of access.
14	Research Data Integrity	The maintenance of, and the assurance of the completeness, accuracy, and consistency of research data over its entire lifecycle from inconsistencies, accidental or malicious alteration or destruction.
15	Research Data Licensing	A license applied to research data or research datasets that clearly defines how and in which conditions it can be reused and guides future re-users.
16	Research Data Sharing	The process of taking any type of research data and making it available for other researchers to examine or use.

Table 7.2: The "WHICH" RDG areas categorized under Relational Mechanism

No	Relational mechanism	Description
1	Decision-making coordination	Provides strategies for achieving alignment across departments on research data governance implementation and practices.
2	Communication	Attempts to raise awareness of the research data governance program among stakeholders constantly.
3	Training	A set of activities designed to improve research data management skills, enhancing performance, productivity, capacity building, and knowledge transfer

The procedural mechanism within RDG forms the backbone of a robust and ethical approach to RDM. Each area outlined above plays a crucial role in shaping the landscape of RDG, ensuring the responsible handling and utilization of research data. As technology evolves and the production of research data expands, a strategic focus on procedural aspects becomes imperative to harness the full potential of research outputs and contribute to advancements in various fields. In the dynamic landscape of RDG, the synergy between procedural and relational mechanisms is paramount to the effective and ethical management of research data. While procedural mechanism focuses on the systematic aspects, relational mechanism fosters collaboration among stakeholders through communication, decision-making coordination, and training initiatives.

As organizations work towards maintaining ethical standards, following regulations, and maximizing the value of research data, working in silos is not a compromise. By emphasizing effective communication, coordinated decision-making, and proper training, all parties can actively foster a culture of responsible and efficient research data management. This not only ensures compliance, but also promotes innovation and excellence in utilizing RDG.

7.2.2.3 RDG Decision Domains

Meanwhile, the upper segment of the functional dimension mirrors the "HOW" question, influencing how RDG is implemented by viewing these decision domains as its focal areas. In the context of RDG, decision domains denote specific areas or aspects that necessitate control, where decisions are made to govern and manage data effectively. This ensures that research data is handled in a manner aligning with organizational goals, complying with regulations, and meeting the needs of stakeholders. The eight identified decision domains within RDG are listed in Table 7.3.

Table 7.3: The "HOW" RDG Decision Domains

No	Decision Domain	Description
1	Research Data Principle	Setting the boundaries for the intended uses of research data, the organization's standards for data quality, and ultimately the foundation for how users will access and interpret the data.
2	Research Data Lifecycle	The process through which research data flow from creation, to processing, analysis, preservation, distribution/sharing and re-use.
3	Research Data Architecture	Set of rules, policies, standards, and models that govern and define the type of data collected and how it is used, stored, managed, and integrated within an organization and its database systems.
4	Research Data Storage & Infrastructure	The technology, policies, standards, and human resources necessary to acquire, process, store, distribute, and improve utilization of research data, including the retention of research data using technology specifically developed to keep that data and have it as accessible as necessary.
5	Research Data Security	The protection of research data from loss, unauthorized access, and unauthorized modification. Security must be maintained while research data is both at rest and in transit.
6	Metadata Management	A set of processes for determining structured information about the attributes of a dataset that enables the research data to be identified, retrieved, and managed over time.
7	Research Data Quality	A measure of the condition of research data based on factors such as accuracy, completeness, consistency, reliability and whether it is up to date.
8	Research Data Access	The degree to which the research data collected is made available to individuals, groups and/or the public.

These decision domains collaborate to form the framework for responsible and efficient RDG. Together, they prioritize the ethical handling of research data and align with organizational objectives. This ensures that research data is not only accessible but also contributes significantly to knowledge and innovation.

7.2.2.4 Vision/Mission/Driver

At the highest level, the roof, shaped like a triangle, embodies the "WHY" question. The RPO strategically identifies its visions, missions, and/or drivers for implementing RDG. Positioned at the summit, the roof structure offers vital support for the entire building, distributing the load of the building's components and ensuring stability.

The drivers for RDG implementation encompass compelling reasons or factors that motivate an RPO to adopt robust RDG practices. It is imperative to establish driving principles and values that underpin the RDG framework, aligning RDG with organizational vision, mission, strategic objectives, and the specific needs of their research community. Furthermore, aligning these driving principles with ethical considerations and legal requirements is crucial for RPOs. Possible drivers, as outlined by the leading RPOs in their RDG/RDM policy, include:

a) *Openness and accessibility* - Emphasizing the FAIR principles of Open Science and the aspiration to optimize the use of publicly funded scientific publications and research data (University of Twente, 2018) is also stressed in the literature (Austin et al., 2021; Fürholz & Jaekel, 2021; Leonelli, 2018; Turkyilmaz-van der Velden et al., 2020).

b) *Research quality* - Highlighting good practices in RDM and open access to research data as integral to high-quality research (Ghent University, 2016; Menendez-Alonso, 2020), is also highlighted by the authors (Boufarss & Laakso, 2020; Vernon et al., 2021)

c) *Data reproducibility* - Underscoring best practices for ensuring the long-term reproducibility of scientific arguments and results (University of Cambridge, 2021; University of Cape Town, 2018) is also underscored by Leonelli et al. (2015) and Tu & Shen (2023).

d) *Data security* - Stressing responsible management of research data, including the secure storage of personal data and the protection of intellectual capital developed by researchers within the organization (Dunning et al., 2018). This emphasis is also echoed in the literature (Al-Ruithe et al., 2018; Nasir et al., 2023; Ng'eno & Mutula, 2018; Peukert et al., 2022)

e) *Data visibility* - Prioritizing research excellence by enhancing the visibility of academic work, leading to the recognition of the overall quality of the research process (Open University, 2020). Fürholz & Jaekel (2021) also stress the importance of data visibility through open research data.

f) *Funders requirements* - Focusing on improved practices for meeting the requirements of funders and publishers concerning research data management and sharing (Dunning et al., 2018) is also stressed in the literature (Anger et al., 2024; Cox et al., 2019b; Gaba et al., 2020)

g) *Research data as asset* - Acknowledging that research data are valuable forms of scientific output that should be made accessible to the research community (London School of Hygiene & Tropical Medicine, 2022). These literature (Abraham et al., 2019; Alhassan et al., 2019a; DAMA International, 2017;

Ladley, 2012; Omar & Almaghthawi, 2020) have also emphasized research data as an asset.

h) *Data preservation* - Underscoring the importance of the long-term availability of research data (Macquarie University, 2021; Nanyang Technological University, 2019) is also mentioned by Adika & Kwanya (2020) and Liu et al. (2020).

i) *Data compliance* - Emphasizing the significance of achieving compliance with data policies (University of Cambridge, 2021). Data compliance is also underscored in the literature (Abraham et al., 2019; Al-Ruithe et al., 2019; Chigwada et al., 2017; Marlina et al., 2022; Peng et al., 2016; Wende & Otto, 2007; Yebenes & Zorrilla, 2019).

Understanding the drivers that propel its implementation is essential for RPOs. By aligning RDG practices with organizational goals, ethical considerations, and legal requirements, institutions can foster a culture that not only meets the demands of funders and publishers but also promotes openness, accessibility, and the long-term impact of valuable research data. This approach supports the principles of FAIR and CARE, as outlined in Section 2.3.1.

In essence, although the RDG framework is crafted with a specific national context in mind for RPOs in Malaysia, its principles and insights resonate for wider RPO community. It stands as a testament to the collaborative nature of advancing RDG practices, offering a shared resource that transcends borders and encourages a collective effort towards effective RDG.

7.3 Summary of Chapter Seven

The discussion in this chapter encompasses various aspects of RDG, providing insights into the construction of its framework. The metaphor likens RPO as a structure building, emphasizing the importance of foundational values and the collaborative efforts of individuals at different levels of governance. It underscores perspectives of both effectiveness and efficiency. In addition, it highlights on various facets of RDG, and the continuous updates required for optimal functionality. Furthermore, the chapter examines the drivers behind RDG implementation, emphasizing the need to align RDG with organizational vision, mission, and ethical considerations. Meanwhile, Chapter Eight discusses the findings to confirm their consistency and relevance with previous literature.

CHAPTER 8: DISCUSSION AND CONCLUSIONS

"Invite to the way of your Lord with wisdom and good instruction, and argue with them in a way that is best..."
(Surah An-Nahl, 16:125)

8.1 Introduction

This final chapter synthesizes the key findings discovered throughout the study. It engages in a discussion against the existing literature, evaluating how the findings align with or diverge from previous studies, in relation to the following research objectives:

RO1: To investigate research data governance activities in relation to both actual implementation and perceived importance among data practitioners within research performing organizations.

RO2: To identify the stakeholders of research data governance roles within research performing organizations, including their governance structure.

RO3: To develop a research data governance framework for research performing organizations.

The discussion highlights the significance of comprehending and implementing RDG to achieve long-term objectives and ensure data integrity, efficiency, and compliance within RPOs. This chapter consolidates the theoretical and practical implications of the study, reflecting on the established framework, its dimensions, and the operationalization of RDG practices. The key elements are revisited, including the roles and responsibilities defined at various levels, the strategic, tactical, and operational tasks, and the interconnected mechanisms that drive effective data governance. This comprehensive overview emphasizes the importance of a structured approach to RDG, and proposes recommendations for future research, contributing to the ongoing discourse in the field.

8.2 Discussion of Findings

The discussion is organized in alignment with each research objective (RO).

8.2.1 Research Data Governance Activities

This section addresses RO1: *To investigate research data governance activities in relation to both actual implementation and perceived importance among data practitioners within research performing organizations.*

8.2.1.1 Implementation of RDG Activities by Malaysia RPOs

The findings of the desk research were used as foundational information to benchmark what has been implemented in Malaysia's RPOs regarding RDG. The desk research employed a rigorous coding procedure to analyze the RDG activities of the top 36 leading RPOs, predominantly from the United Kingdom. As highlighted by Tuesta et al. (2019), Germany, the United Kingdom, Canada, and Italy were noted for having a higher number of universities in top positions. Understanding how these top universities prioritize and execute RDG activities provides valuable insights into best practices and trends within the academic research landscape. It becomes evident that the leading RPOs worldwide prioritize 'implementing' activities over 'defining' or 'monitoring' activities. This emphasis on implementation underscores the proactive stance of these institutions towards actualizing their data governance frameworks. This aligns with the previous study (Alhassan et al., 2018), which revealed that practice-oriented organizations tend to have a greater emphasis on data governance activities related to implementation. This suggests that RPOs are increasingly recognizing the importance of RDG in maximizing the value of research data (Abraham et al., 2019; Brous & Janssen, 2020; Kouper et al., 2020; Omar & Almaghthawi, 2020).

By analyzing individual RDG activities, a more detailed understanding of the specific actions that the leading RPOs prioritize in their RDM practices was identified. The findings indicate a strong emphasis on activities related to compliance monitoring for data principles, as evidenced by the high frequency of both 'Implement Compliance Monitoring for Data Principle' and 'Monitor Compliance Monitoring for Data Principle' activities. These findings highlight RPOs' dedication to upholding ethical and legal standards for RDM, aligning with observations in previous studies (Kouper et al., 2020; Liu et al., 2020; Marlina et al., 2022). Data protection is governed by relevant legislation, covering its creation and usage (Lee et al., 2017; Parmiggiani & Grisot, 2020). Consequently, RDG should not be considered optional, as it ensures consistent and compliant operations, thereby enhancing organizational performance (Koltay, 2016).

Moreover, the results highlight the importance of developing a proactive approach toward RDM (Alhassan et al., 2019a; Willaert et al., 2019), as evidenced by the significant number of activities related to 'Define Data Strategy for Data Principle'. By developing clear strategies for managing research data effectively, RPOs can better align their RDM practices with their organizational objectives and ensure compliance with regulatory and funding agency requirements (Borghi & Van Gulick, 2022).

In general, the desk research offers valuable insights into the RDG activities of top-leading RPOs and underscores the importance of establishing comprehensive RDM policies to comply with regulatory and funding agency requirements. By gaining a better understanding of the RDG activities implemented by these leading RPOs, the results are seen to make a crucial contribution to developing the instruments for the Delphi study. This desk research holds importance in the overall investigation, acting as a preliminary phase before the modified Delphi study (Albotoush & Shau-Hwai, 2023; Münch et al.,

2021), and contributes qualitative insights into the RDG practices employed by leading RPOs worldwide (Kabir, 2016).

Comparing the data obtained from prior desk research reveals that, despite Malaysia's progressive efforts towards advocating open science and the early stages of RDM implementation, none of the RDG tasks fall under the "not implemented" category. While it is understandable that none of the tasks are classified as 'very highly implemented', due to the relatively nascent stage of these activities. This balance reflects a positive adaptability and growing intention of RDG practices among RPOs in Malaysia. The absence of 'not implemented' tasks suggest a strong foundational commitment, while the lack of 'very highly implemented' tasks indicate ongoing development and room for enhancement. This promising trend underscores Malaysia's potential to have good RDG as its practices continue to mature and evolve. According to research (Liu et al., 2020; Manik et al., 2022; Mansfield-Devine, 2017; Napis et al., 2019), widespread awareness of RDM and buy-in from all top management leaders of RPOs regarding the strategic importance of research data protection and governance, are crucial factors for easier and successful implementation of RDM initiatives. This is because most of the RDG tasks have been collectively familiarized and are currently being implemented to a moderate extent.

While some may argue that the level of implementation rated by the panelists is based on perception rather than actual implementation, in a post-positivist research, it is essential to acknowledge this challenge. In striving for objective responses, it is essential to recognize and accept the presence of an objective reality independent of human perception (Ryan, 2006). However, it is important to note that the understanding of this reality is inherently limited and shaped by various factors such as the subjectivity, interpretation, context, and analysis (Beeftink, 2005; Khanal, 2014; Maxwell &

Mittapalli, 2010), all of which influence the perceptions of the panelists who participated in the study. This acknowledgment of subjectivity is not a weakness but rather an inherent aspect of qualitative research, particularly in understanding complex phenomena such RDG. Hence, utilizing both desk research and the Delphi technique to gather quantitative and qualitative data is regarded as the most preferable option, as it helps to mitigate the risk of bias (Khanal, 2014; Sadler & Hammerman, 1999). This combined methodology allows for a more comprehensive understanding of RDG implementation, incorporating diverse perspectives and minimizing the influence of individual biases, thereby enhancing the robustness and reliability of the study's findings.

8.2.1.2 Importance of RDG Activities for Malaysia RPOs

The research findings provide a comprehensive understanding of the importance level attributed by expert panels to various governance tasks associated with different RDG roles. The findings underscore the important tasks, among others are related to ensuring data integrity, data quality, data security, and long-term accessibility, along with the protection of sensitive and confidential research data in accordance with legal and ethical standards, as supported in the literature (Brous et al., 2016; DAMA International, 2017; Thompson et al., 2015). These insights offer guidance for prioritizing and focusing on key RDG activities, which, in turn, can inform decision-making and resource allocation for the enhancement of RDG practices and implementation within RPOs.

Preserving researchers' rights and intellectual property rights emerged as significant priorities in this study. Furthermore, acknowledging the contributions of researchers and ensuring the responsible utilization of research materials are also deemed highly important tasks. These findings emphasize the need to foster an environment that promotes innovation (Sharif et al., 2018), recognizes researchers' contributions, and upholds ethical practices in research (Hendey et al., 2018; Nielsen, 2017; Parmiggiani &

Grisot, 2020). In addition, providing training on RDM were identified as important tasks too. These findings highlight the importance of enhancing capacity building in effective RDM, a point emphasized by studies by Gunjal & Gaitanou (2017), Kabanda et al. (2023); Lefebvre et al. (2018), Liu et al., 2020), and Palsdottir (2021). It is evident that RPOs should invest in initiatives that facilitate stakeholder involvement and provide comprehensive training programs to equip researchers with the necessary skills for managing and sharing data effectively.

Furthermore, the findings indicate that certain tasks require collaboration among multiple stakeholders, a concept also highlighted in studies by Abraham et al. (2019), Cerrillo-Mártinez & Casadesús-de-Mingo (2021), Cox et al. (2017), Lee et al. (2017), and Paskaleva et al. (2017). For example, a consensus emerged among the Delphi panelists regarding the importance of providing sufficient metadata for the discoverability and reusability of research data, which they deemed to be the responsibility of a 'Researcher'. Simultaneously, they suggested that 'National or Institutional Repositories' should delineate the necessary contextual information and metadata for deposited data. Combining these tasks undertaken by different entities could lead to a successful implementation of sufficient metadata quality level within the data repository (Johnston et al., 2018; Lee & Stvilia, 2017), ensuring the documentation, preservation, and discovery of the data (Alter & Gonzalez, 2018; Cox & Pinfield, 2014). 'Defining' the standard metadata required for entry into the repository by the national or institutional repositories would assist researchers in correctly 'implementing' and ensuring the adequacy of metadata within the system. For instance, consider the MOSP, serving as a national data repository; it should establish or identify the requisite standard metadata. With these explicit requirements in place, all contributors from universities and institutions involved would have a clear understanding of the expected standard metadata.

In this study, some tasks were regarded as less critical when compared to others. For example, the panelists assigned lower importance to the Research Data Governance Committee periodically reporting on the value contributed by RDM. On the other hand, they emphasized the significance of the Data Governance Leader in conveying the outcomes of the Research Data Governance Committee, instead. Similarly, the panelists assigned lower importance ratings to researchers budgeting for the costs associated with capturing, managing, archiving, and sharing research data throughout a project's lifespan, as well as the time investment required for RDM. However, they suggested that these tasks should fall under the responsibility of Research Data Curators instead. Consequently, the perceived importance of tasks is assessed in alignment with the assigned roles. This finding emphasizes the contextual nature of task importance within different roles, aligning with research findings by Abraham et al. (2019) and Alhassan et al (2016). It highlights the importance of considering individual perspectives and responsibilities when evaluating the significance of tasks. Therefore, it is important to note that a lower CL does not imply lack of importance; rather, it indicates that the task is considered less critical for that particular role. This finding encourages a nuanced understanding of task prioritization within specific organizational contexts.

Meanwhile, based on the findings of the gap analysis, a negative mean difference is evident for all RDG tasks, indicating a notable gap between the perceived importance of these tasks and their actual implementation. This discrepancy raises questions about the extent to which RPOs recognize the significance of these tasks in the context of RDG. Kabanda et al. (2023) highlight the importance for organizations to support or fund RDM and related infrastructure. The observed gap may indicate challenges in resource allocation or implementation strategies, potentially due to constraints in budget, expertise, or technology (Bakri et al., 2017; Cox et al., 2019a; Huang et al., 2021; Ministry of Education Malaysia, 2015), a lack of awareness or a cultural barrier within RPOs

regarding the importance of specific RDG tasks, or a need for training and skill development among personnel responsible for RDG task implementation. Establishing awareness and fostering a culture that values and prioritizes RDG are critical steps in addressing this gap (Abraham et al., 2019; Gunjal & Gaitanou, 2017; Hamad et al., 2019; Kabanda et al., 2023; Lefebvre et al., 2018; Liu et al., 2020; Manik et al., 2022; Napis et al., 2019; Omar & Almaghthawi, 2020; Palsdottir, 2021; Singeh et al., 2013b). Additionally, providing adequate training can empower individuals to effectively carry out these tasks, bridging the disparity between perceived importance and actual implementation.

8.2.2 Research Data Governance Roles

This section addresses RO2: *To identify the stakeholders of research data governance roles within research performing organizations, including their governance structure.*

8.2.2.1 Stakeholders of the Identified RDG Roles

The RDG roles identified in the study are ideally tailored for RPOs. Key roles, such as Executive Sponsor, Data Governance Leader, Data Governance Committee, Office of Data Governance, Data Steward, Data Curator, Data Owner, Researcher, Data Consumer, Administrative Offices, and External Bodies, have been unanimously agreed upon by the panelists. Within the Administrative Offices, among other entities, lie the library, IT services, researchers, and the research support office, which are commonly cited as collaborators (Andrikopoulou et al., 2022; Cox et al., 2017; Faniel & Connaway, 2018; Piracha & Ameen, 2019; Verbaan & Cox, 2014). Additionally, panelists recognized legal office (Pinfield et al., 2014), while also highlighting research ethics committees (Patterton et al., 2018) as potential partners.

In addition, the current study suggests new nomenclatures for some roles, such as Data Governance Executive, Data Governance Director, Research Data Steward, and Research Data Curator, specifically tailored for best practices in Malaysia RPOs. These changes are based on the study's findings and distinguish from other existing frameworks in the literature (Abraham et al., 2019; Korhonen et al., 2013; Solomonides, 2019; Wende & Otto, 2007; Yebenes & Zorrilla, 2019). In addition, the descriptions of each role have also been derived from the study and agreed upon by the panelists.

8.2.2.2 Governance Structure of RDG Roles

Benfeldt and colleagues (2020) describe data governance as the organizational and execution principles that establish decision-making and responsibility for an organization's data assets. The emphasis is on organizational structure and the attribution of power for data quality management (Jang & Kim, 2016; Weber et al., 2009). In terms of the structure, there are some similarities and differences when compared to the Data Governance Instantiation of Agile Governance Model by Korhonen et al. (2013). The similarities lie in the fundamental levels of governance, maintaining strategic steering, strategic implementation, tactical, and operational levels without involvement in day-to-day operations. Another common element is the emphasis on effectiveness and efficiency aspects, as the findings believe these to be integral in the RDG framework. However, the differences are found in the governance roles themselves, where the roles are more specific to cater to RDM, whereas the model proposed by Korhonen et al. (2013) focuses on a general data governance model with roles that are broader and based on the analysis of the literature on data governance topics. Nevertheless, these findings have achieved consensus among expert panels, comprising various data practitioners from different management levels, who believe the roles are suitable for implementation by RPOs in Malaysia.

Each role involves the contribution of ideal stakeholders who are responsible for specific governance tasks, ensuring the seamless operation of RDM practices. Providing flexibility in how people can contribute ultimately supports the creation of a culture of inclusion where individuals' contributions are valued (Wong, 2024). However, the acknowledgment that stakeholders holding these roles may differ and vary based on the size and needs of the organization (Chignard & Glatron, 2023; Suzuki et al., 2023), introduces a level of flexibility and adaptability in implementing the findings. Smaller organizations may have more streamlined governance structures with fewer dedicated roles due to limited resources, while larger organizations might require a more complex and hierarchical governance model to manage extensive research data. Academic settings and research institutions may also differ in terms of the availability of the positions. The acknowledgment of variability in stakeholders implies flexibility in applying the findings, allowing organizations to tailor the proposed roles to better fit their specific circumstances. Flexibility enables a customized approach, where organizations can select, combine, or modify roles based on their unique requirements.

8.2.3 Research Data Governance Framework

This section addresses RO3: *To develop a research data governance framework for research performing organizations*.

Given the government's growing focus on driving innovation and cultivating a knowledge-based economy, as well as initiatives aimed at promoting open science in Malaysia, there is a pressing need for RDM among RPOs (Jusoh, 2018; Ibrahim & Wei, 2023; Ministry of Education Malaysia, 2015). Likewise, the successful implementation of RDG requires a well-balanced combination of people, processes, and technologies (Akoum & Bu Hazzaa, 2019; Brous & Janssen, 2020; DAMA International, 2017; Kaewkamol, 2022; Koltay, 2016; Paskaleva et al., 2017). This aligns with the People,

Process, and Technology (PPT) framework developed by corporate management specialist Harold Leavitt in the early 1960s for effecting organizational change (Leavitt, 1965, as cited by David et al., 2023).

Committed individuals from various levels of governance, possessing diverse expertise, knowledge, and backgrounds, ensure the successful execution of RDG within the RPO (Ashiq et al., 2020; Kouper et al., 2020; Rosenbaum, 2020; Vilar & Zabukovec, 2019; Zhou, 2018). They collaborate to address both the effectiveness and efficiency of the implementation process (Korhonen et al., 2013). Ultimately, RDG is a collective responsibility shared by data management professionals (DAMA International, 2017). At the same time, organizational leaders and policymakers play a strategic role in shaping the overall structure. Their decisions influence the tactical and operational levels, affecting all aspects of RDG practices and their implementation (Abraham et al., 2019). Similarly, the use of advanced technological solutions and tools is crucial for ensuring the smooth running of the entire RDG processes, carried out by various governance roles (Al-Ruithe et al., 2018; Hannila et al., 2019; Koltay, 2016; Madison, 2020; Reis et al., 2018).

Analogously, having drivers, missions, and visions of RDG ensures that all entities involved in implementation align with the organizational direction (Al-Ruithe et al., 2019; Cato et al., 2015; DAMA International, 2017; Yebenes & Zorrilla, 2019). For a vision to be realistic, everybody involved in the change process must have a common understanding of the activity's goal and direction and a shared outlook on the desired future (DAMA International, 2017), e.g.; members of the organization must have a common understanding of research data accessibility (Austin et al., 2021). Aligning the treatment of data assets with the organization's goals increases their value (Ladley, 2012).

The development of an RDG framework based on the current study is indeed a significant endeavour with implications for the broader research community. The framework's alignment with common practices, particularly in terms of strategic steering, strategic implementation, tactical, and operational positions, establishes a unified basis for best practices. The uniqueness of the nomenclatures for certain roles identified in the framework, as defined through the Delphi study with expert consensus, adds a distinct layer of specificity. It recognizes diverse needs and nuances within the framework. In addition, the tasks allocated to each role have been identified and acknowledged by the panelists as pivotal responsibilities to be implemented in Malaysia RPOs. This framework aligns closely with Janssen et al. (2020), who emphasize that data governance involves organizations and their personnel defining, applying, and monitoring rules and authorities to ensure proper functioning and accountability throughout the data and algorithm life cycle.

In conclusion, the significance of research data as an asset for Malaysia RPOs underscores the critical need for an RDG framework. RPOs in Malaysia should invest resources and effort into developing and implementing effective governance frameworks that will contribute to enhancing RDM practices (Kabanda et al., 2023; Lefebvre & Spruit, 2021; Marlina & Purwandari, 2019; Palsdottir, 2021; Wong et al., 2020; Abraham et al., 2019). Additionally, the gap between the implementation of RDG tasks and their perceived importance among data practitioners in Malaysia demonstrates the need for an RDG framework to guide the process, aiming to improve practices and align them with their expected importance.

8.3 Contribution of the Study

The study contributes significantly to the existing body of knowledge by providing insights into the RDG in terms of:

a) *Theoretical contribution:* The findings of the study could contribute to the development of a theoretical foundation for understanding RDG implementation within the Malaysian context. This involves integrating nuanced insights into specific practices and activities identified as crucial by data practitioners within RPOs. Additionally, the study may inform critical roles in governing research data, advancing perspectives on organizational structures for effective RDG.

b) *Methodological contribution:* (i) The methodology, combining desk research and a Delphi study, serves as a valuable reference for future researchers exploring RDG practices. This dual approach provides a comprehensive understanding, blending insights from existing literature with expert consensus, potentially setting a precedent for similar studies in other contexts. The integration of desk research and Delphi study highlights the advantages of employing qualitative approaches to obtain quantitative results in RDG research. Future studies could consider adopting qualitative approaches to triangulate findings, providing a more robust understanding of RDG practices in diverse organizational settings. This approach allows researchers to gather rich qualitative insights while also generating quantitative data for statistical analysis. (ii) Besides that, the use of a modified Delphi technique for consensus building ensured participant anonymity, granting freedom of expression without restrictions. This approach effectively minimized the influence of dominant personalities and bias. Additionally, the technique is cost-effective and flexible, allowing data practitioners to contribute from

various geographical locations through online surveys, eliminating the need for physical contact. The study also benefited from the diverse professional backgrounds of the panelists who are data practitioners involved in handling research and/or publication data.

c) *Empirical contribution*: By leveraging these findings, RPOs can refine its RDG policies and practices to enhance the overall research ecosystem. Additionally, the study offers empirical insights into specific key roles crucial for governing research data within RPOs. Engaging with a diverse range of stakeholders and incorporating their expertise enables RPOs to establish an RDG framework aligned with best practices, meeting the evolving needs of researchers and the wider scientific community.

d) *Practical contribution*: (i) The adaptability of the framework for use in diverse contexts emphasizes its potential as a source of inspiration rather than a rigid guide. The findings should be seen as a living document that can be updated based on the organization's experience and changing circumstances. RPOs can leverage the framework as a foundation, tailoring the roles and structures to align with their specific needs and organizational dynamics. This adaptability promotes a spirit of shared learning, fostering collaboration, and the continuous improvement of RDG practices on a more global scale,

(ii) While the framework is deemed ideal for RPOs in Malaysia, its utility may extend beyond national borders. RPOs from other countries can view it as a valuable collection of best practices, providing insights into the effective structuring of RDG roles and the assignment of important RDG tasks for each role in the governance of research data. The framework becomes a collective reference point, allowing RPOs, including those outside Malaysia, to initiate

and enhance their RDG practices by drawing on the experiences and expertise encapsulated in the Malaysian model.

e) *Literature contribution*: The study makes a significant literature contribution by building upon the frameworks established by Alhassan et al. (2018) and Korhonen et al. (2013). By relying on these foundational works, the study not only reinforces existing theories but also provides a structured analysis of data governance activities and the roles. The study employs the Data Governance Activities Model proposed by Alhassan et al. (2018), which classifies data governance activities into specific categories. This model serves as a vital tool for organizing and interpreting data, facilitating a more systematic approach to data collection, coding, and analysis. By utilizing this framework, the study highlights how structured categorization enhances understanding of RDG activities and promotes best practices. Furthermore, the study incorporates the Agile Governance Model from Korhonen et al. (2013). This model assists in defining the governance structure of various roles identified in the study. By situating these roles within an RDG framework, the study illustrates how agile principles can enhance flexibility and responsiveness in data governance.

In conclusion, the findings arising from this study provide invaluable guidance for advancing RDG activities and identifying key roles within the organization. By diligently addressing the identified areas of significance and responsible key roles, RPOs can foster an environment deeply rooted in responsible and ethical RDM. Ultimately, this will bestow considerable benefits upon researchers and the broader scientific community. In essence, the contributions of the study revolve around providing theoretical advancements, methodological contributions, practical insights, empirical knowledge, and literature contribution, specifically tailored for RPOs in Malaysia, as they navigate the complexities of RDG.

8.4 Implications of the Findings

The implications of the study in establishing an RDG framework for Malaysia RPOs include:

a) *Practical implications:* (i) The identification of crucial RDG activities and roles offers practical guidance for RPOs in Malaysia. Organizations can use these findings to refine their existing RDG frameworks, ensuring alignment with best practices and addressing specific needs identified through the study. Through a thorough analysis of entities responsible for governing research data, the study recognizes diversity in backgrounds, expertise, and knowledge. This diversity enables effective collaboration, contributing to improved decision-making processes and overall data quality within RPOs. (ii) The study aims to effect positive social change by exploring the significance of RDG activities and roles, guiding data practitioners in understanding their role in maintaining data integrity. The potential positive social change resulting from the implementation of the study's findings is substantial, as it has the capacity to increase societal value by enhancing activities, providing greater protection, and expanding access to research data for a larger population. RPO leaders can consider the implementation of RDG efforts in their organizations by adopting the set of best practices identified in the study.

b) *Policy implications:* Beyond organizational boundaries, the insights gained may influence societal and policy changes. The study's findings can inform the development or refinement of RDG policies, standards, and guidelines at organizational and national levels, prompting regulatory bodies and policymakers to establish or revise regulations promoting research data quality and integrity across RPOs. Policymakers can use insights to create

regulations aligning with identified crucial activities and roles, fostering a standardized approach to RDG across Malaysia.

c) *Institutional implications:* (i) RPOs can allocate resources more effectively by focusing on important activities and roles. Additionally, the identified roles can guide the development of targeted training initiatives to enhance the skills of individuals responsible for RDG. The study could contribute to a cultural shift within RPOs by raising awareness about the significance of RDG, leading to a more proactive approach to data governance and fostering a culture of responsibility and compliance. (ii) Meanwhile, the findings of the gap analysis underscore the need for strategic interventions to align the implementation of RDG tasks with their perceived importance. Addressing these discrepancies is crucial for optimizing RDM processes, ensuring compliance, and fostering a culture of effective RDG within the RPOs. Continuous monitoring and feedback mechanisms will be essential to track progress and adapt strategies as needed.

In summary, the implications of the findings on the RDG framework for Malaysia RPOs are extensive, encompassing practical guidance for policy development and organizational enhancement. These implications collectively contribute to the evolution of a context-specific RDG framework within the Malaysian research landscape, offering valuable insights for decision-making processes and guiding resource allocation strategies aim at improving RDG practices and implementation.

8.5 Limitations of the Study

While considering the aforementioned implications of the study, it is imperative to acknowledge the inherent limitations within the study:

a) *Limited generalizability*: The findings of the study may be specific to the context or characteristics of the RPOs included in the research sample, limiting the generalizability of the results to other types of institutions or research settings. Additionally, more responses were from universities compared to research institutes in both desk research and the Delphi study. Thus, the findings may have leaned towards the university setting compared to research institutes.

b) *Data availability and completeness*: The study may be limited by the availability or completeness of data on RDG practices within RPOs due to decreasing response rate in subsequent rounds, which could impact the comprehensiveness and depth of the analysis. For instance, in Round I, 47 data practitioners participated. This number further reduced in Round II, with 34 out of 47 invited data practitioners responding.

c) *Timing of data collection*: The study's findings may be influenced by the timing of data collection (April 2023 to September 2023), as RDG practices and policies within RPOs may evolve over time, potentially affecting the relevance and applicability of the results. It is due to the insights derived from the data collected during this period may become outdated if significant changes occur soon after the data collection period. This may affect the long-term applicability and relevance of the study's conclusions.

d) *Focus on specific aspects*: The study narrowly focuses on specific aspects of RDG within RPOs, primarily research data management policies, while overlooking other critical dimensions that could significantly impact RDG

effectiveness. These dimensions include capacity building, infrastructure, and fostering data sharing and collaboration. This oversight may lead to biased findings and recommendations that do not adequately address the diverse needs and challenges encountered in RDG practices.

Overall, despite these limitations, organizations seeking to implement or assess the current state of an RDG program may find the findings of the study useful. Acknowledging these limitations is essential for a comprehensive understanding of the scope and potential implications of the study. Researchers should consider these constraints when interpreting and applying the findings in the broader RDG domain.

8.6 Future Research

The study identifies research gaps and opportunities for RDG research. This assists researchers in selecting topics and refining methods for future studies. The exploration of emerging trends and new technologies in RDG practices is essential through ongoing research in this field to better support researchers, institutions, and society. The study recommends further research based on its findings, as discussed below.

a) *Empirical Assessment of the RDG Framework*: This study contributes to the field by recommending that future research empirically assess the developed RDG framework to evaluate its effectiveness and applicability in various contexts.

b) *The Integration between RDG and other Governance Models*: Future research should investigate how RDG roles and its governance structure can be effectively integrated with other governance frameworks. It should focus on the processes and mechanisms that ensure RDG works alongside, rather than in isolation, helping to create more adaptive, collaborative, and resilient organizational structures.

c) *Enhancing RDG Effectiveness through Stakeholder Engagement and Strategic Insights*: Based on the findings of the gap analysis, since the perceived importance is consistently greater than the actual implementation, RPOs may have the opportunity to concentrate on bridging the gap. By identifying and addressing the contributing factors to this misalignment, RDG practices could be made more effective. Researchers may interact with stakeholders, such as organizational leaders and data practitioners, to obtain their insights regarding the perceived importance and actual implementation of these tasks. Further examination of the factors, obstacles, and challenges that impede the effective execution of RDG tasks, notwithstanding their perceived significance, is necessary to gain a more comprehensive understanding of the situation. Gaining an understanding into the factors that influence individuals' expectations, limitations, and motivations can provide valuable insights for developing effective strategies to enhance collaboration and participation and to overcome the challenges, thereby improving the alignment between the perceived importance and actual implementation. Examining the impact of training programs and workshops on improving RDG awareness and practices among the research community is also recommended.

d) *Interdisciplinary Collaboration in RDG*: It is of the utmost importance to analyze the opportunities and challenges that arise from interdisciplinary collaboration and research in the context of RDG. RPOs frequently encompass a wide range of disciplines; therefore, it is critical to comprehend the integration and management of research data from various domains within a unified governance framework. An examination of the impact of RDG on promoting collaboration and the exchange of knowledge among researchers

from diverse fields of study enhances the overall comprehension of the complexities inherent in interdisciplinary research.

e) *Intersection of RDG and Open Science*: Aside from that, investigating the way in which RDG and open science initiatives intersect is crucial too. Examine the ways in which RDG practices can support and correspond with the tenets of transparency, replication, and cooperation in research within the dynamic realm of academic communication. It is also imperative for RPOs to confront the ethical aspects of RDG, with a particular focus on privacy, consent, and responsible data utilization. Investigating the effects that ethical guidelines and regulations have on the formulation and execution of RDG policies is essential to guarantee adherence to the tenets of responsible research conduct. The governance landscape is significantly influenced by ethical considerations; therefore, it is imperative that research pursues a harmonious balance between ensuring data accessibility and upholding ethical standards.

f) *Ensuring Quality in RDG*: The role of RDG in ensuring and enhancing the quality of research data throughout its lifecycle is a critical aspect. Exploring methodologies for assessing and improving data quality, including the integration of automated tools and validation processes within RDG framework, can significantly contribute to the reliability and credibility of research outcomes. Practical strategies for maintaining high data quality standards in diverse research settings should be developed.

g) *Research Data Security and Compliance*: RPOs may encounter compliance and security issues pertaining to research data, specifically in domains where sensitive or confidential data is involved. It is crucial to conduct research on approaches to guarantee research data security, fulfil regulatory obligations,

and handle compliance concerns. Understanding the distinct obstacles that RPOs face in upholding research data security and compliance, suggesting efficacious resolutions, and harmonizing RDG procedures with the ever-changing regulatory environments ought to be the primary objectives of research.

By considering these potential areas for further investigation, the academic community can make a valuable contribution to the ongoing enhancement and adjustment of RDG methodologies to align with the changing demands of the scientific domain.

8.7 Conclusions

The study has provided a thorough exploration of RDG practices within Malaysia RPOs. By combining desk research and the Delphi study, the aim was to identify key RDG activities, pinpoint essential stakeholders in relevant roles, and define their governance structure within these organizations.

Through an analysis of policy documents on RDG/RDM and the synthesis of expert opinions, a set of activities and the assigned roles deemed pivotal for effective RDG implementation within Malaysia RPOs have been identified. In addition, the study has illuminated the organizational structures required for successful RDG, emphasizing the strategic placement of roles responsible for governance activities. This insight into structural positions provides practical recommendations for optimizing the organizational hierarchy, thereby enhancing the overall efficiency and effectiveness of RDG practices. Therefore, the findings contribute valuable insights to the development of RDG framework, offering both structural and functional dimensions. These findings serve as a guide for organizations seeking to refine their existing RDG framework and contribute to the broader discourse on responsible RDG practices.

A comprehensive RDG framework must incorporate people, processes, and technologies. Continuous monitoring and adaptability are necessary to keep the framework relevant to changing research needs and technology. A dynamic and adaptive governance framework will be required to uphold data integrity, security, and ethics as research environments and data landscapes evolve. It is important to note that the findings serve as a guidance rather than strict rules, recognizing that each organization has its own dynamics and context. This approach allows for a balance between providing a structured framework and allowing room for adaptation. For organizations initiating RDG, the findings act as a starting point, offering insights into potential roles and stakeholders. However, they are not prescriptive, it is recommended that the organizations engage in regular evaluations and reassessments of their governance roles to account for evolving requirements, technological progress, and shifts in the research environment.

In essence, this study represents a stepping stone in the ongoing dialogue on effective RDG practices. It provides a foundation for future investigations, encourages a culture of responsibility and compliance within the research community, and contributes to the continual refinement of RDG framework to meet the dynamic challenges of the research landscape. As the scholarly community continues to evolve, adapt, and innovate, the insights gained from this study will hopefully catalyze further advancements in responsible and impactful RDG.

While acknowledging these contributions, it is important to acknowledge the study's limitations, which involve challenges such as limited generalizability, data availability and completeness, timing of data collection, and a predominant focus on policy aspects. These limitations should be taken into consideration when interpreting the results and suggest avenues for further research. Future research endeavours could explore closing gaps in RDG practices, involving RPOs to identify issues and overcome obstacles. This

exploration should also cover emerging trends, innovative solutions, and how RDG impacts collaboration, knowledge exchange, and ethics. Research could further analyze RDG's role in ensuring data quality, emphasizing strategies for reliability. Addressing compliance and security concerns, adapting to regulations, and promoting ongoing improvements in an inclusive manner are essential. Embracing these recommendations for research data evolution aligns with scientific demands, fostering effective RDG.

REFERENCES

Abd Rahman, N. (2019). The need for open science. *Journal of Research Management and Governance, 2*(1), 22–30. https://doi.org/10.22452/jrmg.vol2no1.3

Abduldayan, F. J., Abifarin, F. P., Oyedum, G. U., & Alhassan, J. A. (2021). Research data management practices of chemistry researchers in federal universities of technology in Nigeria. *Digital Library Perspectives, 37*(4), 328–348. https://doi.org/10.1108/DLP-06-2020-0051

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

Adika, F. O., & Kwanya, T. (2020). Research data management literacy amongst lecturers at Strathmore University, Kenya. *Library Management, 41*(6/7), 447–466. https://doi.org/10.1108/LM-03-2020-0043

Aghimien, L. M., Aigbavboa, C. O., Anumba, C. J., & Thwala, W. D. (2023). A confirmatory factor analysis of the challenges of effective management of construction workforce in South Africa. *Journal of Engineering, Design and Technology*, *21*(4), 1134–1152. https://doi.org/10.1108/JEDT-05-2021-0264

Ahmadi, S., Tavana, M. M., Shokouhyar, S., & Dortaj, M. (2022). A new fuzzy approach for managing data governance implementation relevant activities. *The TQM Journal*, *34* (5), 979-1012. https://doi.org/10.1108/TQM-01-2021-0015

Ahmed, M., & Othman, R. (2021). Readiness towards the implementation of open science initiatives in the Malaysian comprehensive public universities. *The Journal of*

Academic Librarianship, 47(5), 102368. https://doi.org/10.1016/j.acalib.2021.102368

Akartuna, E. A., Johnson, S. D., & Thornton, A. (2022). Preventing the money laundering and terrorist financing risks of emerging technologies: An international policy Delphi study. *Technological Forecasting and Social Change*, *179*, 121632. https://doi.org/10.1016/j.techfore.2022.121632

Akins, R. B., & Cole, B. R. (2005). Stability of response characteristics of a Delphi panel: Application of bootstrap data expansion. *BMC Medical Research Methodology, 5*(37). https://doi.org/10.1097/ACM.0000000000003318

Akoum, M., & Bu Hazzaa, H. (2019, October). *A data governance framework: The foundation for data management excellence* [Paper presentation]. SPE Gas & Oil Technology Showcase and Conference, Dubai, UAE. D021S009R002. https://doi.org/10.2118/198593-MS

Al Kez, D., Foley, A. M., Laverty, D., Del Rio, D. F., & Sovacool, B. (2022). Exploring the sustainability challenges facing digitalization and internet data centers. *Journal of Cleaner Production, 371,* 13363. https://doi.org/10.1016/j.jclepro.2022.133633

Al-Ruithe, M., Benkhelifa, E., & Hameed, K. (2016). A conceptual framework for designing data governance for cloud computing. *Procedia Computer Science*, *94*, 160–167. https://doi.org/10.1016/j.procs.2016.08.025

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

Al-Ruithe, M., Benkhelifa, E., Jararweh, Y., & Ghedira, C. (2018). Addressing data governance in cloud storage: Survey, techniques and trends. *Journal of Internet Technology*, *19*, 1763–1775. https://doi.org/10.3966/160792642018111906013

Al-Wahshi, J. J., Abbott, J., & Abbott, P. (2022). An investigation into the role of data governance in improving data quality: A case study of the Omani banking sector. *ECIS 2022 Research Papers*, 121. https://aisel.aisnet.org/ecis2022_rp/121

Alanazi, A. (2016). A critical review of Constructivist Theory and the emergence of Constructionism. *American Research Journal of Humanities and Social Sciences, 2*, 1-8. https://doi.org/10.21694/2378-7031.16018

Albotoush, R., & Shau-Hwai, A. T. (2023). Overcoming worldwide Marine Spatial Planning (MSP) challenges through standardizing management authority. *Ocean & Coastal Management*, *235*, 106481. https://doi.org/10.1016/j.ocecoaman.2023.106481

Aleixandre-Benavent, R., Sapena, A. F., Ferrer, S. C., Peset, F., & García, A. G. (2019). Policies regarding public availability of published research data in pediatrics journals. *Scientometrics*, *118*(2), 439–451. Scopus.

Alfaras, R. S. A. (2023). Integrated urban solid waste management: Knowledge, practices, and implementation. *Journal of Environmental Management and Tourism*, *14*(5), 2328. https://doi.org/10.14505/jemt.v14.5(69).14

Alhassan, I., Sammon, D., & Daly, M. (2016). Data governance activities: An analysis of the literature. *Journal of Decision Systems*, *25*(sup1), 64–75. https://doi.org/10.1080/12460125.2016.1187397

Alhassan, I., Sammon, D., & Daly, M. (2018). Data governance activities: A comparison between scientific and practice-oriented literature. *Journal of Enterprise Information Management*, *31*(2), 300–316. https://doi.org/10.1108/JEIM-01-2017-0007

Alhassan, I., Sammon, D., & Daly, M. (2019a). Critical success factors for data governance: A telecommunications case study. *Journal of Decision Systems*, *28*(1), 41–61. https://doi.org/10.1080/12460125.2019.1633226

Alhassan, I., Sammon, D., & Daly, M. (2019b). Critical success factors for data governance: A theory building approach. *Information Systems Management*, *36*(2), 98–110. https://doi.org/10.1080/10580530.2019.1589670

Allela, A., & Mwai, N. (2019). Strategies for research data management (RDM) at the Technical University of Kenya. In T. Kwanya, J. Kiplang'at, J. Wamukoya, & D. Njiraine (Eds.), *Digital technologies for information and knowledge management* (pp. 155–158). The Technical University of Kenya.

Allen, M., & Cervo, D. (2015). *Multi-domain master data management: Advanced MDM and data governance in practice* (1st ed.). Morgan Kaufmann Publishers Inc.

Alsheikh-Ali, A. A., Qureshi, W., Al-Mallah, M. H., & Ioannidis, J. P. A. (2011). Public availability of published research data in high-impact journals. *PLoS ONE*, *6*(9), 1–4. https://doi.org/10.1371/journal.pone.0024357

Alter, G., & Gonzalez, R. (2018). Responsible practices for data sharing. *American Psychologist*, *73*(2), 146–156. https://doi.org/10.1037/amp0000258

Alvarez-Romero, C., Martínez-García, A., Bernabeu-Wittel, M., & Parra-Calderón, C. L. (2023). Health data hubs: An analysis of existing data governance features for

research. *Health Research Policy and Systems*, *21*(1), 70. https://doi.org/10.1186/s12961-023-01026-1

Amanullah, S. W., & Abrizah, A. (2023). The landscape of research data management services in Malaysian academic libraries: Librarians' practices and roles. *The Electronic Library*, *41*(1), 63–86. https://doi.org/10.1108/EL-06-2022-0135

Ambrose, K. (2016). *Scientific Research Data Management Policy*. The Francis Crick Institute. 1 (crick.ac.uk)

Andrikopoulou, A., Rowley, J., & Walton, G. (2022). Research data management (RDM) and the evolving identity of academic libraries and librarians: A literature review. *New Review of Academic Librarianship*, *28*(4), 349-365. https://doi.org/10.1080/13614533.2021. 1964549

Anduvare, E. M., & Mutula, S. M. (2019). Research data management challenges in Kenya: The case of private universities in Nairobi County. In D. Ocholla & N. Evans (Eds.), *Proceedings of the 20th IS Annual Conference 2020* (pp. 258–278). Rosslyn Publishers.

Angelopoulos, M. K., & Pollalis, Y. A. (2020). Use of open data as a tool for successful lean management in public services: Evidence from Greece. *SPOUDAI Journal of Economics and Business*, *70*(1–2), 117–127.

Anger, M., Wendelborn, C., & Schickhardt, C. (2024). German funders' data sharing policies: A qualitative interview study. *PLOS ONE*, *19*(2), e0296956. https://doi.org/10.1371/journal.pone.0296956

Aronson, B. D., Janke, K. K., & Traynor, A. P. (2012). Investigating student pharmacist perceptions of professional engagement using a modified Delphi process.

American Journal of Pharmaceutical Education, 76(7), 125. https://doi.org/10.5688/ajpe767125

Ashiabi, E. E., & Avea, A. P. (2019). The absence of a disability measurement system in the disbursement of the district assembly common fund for persons with disabilities in Ghana: How the most vulnerable are denied access. *Review of Disability Studies: An International Journal*, *15*(4), 1–17.

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

Austin, T., Bei, K., Efthymiadis, T., & Koumoulos, E. (2021). Lessons learnt from engineering science projects participating in the Horizon 2020 open research data pilot. *Data, 6*(96), 1–19. https://doi.org/10.3390/data6090096

Azeroual, O. (2020). Treatment of bad big data in research data management (RDM) systems. *Big Data and Cognitive Computing*, *4*(29), 1–11.

Azeroual, O., Schöpfel, J., Ivanovic, D., & Nikiforova, A. (2022). Combining data lake and data wrangling for ensuring data quality in CRIS. *Procedia Computer Science*, *211*, 3–16. https://doi.org/10.1016/j.procs.2022.10.171

Azimi, N., Caldera, F., Cohen, S., Conners, J., Fernandes, T., Han, M., Strand, V., Tapson, V., Weinberg, A., Weinberg, J., & Yarur, A. (2021). Immune-mediated diseases and thromboembolic events: A modified Delphi panel. *Current Medical Research and Opinion*, *37*(8), 1283–1291. https://doi.org/10.1080/03007995.2021.1932450

Bahlai, C., Bartlett, L., Burgio, K., Fournier, A., Keiser, C., Poisot, T., & Whitney, K. (2019). Open science isn't always open to all scientists. *American Scientist*, *107*(2), 78–82. https://doi.org/10.1511/2019.107.2.78

Bakri, A., Azura, N. M., Nadzar, M., Ibrahim, R., & Tahira, M. (2017). Publication productivity pattern of Malaysian researchers in Scopus from 1995 to 2015. *Journal of Scientometric Research*, *6*(2), 86–101. https://doi.org/10.5530/jscires.6.2.14

Barber, P., Pontillo, M., Bellm, E., & Davies, G. (2023). Objective and subjective measures to guide upper extremity return to sport testing: A modified Delphi survey. *Physical Therapy in Sport*, *62*, 17–24. https://doi.org/10.1016/j.ptsp.2023.05.009

Barrett, D., & Heale, R. (2020). What are Delphi studies? *Evidence Based Nursing*, *23*(3), 68–69. https://doi.org/10.1136/ebnurs-2020-103303

Beeftink, K. (2005). Positivistic philosophy: An interpretivistic case study of tourist expectations & satisfaction. In J. G. Peden, & R. M. Schuster (Eds.), *Proceedings of the 2005 Northeastern Recreation Research Symposium* (pp. 345–354). U.S. Forest Service, Northeastern Research Station.

Beiderbeck, D., Evans, N., Frevel, N., & Schmidt, S. L. (2023). The impact of technology on the future of football: A global Delphi study. *Technological Forecasting and Social Change*, *187*, 122186. https://doi.org/10.1016/j.techfore.2022.122186

Beiderbeck, D., Frevel, N., von der Gracht, H. A., Schmidt, S. L., & Schweitzer, V. M. (2021). Preparing, conducting, and analyzing Delphi surveys: Cross-disciplinary

practices, new directions, and advancements. *MethodsX*, *8*(101401), 1–20. https://doi.org/10.1016/j.mex.2021.101401

Belcher, B. M., Rasmussen, K. E., Kemshaw, M. R., & Zornes, D. A. (2016). Defining and assessing research quality in a transdisciplinary context. *Research Evaluation*, *25*(1), 1–17. https://doi.org/10.1093/reseval/rvv025

Benfeldt, O., Persson, J. S., & Madsen, S. (2020). Data governance as a collective action problem. *Information Systems Frontiers*, *22*, 299–313. https://doi.org/10.1007/s10796-019-09923-z

Beno, M., Figl, K., Umbrich, J., & Polleres, A. (2017). Open data hopes and fears: Determining the barriers of open data. *2017 Conference for E-Democracy and Open Government (CeDEM)*, *Krems, Austria*, 69–81. https://doi.org/10.1109/CeDEM.2017.22

Berg, B. L., & Lune, H. (2017). *Qualitative research methods for the social sciences* (Ninth edition). Pearson.

Berger, E., Baidawi, S., D'Souza, L., Mendes, P., Morris, S., Bollinger, J., & Purtell, J. (2023). Educational experiences and needs of students in out-of-home care: A Delphi study. *European Journal of Psychology of Education*, 1-22. https://doi.org/10.1007/s10212-023-00714-4

Bertagnolli, M. M., Khozin, S., & Murphy, M. J. (2017). Data-sharing models. *New England Journal of Medicine*, *376*(23), 2305–2306. https://doi.org/10.1056/NEJMc1705477

Beyan, O., Choudhury, A., van Soest, J., Kohlbacher, O., Zimmermann, L., Stenzhorn, H., Karim, Md. R., Dumontier, M., Decker, S., da Silva Santos, L. O. B., &

Dekker, A. (2020). Distributed analytics on sensitive medical data: The personal health train. *Data Intelligence*, *2*(1–2), 96–107. https://doi.org/10.1162/dint_a_00032

Bhaskar, R. (1975). *A realist theory of science*. Leeds Books. https://archive.org/details/realisttheoryofs0000bhas/page/n7/mode/2up

Bietz, M. J., Bloss, C. S., Calvert, S., Godino, J. G., Gregory, J., Claffey, M. P., Sheehan, J., & Patrick, K. (2016). Opportunities and challenges in the use of personal health data for health research. *Journal of the American Medical Informatics Association*, *23*(e1), e42–e48. https://doi.org/10.1093/jamia/ocv118

binti Azizatun Nafi'ah. (2021). Data governance mechanisms in Indonesia's Covid-19 data integration system. *Sosiohumaniora*, *23*(3), 339. https://doi.org/10.24198/sosiohumaniora.v23i3.28437

Bloor, M., Sampson, H., Baker, S., & Dahlgren, K. (2015). Useful but no Oracle: Reflections on the use of a Delphi group in a multi-methods policy research study. *Qualitative Research*, *15*(1), 57–70. https://doi.org/10.1177/1468794113504103

Borda, A., Gray, K., & Fu, Y. (2020). Research data management in health and biomedical citizen science: Practices and prospects. *JAMIA Open*, *3*(1), 113–125. https://doi.org/10.1093/jamiaopen/ooz052

Borghi, J., & Van Gulick, A. (2022). Promoting open science through research data management. *Harvard Data Science Review*, *4*(3). https://doi.org/10.1162/99608f92.9497f68e

Borkakoti, R. (2021). Research data management in central universities and institutes of national importance: A perspective from North East India. *Library Philosophy and Practice*, May, 5848. https://doi.org/10.1177/09610006211070282

Boufarss, M., & Laakso, M. (2020). Open Sesame? Open access priorities, incentives, and policies among higher education institutions in the United Arab Emirates. *Scientometrics*, *124*(2), 1553–1577. https://doi.org/10.1007/s11192-020-03529-y

Boullenois, C. (2021, October). China's data strategy: Creating a state-led market. *Brief*, 1–8.

Bourne, P. E., Polka, J. K., Vale, R. D., & Kiley, R. (2017). Ten simple rules to consider regarding preprint submission. *PLOS Computational Biology*, *13*(5), e1005473. https://doi.org/10.1371/journal.pcbi.1005473

Bowser, A., Cooper, C., De Sherbinin, A., Wiggins, A., Brenton, P., Chuang, T. R., Faustman, E., Haklay, M., & Meloche, M. (2020). Still in need of norms: The state of the data in citizen science. *Citizen Science: Theory and Practice*, *5*(1). Scopus. citizenscienceassociation.org

Briner, K., & Rome, J. (2018). The need for and value of data governance. In K. L. Webber (Ed.), *Building capacity in institutional research and decision support in higher education* (pp. 67–80). Springer International Publishing. https://doi.org/10.1007/978-3-319-71162-1_5

Brous, P., & Janssen, M. (2020). Trusted decision-making: Data governance for creating trust in data science decision outcomes. *Administrative Sciences*, *10*, 81. https://doi.org/10.3390/admsci10040081

Brous, P., Janssen, M., & Vilminko-Heikkinen, R. (2016). Coordinating decision-making in data management activities: A systematic review of data governance principles. In H. J. Scholl, O. Glassey, M. Janssen, B. Klievink, I. Lindgren, P. Parycek, E. Tambouris, M. A. Wimmer, T. Janowski, & D. Sá Soares (Eds.), *Electronic government* (Vol. 9820, pp. 115–125). Springer International Publishing. https://doi.org/10.1007/978-3-319-44421-5_9

Buhomoli, O. S., & Muneja, P. S. (2022). Research data handling by researchers in the selected universities in Tanzania. *University of Dar Es Salaam Library Journal*, *16*(2), 53–69. https://doi.org/10.4314/udslj.v16i2.5

Bunkar, A. R., & Bhatt, D. D. (2020). Perception of researchers & academicians of Parul University towards research data management system & role of library: A study. *DESIDOC Journal of Library & Information Technology*, *40*(3), 139–146.

Burgi, P.-Y., Blumer, E., & Makhlouf-Shabou, B. (2017). Research data management in Switzerland: National efforts to guarantee the sustainability of research outputs. *IFLA Journal*, *43*(1), 5–21. https://doi.org/10.1177/0340035216678238

Busher, H., & James, N. (2007). Ethical issues in online educational research: Protecting privacy, establishing authenticity in email interviewing. *International Journal of Research & Method in Education*, *30*(1), 101–113. https://doi.org/10.1080/17437270701207868

Cahyaningtyas, D. F., & Priyanto, P. (2021). Strategies and challenges in archiving and sharing research data. *Seventh International Conference of Asian Special Libraries*, 1–6. https://icoasl2021.mlive.kr/program.php

Campbell, K. A., Wood, J. N., Lindberg, D. M., & Berger, R. P. (2021). A standardized definition of near-fatal child maltreatment: Results of a multidisciplinary Delphi process. *Child Abuse & Neglect*, *112*, 104893. https://doi.org/10.1016/j.chiabu.2020.104893

Carley, K. (1993). Coding choices for textual analysis: A comparison of content analysis and map analysis. *Sociological Methodology*, *23*(1993), 75–126. https://doi.org/10.2307/271007

Carroll, S. R., Herczog, E., Hudson, M., Russell, K., & Stall, S. (2021). Operationalizing the CARE and FAIR principles for Indigenous data futures. *Scientific Data, 8*(1), 108. https://doi.org/10.1038/s41597-021-00892-0

Cato, P., Gölzer, P., & Demmelhuber, W. (2015). An investigation into the implementation factors affecting the success of big data systems. *2015 11th International Conference on Innovations in Information Technology (IIT)*, 134–139. https://doi.org/10.1109/INNOVATIONS.2015.7381528

Cerrillo-Mártinez, A., & Casadesús-de-Mingo, A. (2021). Data governance for public transparency. *Profesional de La Información / Information Professional*, *30*(4), Article 4. https://doi.org/10.3145/epi.2021.jul.02

Cervone, H. F. (2017). Implementing IT governance: A primer for informaticians. *Digital Library Perspectives*, *33*(4), 282–287. https://doi.org/10.1108/DLP-07-2017-0023

Chalmers, I., Glasziou, P., & Godlee, F. (2013). All trials must be registered and the results published. *BMJ*, *346*, f105–f106. https://doi.org/10.1136/bmj.f105

Chambers, D. A., Vinson, C. A., & Norton, W. E. (2019). *Advancing the science of implementation across the cancer continuum*. Oxford University Press.

Chan, A. W., Song, F., Vickers, A., Jefferson, T., Dickersin, K., Gøtzsche, P. C., Krumholz, H. M., Ghersi, D., & van der Worp, H. B. (2014). Increasing value and reducing waste: Addressing inaccessible research. *Lancet*, *383*(9913), 257–266. https://doi.org/10.1016/S0140-6736(13)62296-5

Chapman, J. (2019). How safe is your data? Cyber-security in higher education. *HEPI Policy Note 12,* 1-6. https://www.hepi.ac.uk/wp-content/uploads/2019/03/Policy-Note-12-Paper-April-2019-How-safe-is-your-data.pdf

Che Jaafar, C. R. (2021, March 11). *Research data management and MyRA: Experience of Hamzah Sendut Library* [Video]. The Official IIUM Library Channel. https://www.youtube.com/watch?v=HG2D72Sl9xI

Check, J., & Schutt, R. K. (2012). *Research methods in education* (1st ed). Sage Publications Ltd.

Chignard, S., & Glatron, M. (2023). Data collaborations at a local scale: Lessons learnt in Rennes (2010–2021). *Data & Policy*, *5*, e20. https://doi.org/10.1017/dap.2023.16

Chigwada, J., Chiparausha, B., & Kasiroori, J. (2017). Research data management in research institutions in Zimbabwe. *Data Science Journal*, *16*(31), 1–9. https://doi.org/10.5334/dsj-2017-031

Chik, E. R. E., Rouse, F. M., Jaafar, C. R. C., Ismail, M. I., Azmi, N. A., Ghazali, M. M., & Ahmat, A. (2018). *Report on the contribution of universities and research*

institutions in scientific development of Islamic countries: The case of Malaysia (p. 22). Islamic World Science Citation Center.

Chiware, E. R. T., & Mathe, Z. (2015). Academic libraries' role in research data management services: A South African perspective. *South African Journal of Libraries and Information Science, 81*(2), 1–10. https://doi.org/10.7553/81-2-1563

Chyung, S. Y. Y., Roberts, K., Swanson, I., & Hankinson, A. (2017). Evidence-based survey design: The use of a midpoint on the Likert scale. *Performance Improvement, 56*(10), 15–23. https://doi.org/10.1002/pfi.21727

Colonna, R., Tucker, P., Holmes, J., Wilson, J., & Alvarez, L. (2022). Mobile-based brief interventions targeting cannabis-impaired driving among youth: A Delphi study. *Journal of Substance Abuse Treatment*, *141*, 108802. https://doi.org/10.1016/j.jsat.2022.108802

Cox, A. M., & Pinfield, S. (2014). Research data management and libraries: Current activities and future priorities. *Journal of Librarianship and Information Science*, *46*(4), 299–316. https://doi.org/10.1177/0961000613492542

Cox, A. M., Kennan, M. A., Lyon, E. J., Pinfield, S., & Sbaffi, L. (2019b). Progress in research data services: An international survey of university libraries. *International Journal of Digital Curation*, *14*(1), Article 1. https://doi.org/10.2218/ijdc.v14i1.595

Cox, A. M., Kennan, M. A., Lyon, L., & Pinfield, S. (2017). Developments in research data management in academic libraries: Towards an understanding of research

data service maturity. *Journal of the Association for Information Science and Technology, 68*(9), 2182–2200. https://doi.org/10.1002/asi.23781

Cox, A. M., Kennan, M. A., Lyon, L., Pinfield, S., & Sbaffi, L. (2019a). Maturing research data services and the transformation of academic libraries. *Journal of Documentation, 75*(6), 1432–1462. https://doi.org/10.1108/JD-12-2018-0211

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed). Pearson.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE.

Creswell, J. W., & Guetterman, T. C. (2019). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (6th ed.). Pearson.

Creswell, J. W., & Poth, C. N. (2017). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications, Inc.

Cruz-Castro, L., Martínez, C., Peñasco, C., & Sanz-Menéndez, L. (2020). The classification of public research organizations: Taxonomical explorations. *Research Evaluation, 29*(4), 377–391. https://doi.org/10.1093/reseval/rvaa013

Cruz, M., Dintzner, N., Dunning, A., van der Kuil, A., Plomp, E., Teperek, M., Turkyilmaz-van der Velden, Y., & Versteeg, A. (2019). Policy needs to go hand in hand with practice: The learning and listening approach to data management. *Data Science Journal, 18*(45), 1–11. https://doi.org/10.5334/dsj-2019-045

Cuhls, K. (2023). The Delphi method: An introduction. In M. Niederberger, O. Renn (Eds.), *Delphi methods in the Social and Health Sciences*. Springer, Wiesbaden. https://doi.org/10.1007/978-3-658-38862-1_1

Custer, R. L., Scarcella, J. A., & Stewart, B. R. (1999). The modified Delphi technique—A rotational modification. *Journal of Vocational and Technical Education, 15*(2), 1-10. https://doi.org/10.21061/jcte.v15i2.702

CWTS. 2022. *CWTS Leiden ranking*. Centre for Science and Technology Studies (CWTS). http://www.leidenranking.com

Dalkey, N. C. (1969). *The Delphi method: An experimental study of group opinion* (RM-5888-PR, pp. 1–79). RAND Corporation. https://www.rand.org/pubs/research_memoranda/RM5888.html

DalleMule, L., & Davenport, T. H. (2017). What's your data strategy? *Harvard Business Review*, *May-June*. https://hbr.org/2017/05/whats-your-data-strategy

DAMA International. (2009). *The DAMA guide to Data Management Body of Knowledge (DAMA-DMBOK guide)* (M. Mosley, M. Brackett, & S. Earley, Eds.). Technics Publications.

DAMA International. (2017). *The DAMA guide to the Data Management Body of Knowledge (DMBOK2)*. DAMA International.

David, A., Yigitcanlar, T., Li, R. Y. M., Corchado, J. M., Cheong, P. H., Mossberger, K., & Mehmood, R. (2023). Understanding local government digital technology adoption strategies: A PRISMA review. *Sustainability*, *15*(12), 9645. https://doi.org/10.3390/su15129645

De Lima, F. A., & Seuring, S. (2023). A Delphi study examining risk and uncertainty management in circular supply chains. *International Journal of Production Economics*, *258*, 108810. https://doi.org/10.1016/j.ijpe.2023.108810

de Loë, R. C., Melnychuk, N., Murray, D., & Plummer, R. (2016). Advancing the state of policy Delphi practice: A systematic review evaluating methodological evolution, innovation, and opportunities. *Technological Forecasting and Social Change*, *104*, 78–88. https://doi.org/10.1016/j.techfore.2015.12.009

Delaney, B. (2023). Designing and delivering experiential journalism tasks in online environments: A framework grounded in cognitive apprenticeship theory. *Journalism Studies*, *24*(1), 128–151. https://doi.org/10.1080/1461670X.2022.2150263

Denny, S.G., Silaigwana, B., Wassenaar, D., Bull, S., & Parker, M. (2015). Developing ethical practices for public health research data sharing in South Africa: The views and experiences from a diverse sample of research stakeholders. *Journal of Empirical Research on Human Research Ethics, 10*, 290-301.

Denzin, N. K. (2017). Critical qualitative inquiry. *Qualitative Inquiry*, *23*(1), 8–16. https://doi.org/10.1177/1077800416681864

Derakhshannia, M., Gervet, C., Hajj-Hassan, H., Laurent, A., & Martin, A. (2020). Data lake governance: Towards a systemic and natural ecosystem analogy. *Future Internet*, *12*(126), Article 8. https://doi.org/10.3390/fi12080126

Dijkers, M. P. (2019). A beginner's guide to data stewardship and data sharing. *Spinal Cord*, *57*, 169–182. https://doi.org/10.1038/s41393-018-0232-6

Dijkxhoorn, Y., Plaisier, C., Verwaart, T., Van Wagenberg, C., & Ruben, R. (2019). Trusted sorghum: Simulating interactions in the sorghum value chain in Kenya using games and agent-based modelling. *Journal of Development Effectiveness*, *11*(2), 146–164. https://doi.org/10.1080/19439342.2019.1624596

Donaldson, D. R., & Koepke, J. W. (2022). A focus groups study on data sharing and research data management. *Scientific Data*, *9*(1), Article 1. https://doi.org/10.1038/s41597-022-01428-w

Downs, R. R. (2021). Improving opportunities for new value of open data: Assessing and certifying research data repositories. *Data Science Journal*, *20*(1), 1–11. https://doi.org/10.5334/dsj-2021-001

Drossaert, F., Giannopoulos, A., & Forde, M. (2018). The detection of abandoned mineshafts by railway track bed using transmitted seismic waves using broadside shot gathers. *Construction and Building Materials*, *183*(D-16-03370), 88–94. https://doi.org/10.1016/j.conbuildmat.2018.04.105

Drury, A., De Los Rios La Serna, C. D., Bağçivan, G., Dowling, M., Kotronoulas, G., Shewbridge, A., Sheehan, S., Erdem, S., Aroyo, V., & Wiseman, T. (2023). Consensus views on an advanced breast cancer education curriculum for cancer nurses: A Delphi study. *Nurse Education Today*, *124*, 105757. https://doi.org/10.1016/j.nedt.2023.105757

Drysdale, T. (2019). Research data management in a cultural heritage organisation. *International Journal of Digital Curation*, *14*(1), Article 1. https://doi.org/10.2218/ijdc.v14i1.647

Dunning, A., van der Kuil, A., de Smaele, M., Teperek, M., & Versteeg, A. (2018). *TU Delft Research Data Framework Policy*. TU Delft.

Dwiatmoko, S., Suwitri, S., Warsono, H., Dwimawanti, I. H., & Handayani, A. (2018). Comparison study of three industrial parks in central Java Indonesia. *The Indonesian Journal of Planning and Development*, *3*(2), 67–75. https://doi.org/10.14710/ijpd.3.2.67-75

Education Development Center. (2022). *Profile of the data practitioner*. Oceans of Data Institute. http://oceansofdata.org/our-work/profile-data-practitioner

Education Portal. (2013). *Types of planning: Strategic, tactical, operational & contingency planning*. https://dcyf.worldpossible.org/rachel/modules/en-oya-static/business-101/webpages/Chapter5/2-types-of-planning-strategic-tactical-operational-contingency-planning.html

Eisma, D. V. (2015). Extent of implementation and level of students' satisfaction on student services programs of DMMMSU. *Proceedings of ISER 12th International Conference*, 1–10.

Eke, D., Ridwan, O., Borokini, F., Akintoye, S., Ochang, P., Lebura, S., Adeyeye, M., & Wale-Oshinowo, B. (2022). *Responsible data governance in Africa: Institutional gaps and capacity needs*. Centre for the Study of African Economies (CSEA).

El Miedany, Y., Abu-Zaid, M. H., El Gaafary, M., El Naby, M. M. H., Fathi, N., Saber, H. G., Hassan, W., Eissa, M., Mohannad, N., Khaled, H., Mortada, M., Nasef, S. I., Galal, S., Ghaleb, R., Tabra, S. A. A., Mohamed, S. S., Medhat, B. M., Aly, H. M., Elolemy, G., … Gadallah, N. (2021). Egyptian consensus on treat-to-target approach for osteoporosis: A clinical practice guideline from the Egyptian

Academy of Bone Health and Metabolic Bone Diseases. *Egyptian Rheumatology and Rehabilitation*, *48*(1), 5. https://doi.org/10.1186/s43166-020-00056-9

Elsayed, A. M., & Saleh, E. I. (2018). Research data management and sharing among researchers in Arab universities: An exploratory study. *IFLA Journal*, *44*(4), 281–299. https://doi.org/10.1177/0340035218785196

Else, H. (2018). Radical open-access plan could spell end to journal subscriptions. *Nature*, *561*, 17–18. https://doi.org/10.1038/d41586-018-06178-7

Emam, I., Elyasigomari, V., Matthews, A., Pavlidis, S., Rocca-Serra, P., Guitton, F., Verbeeck, D., Grainger, L., Borgogni, E., Del Giudice, G., Saqi, M., Houston, P., & Guo, Y. (2019). PlatformTM, a standards-based data custodianship platform for translational medicine research. *Scientific Data*, *6*(149), 1–13. https://doi.org/10.1038/s41597-019-0156-9

Esfandiari, E., Chudyk, A. M., Grover, S., Lau, E. Y., Hoppmann, C., Mortenson, W. B., Mulligan, K., Newton, C., Pauly, T., Pitman, B., Rush, K. L., Sakakibara, B. M., Symes, B., Tsuei, S., Petrella, R. J., & Ashe, M. C. (2023). Social Prescribing Outcomes for Trials (SPOT): Protocol for a modified Delphi study on core outcomes. *PLOS ONE*, *18*(5), e0285182. https://doi.org/10.1371/journal.pone.0285182

Etikan, İ., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1. https://doi.org/10.11648/j.ajtas.20160501.11

European Commission. Directorate General for Research & Innovation. (2016). *H2020 programme: Guidelines on FAIR data management in Horizon 2020*.

https://repository.oceanbestpractices.org/bitstream/handle/11329/1259/h2020-hi-oa-data-mgt_en.pdf?sequence=1&isAllowed=y

European Commission. Directorate General for Research and Innovation. (2019). *European Open Science Cloud (EOSC) strategic implementation plan.* Publications Office of the European Union. https://data.europa.eu/doi/10.2777/202370

Faiz, M. F., Arshad, J., Alazab, M., & Shalaginov, A. (2020). Predicting likelihood of legitimate data loss in email DLP. *Future Generation Computer Systems, 110*, 744–757. https://doi.org/10.1016/j.future.2019.11.004

Faniel, I., & Connaway, L. (2018). Librarians' perspectives on the factors influencing research data management programs. *College & Research Libraries, 79*(1). https://doi.org/10.5860/crl.79.1.100

Fanniff, A. M., York, T., & Gutierrez, R. (2023). Developing consensus for culturally informed forensic mental health assessment: Experts' opinions on best practices. *Law and Human Behavior, 47*(3), 385–402. https://doi.org/10.1037/lhb0000531

Fauzy, H., Aryza, S., & Tarigan, A. S. (2021). Implementation of IoT water saving based on smart water flow system. *INFOKUM, 10*(1), Article 1.

Fecher, B., Friesike, S., & Hebing, M. (2015). What drives academic data sharing? *PLOS ONE, 10*(2), e0118053. https://doi.org/10.1371/journal.pone.0118053

Feki, M., & Boughzala, I. (2016). Big data governance: A literature review and research agenda. *15ieme Conference Internationale de Gouvernance de l'AAIG (Association Academique Internationale de Gouvernance)*, 1–11.

Filgueiras, F., & Raymond, A. (2023). Designing governance and policy for disruptive digital technologies. *Policy Design and Practice*, *6*(1), 1–13. https://doi.org/10.1080/25741292.2022.2162241

Florida State University. (2016). *Research Data Management Policy*. Florida State University. FSU Policy 7A-26.pdf

Foster, E. D., & Deardorff, A. (2017). Open Science Framework (OSF). *Journal of the Medical Library Association*, *105*(2). https://doi.org/10.5195/JMLA.2017.88

Fürholz, A., & Jaekel, M. (2021). Data life cycle management pilot projects and implications for research data management at universities of applied sciences. *Revue Électronique Suisse de Science de l'information, Spécial DLCM*, 32–46. https://doi.org/10.21256/zhaw-23073

Gaba, J. F., Siebert, M., Dupuy, A., Moher, D., & Naudet, F. (2020). Funders' data-sharing policies in therapeutic research: A survey of commercial and non-commercial funders. *PLOS ONE*, *15*(8), e0237464. https://doi.org/10.1371/journal.pone.0237464

Garner, B. R., Gotham, H. J., Knudsen, H. K., Zulkiewicz, B. A., Tueller, S. J., Berzofsky, M., Donohoe, T., Martin, E. G., Brown, L. L., & Gordon, T. (2022). The prevalence and negative impacts of substance use disorders among people with HIV in the United States: A real-time Delphi survey of key stakeholders. *AIDS and Behavior*, *26*(4), 1183–1196. https://doi.org/10.1007/s10461-021-03473-9

Gasigwa, T., Maniraho, D. J. F., & Habimana, D. O. (2022). Effectiveness of excel software integration in teaching and learning statistics on students' performance in selected Kicukiro district upper secondary schools in Rwanda. *Journal of*

Research Innovation and Implications in Education, *6*(3), 194–200. https://jriiejournal.com/wp-content/uploads/2022/08/JRIIE-6-3-021.pdf

Ghent University. (2016). *Policy Framework on Research Data Management*. Ghent University.

Glascock, J., Darras, B. T., Crawford, T. O., Sumner, C. J., Kolb, S. J., DiDonato, C., Elsheikh, B., Howell, K., Farwell, W., Valente, M., Petrillo, M., Tingey, J., & Jarecki, J. (2023). Identifying biomarkers of spinal muscular atrophy for further development. *Journal of Neuromuscular Diseases*, *10*(5), 937–954. https://doi.org/10.3233/JND-230054

Goel, K., Emamjome, F., & ter Hofstede, A. H. M. (2021). Data governance for managing data quality in process mining. *Proceedings of the 42nd International Conference on Information Systems (ICIS 2021).*, 1–17. https://eprints.qut.edu.au/226279/1/100232966.pdf

Gonçalves, J. P. D. B., Braghetta, C. C., Alvarenga, W. D. A., Gorenstein, C., Lucchetti, G., & Vallada, H. (2023). Development of a comprehensive flourishing intervention to promote mental health using an e-Delphi technique. *Frontiers in Psychiatry*, *14*, 1064137. https://doi.org/10.3389/fpsyt.2023.1064137

Goodman, C. M. (1987). The Delphi technique: A critique. *Journal of Advanced Nursing*, *12*(6), 729–734. https://doi.org/10.1111/j.1365-2648.1987.tb01376.x

Gotay, J. A. (2020). A classical Delphi study to identify the barriers of green information technologies. *Open Science Journal*, *5*(3). https://doi.org/10.23954/osj.v5i3.2456

Goundar, S. (2012). *Chapter 3: Research methodology and research method*. ResearchGate.

Grabus, S., & Greenberg, J. (2019). The landscape of rights and licensing initiatives for data sharing. *Data Science Journal*, *18*(29), 1–11. https://doi.org/10.5334/dsj-2019-029

Gunjal, B., & Gaitanou, P. (2017). Research data management: A proposed framework to boost research in higher educational institutes. *IASSIST Quarterly*, *41*(1–4), 1–13. https://doi.org/10.29173/iq12

Gupta, N., Arora, S., & Chakravarty, R. (2021). Science mapping and visualization of research data management (RDM): Bibliometric and scientometric study. *Library Philosophy and Practice*, *6096*, 1–25.

Gupta, U., & Cannon, S. (2020a). Data governance frameworks. In U. Gupta & S. Cannon, *A practitioner's guide to data governance* (pp. 101–122). Emerald Publishing Limited. https://doi.org/10.1108/978-1-78973-567-320201005

Gupta, U., & Cannon, S. (2020b). Data strategy. In U. Gupta & S. Cannon, *A practitioner's guide to data governance* (pp. 71–99). Emerald Publishing Limited. https://doi.org/10.1108/978-1-78973-567-320201004

Hall, D. A., Smith, H., Heffernan, E., & Fackrell, K. (2018). Recruiting and retaining participants in e-Delphi surveys for core outcome set development: Evaluating the COMiT'ID study. *PLoS ONE*, *13*(7), e0201378. https://doi.org/10.1371/journal.pone.0201378

Hamad, F., Al-Fadel, M., & Al-Soub, A. (2019). Awareness of research data management services at academic libraries in Jordan: Roles, responsibilities and challenges. *New Review of Academic Librarianship*, 76–96. https://doi.org/10.1080/13614533.2019.1691027

Hannila, H., Silvola, R., Harkonen, J., & Haapasalo, H. (2019). Data-driven begins with DATA; potential of data assets. *Journal of Computer Information Systems*, 1–10. https://doi.org/10.1080/08874417.2019.1683782

Hao, Y., Piao, C., Zhao, Y., & Jiang, X. (2020). Privacy preserving government data sharing based on hyperledger blockchain. In K. M. Chao, L. Jiang, O. K. Hussain, S. P. Ma, & X. Fei (Eds.), *Advances in e-business engineering for ubiquitous computing* (Vol. 41, pp. 373–388). Springer International Publishing. https://doi.org/10.1007/978-3-030-34986-8_27

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, *32*(4), 1008-1015. https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x

Hendey, L., Gold, A., & Pettit, K. L. S. (2018). *NNIP's resource guide to data governance and security* (pp. 1–49). National Neighborhood Indicators Partnership.

Heriot-Watt University. (2015). *Research Data Management Policy*. Heriot-Watt University. (hw.ac.uk)

Higman, R., Bangert, D., & Jones, S. (2019). Three camps, one destination: The intersections of research data management, FAIR and Open. *Insights the UKSG Journal, 32*, 18. https://doi.org/10.1629/uksg.468

Hodonu-Wusu, J. O., Noorhidawati, A., & Abrizah, A. (2020). Malaysian researchers on open data: The first national survey on awareness, practices and attitudes. *Malaysian Journal of Library & Information Science, 25*(2), 1-20. https://doi.org/10.22452/mjlis.vol25no2.1

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Huang, P. B., Yang, C.-C., Inderawati, M. M. W., & Sukwadi, R. (2022). Using modified Delphi study to develop instrument for ESG implementation: A case study at an Indonesian higher education institution. *Sustainability*, *14*(19), 12623. https://doi.org/10.3390/su141912623

Huang, Y., Cox, A. M., & Sbaffi, L. (2021). Research data management policy and practice in Chinese university libraries. *Journal of the Association for Information Science and Technology*, *72*(4), 493–506.

Humphrey-Murto, S., Crew, R., Shea, B., Bartlett, S. J., March, L., Tugwell, P., Maxwell, L. J., Beaton, D., Grosskleg, S., & De Wit, M. (2019). Consensus building in OMERACT: Recommendations for use of the Delphi for core outcome set development. *The Journal of Rheumatology*, *46*(8), 1041–1046. https://doi.org/10.3899/jrheum.181094

Humphrey-Murto, S., Varpio, L., Gonsalves, C., & Wood, T. J. (2017). Using consensus group methods such as Delphi and nominal group in medical education research. *Medical Teacher*, *39*(1), 14–19. https://doi.org/10.1080/0142159X.2017.1245856

Huston, P., Edge, V., & Bernier, E. (2019). Reaping the benefits of open data in public health. *Canada Communicable Disease Report*, *45*(10), 252–256. https://doi.org/10.14745/ccdr.v45i10a01

Ibrahim, S., & Wei, T. H. (2023). Development of research management in Malaysia. In S. Kerridge, S. Poli, & M. Yang-Yoshihara (Eds.), *The Emerald handbook of*

research management and administration around the world (pp. 527–533). Emerald Publishing Limited. https://doi.org/10.1108/978-1-80382-701-820231047

Imperial College. (2022). *Research Data Management Policy*. Imperial College. (imperial.ac.uk)

Ismail, M. I., Jaafar, C. R. C., Azmi, N. A., Makhtar, M. M. Z., Samsuddin, S. F., & Abrizah, A. (2022). Eliciting researchers' behaviour as the foundation of research data management service development. *Library and Information Science Research e-Journal, 32*(1), 44–63. https://doi.org/10.32655/LIBRES.2022.1.4

Jacobsson, T. J., Hultqvist, A., García-Fernández, A., Anand, A., Al-Ashouri, A., Hagfeldt, A., Crovetto, A., Abate, A., Ricciardulli, A. G., & Vijayan, A. (2022). An open-access database and analysis tool for perovskite solar cells based on the FAIR data principles. *Nature Energy, 7,* 107–115. https://doi.org/10.1038/s41560-021-00941-3

James, L. (2013, October 3). *Defining open data*. Open Knowledge Foundation: For a fair, sustainable and open future. https://blog.okfn.org/2013/10/03/defining-open-data/

Jamiu, S. M., Abdullah, N. S., Miskon, S., & Ali, N. M. (2020). Data governance support for business intelligence in higher education: A systematic literature review. In F. Saeed, F. Mohammed, & N. Gazem (Eds.), *Emerging trends in intelligent computing and informatics* (Vol. 1073, pp. 35–44). Springer International Publishing. https://doi.org/10.1007/978-3-030-33582-3_4

Jang, K., & Kim, W.-J. (2016). Trend of research in data governance using time series analysis. *Advanced Science and Technology Letters, 139,* 146–149. http://dx.doi.org/10.14257/astl.2016.139.33

Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy Artificial Intelligence. *Government Information Quarterly*, *37*(3), 101493. https://doi.org/10.1016/j.giq.2020.101493

Janssen, M., Konopnicki, D., Snowdon, J. L., & Ojo, A. (2017). Driving public sector innovation using big and open linked data (BOLD). *Information Systems Frontiers*, *19*, 189–195. https://doi.org/10.1007/s10796-017-9746-2

Jim, C. K., & Chang, H. C. (2018). The current state of data governance in higher education. *Proceedings of the Association for Information Science and Technology*, *55*(1), 198–206. https://doi.org/10.1002/pra2.2018.14505501022

Johnston, L. R., Carlson, J., Hudson-Vitale, C., Imker, H., Kozlowski, W., Olendorf, R., Stewart, C., Blake, M., Herndon, J., McGeary, T. M., & Hull, E. (2018). Data curation network: A cross-institutional staffing model for curating research data. *International Journal of Digital Curation*, *13*(1), 125–140. https://doi.org/10.2218/ijdc.v13i1.616

Joo, S., & Peters, C. (2020). User needs assessment for research data services in a research university. *Journal of Librarianship and Information Science, 52*(3), 633–646. https://doi.org/10.1177/0961000619856073

Jünger, S., Payne, S. A., Brine, J., Radbruch, L., & Brearley, S. G. (2017). Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care:

Recommendations based on a methodological systematic review. *Palliative Medicine*, *31*(8), 684–706. https://doi.org/10.1177/0269216317690685

Jusoh, I. (2018). *Amanat menteri pendidikan tinggi 2018. Tema: Pendidikan tinggi 4.0: Ilmu, industri dan insan*. https://www.mohe.gov.my/muat-turun/teks-ucapan-dan-slaid/teks-2018/amanat-menteri-pendidikan-tinggi-2018/260-text-ucapan-amanat-menteri-pendidikan-tinggi-2018/file

Kabanda, S. M., Cengiz, N., Rajaratnam, K., Watson, B. W., Brown, Q., Esterhuize, T. M. n, & Moodley, K. (2023). Data sharing and data governance in sub-Saharan Africa: Perspectives from researchers and scientists engaged in data-intensive research. *South African Journal of Science*, *119*(5/6). https://doi.org/10.17159/sajs.2023/15129

Kabir, S. M. S. (2016). Methods of data collection. In *Basic guidelines for research: An introductory approach for all disciplines* (pp. 201–276). Book Zone Publication.

Kaewkamol, P. (2022). Data governance framework as initiative for higher educational organisation. *2022 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON)*, 175–178. https://doi.org/10.1109/ECTIDAMTNCON53731.2022.9720396

Kariotis, T., Ball, M. P., Greshake Tzovaras, B., Dennis, S., Sahama, T., Johnston, C., Almond, H., & Borda, A. (2020). Emerging health data platforms: From individual control to collective data governance. *Data & Policy*, *2*, e13. https://doi.org/10.1017/dap.2020.14

Keeney, S., Hasson, F., & McKenna, H. (2006). Consulting the oracle: Ten lessons from using the Delphi technique in nursing research. *Journal of Advanced Nursing*, *53*(2), 205–212. https://doi.org/10.1111/j.1365-2648.2006.03716.x

Khanal, R. C. (2014). Concerns and challenges of data integration from objective post-positivist approach and a subjective non-positivist interpretive approach and their validity/credibility issues. *Journal of the Institute of Engineering*, *9*(1), 115–129. https://doi.org/10.3126/jie.v9i1.10677

Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, *53*(1), 148–152. https://doi.org/10.1145/1629175.1629210

Khodyakov, D., Grant, S., Kroger, J., & Bauman, M. (2023). *RAND methodological guidance for conducting and critically appraising Delphi panels*. RAND Corporation. https://doi.org/10.7249/TLA3082-1

Kiernan, A., Moneley, D., Doyle, F., & Harkin, D. W. (2023). A modified Delphi process to form an expert consensus in Ireland on the essential information to be included in shared decision making for varicose vein surgery. *Phlebology*, *38*(4), 259–269. https://doi.org/10.1177/02683555231158284

Kilbertus, N., Gascon, A., Kusner, M., Veale, M., Gummadi, K. P., & Weller, A. (2018). Blind justice: Fairness with encrypted sensitive attributes. *Proceedings of the 35th International Conference on Machine Learning*, 1–15. https://arxiv.org/abs/1806.03281

Kim, S. (2020). Development of evaluation perspective and criteria for the DataON platform. *Journal of Information Science Theory and Practice*, *8*(2), 68–78. https://doi.org/10.1633/JISTAP.2020.8.2.6

Koltay, T. (2016). Data governance, data literacy and the management of data quality. *IFLA Journal*, *42*(4), 303–312. https://doi.org/10.1177/0340035216672238

Koltay, T. (2020). Quality of open research data: Values, convergences and governance. *Information*, *11*(175), 1–15.

Korhonen, J. J., Melleri, I., Hiekkanen, K., & Helenius, M. (2013). Designing data governance structure: An organizational perspective. *GSTF Journal on Computing*, *2*(4), 11–17.

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

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed). Sage.

Krzton, A. (2018). Supporting the proliferation of data-sharing scholars in the research ecosystem. *Journal of eScience Librarianship, 7*(2): e1145. https://doi.org/10.7191/jeslib.2018.1145

Kuzio, J., Ahmadi, M., Kim, K. C., Migaud, M. R., Wang, Y. F., & Bullock, J. (2022). Building better global data governance. *Data & Policy*, *4*(e25), 1–17. https://doi.org/10.1017/dap.2022.17

Labadie, C., & Legner, C. (2023). Building data management capabilities to address data protection regulations: Learnings from EU-GDPR. *Journal of Information Technology*, *38*(1), 16–44. https://doi.org/10.1177/02683962221141456

Lacagnina, C., David, R., Nikiforova, A., Kuusniemi, M. E., Cappiello, C., Biehlmaier, O., Wright, L., Schubert, C., Bertino, A., Thiemann, H., & Dennis, R. (2022). *Towards a data quality framework for EOSC authorship community*. EOSC Association. https://doi.org/10.5281/zenodo.7515816

Ladley, J. (2012). *Data governance: How to design, deploy and sustain an effective data governance program* (1st ed.). Morgan Kaufmann.

Lancaster, G. A., Dodd, S., & Williamson, P. R. (2004). Design and analysis of pilot studies: Recommendations for good practice. *Journal of Evaluation in Clinical Practice, 10*(2), 307–312. https://doi.org/10.1111/j..2002.384.doc.x

Lancho-Barrantes, B. S., & Cantu-Ortiz, F. J. (2021). Quantifying the publication preferences of leading research universities. *Scientometrics, 126*(3), 2269–2310. https://doi.org/10.1007/s11192-020-03790-1

Lau, H. X., Lee, S. L. C., & Ali, Y. (2021). Effectiveness of data auditing as a tool to reinforce good research data management (RDM) practice: A Singapore study. *BMC Medical Ethics, 22*(103), 1–8. https://doi.org/10.1186/s12910-021-00662-y

Lee, C. J., & Moher, D. (2017). Promote scientific integrity via journal peer review data. *Science, 357*(6348), 256–257. https://doi.org/10.1126/science.aan4141

Lee, D. J., & Stvilia, B. (2017). Practices of research data curation in institutional repositories: A qualitative view from repository staff. *PLOS ONE, 12*(3), e0173987. https://doi.org/10.1371/journal.pone.0173987

Lee, S. U., Zhu, L., & Jeffery, R. (2017). Data governance for platform ecosystems: Critical factors and the state of practice. *Proceedings of the 21st Pacific Asia*

Conference on Information Systems: Societal Transformation Through IS/IT, 1–12.

Lefebvre, A., & Spruit, M. (2021). Laboratory forensics for open science readiness: An investigative approach to research data management. *Information Systems Frontiers*, *25*, 381–399. https://doi.org/10.1007/s10796-021-10165-1

Lefebvre, A., Schermerhorn, E., & Spruit, M. (2018). *How research data management can contribute to efficient and reliable science*. 26th European Conference on Information Systems (ECIS2018), Portsmouth, UK. Scopus.

Leonelli, S. (2018). Rethinking reproducibility as a criterion for research quality. In L. Fiorito, S. Scheall, & C. E. Suprinyak (Eds.), *Research in the History of Economic Thought and Methodology* (Vol. 36, pp. 129–146). Emerald Publishing Limited. https://doi.org/10.1108/S0743-41542018000036B009

Leonelli, S., Spichtinger, D., & Prainsack, B. (2015). Sticks and carrots: Encouraging open science at its source. *Geo: Geography and Environment*, *2*(1), 12–16. https://doi.org/10.1002/geo2.2

Lim, B., Cook, K., Sutherland, D., Weathersby, A., & Tillard, G. (2022). Clinical supervision in Singapore: Allied Health professional perspectives from a two-round Delphi study. *Internet Journal of Allied Health Sciences and Practice*, *20*(1). Article 12. https://doi.org/10.46743/1540-580X/2022.2103

Lim, M., Thompson, B., D'Silva, C., Wang, G. Y., Bhatnagar, P., Palaganas, M., Reid, R., Cairney, J., Varma, D., Smith, D., & Ahmed, I. (2020). Development and reliability of an appropriateness and prioritization instrument for eye care practice:

A modified Delphi process. *Ophthalmic Epidemiology, 27*(1), 19–28. https://doi.org/10.1080/09286586.2019.1678653

Liu, G., Zotoo, I. K., & Su, W. (2020). Research data management policies in USA, UK and Australia universities: An online survey. *Malaysian Journal of Library & Information Science*, *25*(2), 21–42.

Logan, J. A. R., Hart, S. A., & Schatschneider, C. (2021). Data sharing in education science. *AERA Open, 7*(1), 1–11. https://doi.org/10.1177/23328584211006475

London School of Hygiene & Tropical Medicine. (2022). *Research Data Management Policy*. London School of Hygiene & Tropical Medicine.

Luesebrink, M., Huang, H., Bogdan, K., Salo, D., Thomas, C., West, J., & Wilson, L. B. (2014). Curation and policy issues in collaborative research data management communities: Perspectives from key stakeholders. *Proceedings of the American Society for Information Science and Technology*, *51*(1), 1–4. https://doi.org/10.1002/meet.2014.14505101024

Ma, R., & Lam, P. T. I. (2019). Investigating the barriers faced by stakeholders in open data development: A study on Hong Kong as a "smart city." *Cities, 92*, 36–46. https://doi.org/10.1016/j.cities.2019.03.009

Machimbidza, T., Mutula, S., & Ndiweni, M. (2022). Exploring the feasibility of research data management services in Zimbabwean academic libraries: The case of Zimbabwe School of Mines. *Library Philosophy and Practice (e journal),* 1–35. https://digitalcommons.unl.edu/libphilprac/6760

Macquarie University. (2021). *Research Data Management Policy*. Macquarie University. (mq.edu.au)

Madison, M. (2020). *Tools for data governance* (No. 2020-23, pp. 29-43). University of Pittsburgh School of Law. https://openjournals.nl/techreg/article/view/10993

Makhlouf-Shabou, B. (2017). Training, consulting and teaching for sustainable approach for developing research data life-cycle management expertise in Switzerland. *INFuture2017 Integrating ICT in Society*, 79–86.

Mandinach, E. B., & Gummer, E. S. (2013). A systemic view of implementing data literacy in educator preparation. *Educational Researcher*, *42*(1), 32–37. https://doi.org/10.3102/0013189X12459803

Maniam, J. N., & Singh, D. (2020). Towards data privacy and security framework in big data governance. *International Journal of Software Engineering and Computer Systems*, *6*(1), 41–51. https://doi.org/10.15282/ijsecs.6.1.2020.5.0068

Manik, L. P., Akbar, Z., Yaman, A., & Indrawati, A. (2022). Indonesian scientists' behavior relative to research data governance in preventing WMD-applicable technology transfer. *Publications*, *10*(50), 1-29. https://doi.org/10.3390/publications10040050

Mansfield-Devine, S. (2017). Data governance: Going beyond compliance. *Computer Fraud & Security*, *2017*(June), 12–15. https://doi.org/10.1016/S1361-3723(17)30052-0

Markmann, C., Spickermann, A., Von Der Gracht, H. A., & Brem, A. (2021). Improving the question formulation in Delphi-like surveys: Analysis of the effects of abstract language and amount of information on response behavior. *Futures & Foresight Science*, *3*(1), 1-20. https://doi.org/10.1002/ffo2.56

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

Marlina, E., Hidayanto, A. N., & Purwandari, B. (2022). Towards a model of research data management readiness in Indonesian context: An investigation of factors and indicators through the fuzzy Delphi method. *Library & Information Science Research*, *44*(1), 1–12. https://doi.org/10.1016/j.lisr.2022.101141

Masinde, J., Chen, J., Wambiri, D., & Mumo, A. (2021). Research librarians' experiences of research data management activities at an academic library in a developing country. *Data and Information Management*, *5*(4), 412–424. https://doi.org/10.2478/dim-2021-0002

Matthewson, P. (2019). *Research Data Governance & Materials Handling Policy*. University of New South Wales. (unsw.edu.au)

Maxwell, J., & Mittapalli, K. (2010). Realism as a stance for mixed methods research. In *SAGE handbook of mixed methods in social & behavioral research* (2nd ed., pp. 145-168). SAGE Publications, Inc., https://doi.org/10.4135/9781506335193

McDonald, J. (2018). A holistic approach to becoming a data-driven utility. In *Big data application in power systems* (pp. 3–27). Elsevier. https://doi.org/10.1016/B978-0-12-811968-6.00001-2

McKiernan, E. C. (2017). Imagining the "open" university: Sharing scholarship to improve research and education. *PLOS Biology*, *15*(10), e1002614. https://doi.org/10.1371/journal.pbio.1002614

McMillan, J. H. (2012). *Educational research: Fundamentals for the consumer* (6th ed). Pearson.

Menendez-Alonso, E. (2020). *Research Data Policy*. University of Plymouth. Research Data Policy (plymouth.ac.uk)

Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative research: A guide to design and implementation* (4th ed.). John Wiley & Sons.

Micheli, M., Ponti, M., Craglia, M., & Berti Suman, A. (2020). Emerging models of data governance in the age of datafication. *Big Data & Society*, *7*(2), 205395172094808. https://doi.org/10.1177/2053951720948087

Ministry of Education Malaysia. (2015). *Malaysia Education Blueprint 2015-2025 (Higher Education).*

Mohamad Hashim, H. N. (2019). Developing a model guidelines addressing legal impediments to open access to publicly funded research data in Malaysia. *Data Science Journal*, *18*(27), 1–17. https://doi.org/10.5334/dsj-2019-027

Mohammed, M. S., & Ibrahim, R. (2019). Challenges and practices of research data management in selected Iraq universities. *DESIDOC: Journal of Library & Information Technology*, *39*(6), 308–314. https://doi.org/10.14429/djlit.39.06.14443

MOHE. (2024). *Ministry of Higher Education*. https://www.mohe.gov.my/en/home

MOSP. (2020). *Malaysia Open Science Platform (MOSP): Charting the way forward for Malaysia open science*. https://www.akademisains.gov.my/mosp/

MOSP. (2024). *Malaysia Open Science Platform (MOSP)*. https://mosp.gov.my/stewards

MOSTI. (2013). *National Policy on Science, Technology & Innovation (NPSTI) 2013-2020*. https://www.pmo.gov.my/2019/07/national-policy-on-science-technology-innovation-npsti/

MOSTI. (2021). *Dasar Sains, Teknologi dan Inovasi Negara 2021-2030*. https://www.mosti.gov.my/web/dasar-halatuju/

Mozilla. (2020). *Terminology*. Mozilla data documentation. https://docs.telemetry.mozilla.org/concepts/terminology.html#data-practitioner

Münch, C., von, der G. H. A., & Hartmann, E. (2021). The future role of reverse logistics as a tool for sustainability in food supply chains: A Delphi-based scenario study. *Supply Chain Management: An International Journal*, *28*(2), 262–283. https://doi.org/10.1108/SCM-06-2021-0291

Murray-Rust, P. (2008). Open data in science. *Serials Review*, *34*(1), 52–64.

Mustaffa, E. N., & Ghani, M. F. A. (2021). Development of a self: Leadership model using modified Delphi technique. *Journal of Educational Administration Research and Review*, *5*(1), 10.

Myers, M. D. (1997). Qualitative research in Information Systems. *MIS Quarterly*, *21*(2), 241. https://doi.org/10.2307/249422

Nanyang Technological University. (2019). *NTU Research Data Policy*. Nanyang Technological University. NTU Research Data Policy | Research | NTU Singapore

Napis, S., Michael, V. A., Jusoh, Y. Y., Abdullah, R., Sidi, F., Ishak, I., & Marhaban, M. H. (2019). A preliminary study on the development of research data management

(RDM) policy. *International Journal of Advanced Science and Technology*, *28*(2), 8.

Naranjo, F., Menor, L. J., & Johnson, P. F. (2023). Lean supply chain management: A contextual contingent reconceptualization and Delphi method study. *International Journal of Operations & Production Management*, *43*(9), 1456-1480. https://doi.org/10.1108/IJOPM-07-2022-0436

Nasir, N. N. I., Radzuan, S. N., Azhami, B. A., & Hamidon, H. (2023). Cyber security in higher education: Problem and solution. In M. R. Mohamad Rosman, I. Arshad, N. Abdullah, N. E. Abd Aziz, G. Osman, A. Shukry, N. R. Alias, N. Rosli, F. Fadzil, M. S. Md Saleh, M. Z. Mohd Zawawi, & M. Rachmawati (Eds.), *Proceedings of Global Symposium on Information and Social Sciences (GSISS) 2023* (pp. 128–137). https://doi.org/DOI: 10.5281/zenodo.8182762

National Academies of Sciences, Engineering, and Medicine. (2018). *Open science by design: Realizing a vision for 21st century research* (p. 25116). National Academies Press. https://doi.org/10.17226/25116

National Institutes of Health. (2023). *NIH Policy for Data Management and Sharing*. National Institutes of Health.

National Science Council. (2020). *The Malaysian code of responsible conduct in research 2nd edition*. Academy of Sciences Malaysia. https://www.akademisains.gov.my/asm-publication/the-malaysian-code-of-responsible-conduct-in-research-2nd-edition/

Naughton, L., & Kernohan, D. (2016). Making sense of journal research data policies. *Insights*, *29*(1), 84–89. Scopus. https://doi.org/10.1629/uksg.284

Negev, M., Magal, T., & Kaphzan, H. (2023). Attitudes of psychiatrists toward telepsychiatry: A policy Delphi study. *Digital Health*, *9*, 20552076231177132. https://doi.org/10.1177/20552076231177132

Nemati, B., Zandi, S., Aminnejad, B., Davarazar, M., Sheikhnejad, Y., Jahanianfard, D., & Mostafaie, A. (2020). Building information modelling execution in administrative and commercial spaces in Iran: A Fuzzy-Delphi criteria prioritization. *Journal of Settlements and Spatial Planning, Special Issue*(6), 17–27. https://doi.org/10.24193/JSSPSI.2020.6.03

Ng'eno, E., & Mutula, S. (2018). Research data management (RDM) in agricultural research institutes: A literature review. *Journal of Humanities and Social Sciences*, *10*(1), 28–50.

Nie, H., Luo, P., & Fu, P. (2021). Research data management implementation at Peking University library: Foster and promote open science and open data. *Data Intelligence*, *3*(1), 189–204. https://doi.org/10.1162/dint_a_00088

Nielsen, O. B. (2017). A comprehensive review of data governance literature. *Selected Papers of the IRIS*, *8*, 120–133.

Norbib, K., & Abu Bakar, N. A. (2021). Data governance model for the Ministry of Education Malaysia using Enterprise Architecture approach. *Open International Journal of Informatics*, *9*(3), 1–15.

Northumbria University. (2019). *Research Data Management Policy*. Northumbria University. (azureedge.net)

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe,

A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., … Yarkoni, T. (2015). Promoting an open research culture: Author guidelines for journals could help to promote transparency, openness, and reproducibility. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

Nyblom P, Wangen G, Kianpour M et al. (2020). The root causes of compromised accounts at the university. *Proceedings of the 6th International Conference on Information Systems Security and Privacy, Valletta, Malta*. https://doi.org/10.5220/0008972305400551

Office for Human Research Protections (OHRP). (1979). *The Belmont Report* [Text]. https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html

Olesen, A. P., Amin, L., & Mahadi, Z. (2018). Researchers experience of misconduct in research in Malaysian higher education institutions. *Accountability in Research*, *25*(3), 125–141. https://doi.org/10.1080/08989621.2018.1429925

Olsen, A. A., Wolcott, M. D., Haines, S. T., Janke, K. K., & McLaughlin, J. E. (2021). How to use the Delphi method to aid in decision making and build consensus in pharmacy education. *Currents in Pharmacy Teaching and Learning*, *13*(10), 1376–1385. https://doi.org/10.1016/j.cptl.2021.07.018

Omar, A., & Almaghthawi, A. (2020). Towards an integrated model of data governance and integration for the implementation of digital transformation processes in the Saudi universities. *International Journal of Advanced Computer Science and Applications*, *11*(8), 588–593. https://doi.org/10.14569/IJACSA.2020.0110873

Open Knowledge Foundation. (2023). *Open Knowledge Foundation: For a fair, sustainable and open future*. https://okfn.org

Open University. (2020). *Research Data Management Policy*. Open University. https://www.open.ac.uk

Otto, B. (2011). Organizing data governance: Findings from the telecommunications industry and consequences for large service providers. *Communications of the Association for Information Systems, 29*. Article 3. https://doi.org/10.17705/1CAIS.02903

Otto, B. (2013). On the evolution of data governance in firms: The case of Johnson & Johnson consumer products North America. In S. Sadiq (Ed.), *Handbook of data quality* (pp. 93–118). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-36257-6_5

Palsdottir, A. (2021). Data literacy and management of research data: A prerequisite for the sharing of research data. *Aslib Journal of Information Management, 73*(2), 322–341. https://doi.org/10.1108/AJIM-04-2020-0110

Panhwar, D. A. H., Ansari, D., & Shah, A. (2017). Post-positivism: An effective paradigm for social and educational research. *International Research Journal Arts & Humanities (IRJAH), 45,* 253–260.

Parmiggiani, E., & Grisot, M. (2020). Data curation as governance practice. *Scandinavian Journal of Information Systems, 32*(1), 3–38.

Paskaleva, K., Evans, J., Martin, C., Linjordet, T., Yang, D., & Karvonen, A. (2017). Data governance in the sustainable smart city. *Informatics, 4*(4), 41. https://doi.org/10.3390/informatics4040041

Patterton, L., Bothma, T. J. D., & van Deventer, M. J. (2018). From planning to practice: An action plan for the implementation of research data management services in resource-constrained institutions. *South African Journal of Libraries & Information Science, 84*(2), 14-26. https://doi.org/10.7553/84-2-1761

Peng, G., Privette, J. L., Tilmes, C., Bristol, S., Maycock, T., Bates, J. J., Hausman, S., Brown, O., & Kearns, E. J. (2018). A conceptual enterprise framework for managing scientific data stewardship. *Data Science Journal*, *17*, 1–17. https://doi.org/10.5334/dsj-2018-015

Peng, G., Ritchey, N. A., Casey, K. S., Kearns, E. J., Privette, J. L., Saunders, D., Jones, P., Maycock, T., & Ansari, S. (2016). Scientific stewardship in the open data and big data era: Roles and responsibilities of stewards and other major product stakeholders. *D-Lib Magazine*, *22*(5/6), 1–21.

Perrier, L., Blondal, E., & MacDonald, H. (2018). Exploring the experiences of academic libraries with research data management: A meta-ethnographic analysis of qualitative studies. *Library & Information Science Research*, *40*(3/4), 173-183. https://doi.org/10.1016/j.lisr.2018.08.002

Petersen, K. D., Ratcliffe, J., Chen, G., Serles, D., Frøsig, C. S., & Olesen, A. V. (2019). The construct validity of the Child Health Utility 9D-DK instrument. *Health and Quality of Life Outcomes*, *17*(1), 187. https://doi.org/10.1186/s12955-019-1256-0

Peukert, C., Bechtold, S., Batikas, M., & Kretschmer, T. (2022). Regulatory spillovers and data governance: Evidence from the GDPR. *Marketing Science*, *41*(4), 318–340. https://doi.org/10.1287/mksc.2021.1339

Pinfield S., Cox, A.M., & Smith J. (2014). Research data management and libraries: Relationships, activities, drivers and influences. *PLoS ONE 9*(12), e114734. https://doi.org/10.1371/journal.pone.0114734

Piracha, H. A., & Ameen, K. (2019). Policy and planning of research data management in university libraries of Pakistan. *Collection and Curation, 38*(2), 39–44. https://doi.org/10.1108/CC-08-2018-0019

Pisani, E., & AbouZahr, C. (2010). Sharing health data: Good intentions are not enough. *Bulletin of the World Health Organization, 88,* 462–466. https://doi.org/10.2471/BLT.09.074393

Pomykalski, James J. (2020). Privacy considerations throughout the data life cycle. *Journal of Information Systems Applied Research, 13*(1). https://jisar.org/2020-13/n1/JISARv13n1.pdf#page=40

Potter, W. J., & Levine-Donnerstein, D. (1999). Rethinking validity and reliability in content analysis. *Journal of Applied Communication Research, 27*(3), 258–284. https://doi.org/10.1080/00909889909365539

Prayag, G. (2023). Post-positivism and mixed methods research in rural tourism. In H. Mair, *Handbook on tourism and rural community development* (pp. 62–75). Edward Elgar Publishing.

Qi, X. (2021). Research on enterprise data governance based on knowledge map. In: MacIntyre, J., Zhao, J., Ma, X. (eds). *The 2020 International Conference on Machine Learning and Big Data Analytics for IoT Security and Privacy*. SPIOT 2020. Advances in Intelligent Systems and Computing, vol 1283. Springer, Cham. https://doi.org/10.1007/978-3-030-62746-1_85

QS World University Rankings. (2022). *QS world university rankings 2022*. Top Universities. https://www.topuniversities.com/university-rankings/world-university-rankings/2022

Rachmat, D. F., Mulyadi, D., Abdullah, S., Putrianti, S. D., & Nurliawati, R. (2023). Analysis of the governance of One Data Indonesia (SDI) at the Communication and Information Office of Sukabumi City. In A. Rahman, S. Dwiputrianti, & M. N. Afandi (Eds.), *Proceedings of the Fourth International Conference on Administrative Science (ICAS 2022)* (Vol. 776, pp. 42–60). Atlantis Press SARL. https://doi.org/10.2991/978-2-38476-104-3_6

Rajoo, S., Magalingam, P., Samy, G. N., Maarop, N., Idris, N. B., Shanmugam, B., & Perumal, S. (2020). Detecting false messages in the smartphone fault reporting system. In F. Saeed, F. Mohammed, & N. Gazem (Eds.), *Emerging trends in Intelligent Computing and Informatics: Data Science, Intelligent Information Systems and Smart Computing* (Vol. 1073, pp. 759–768). Springer International Publishing. https://doi.org/10.1007/978-3-030-33582-3

Ranathunga, P. A. I., & Wickramarachchi, R. (2021). Designing a data governance model to implement GDPR in Sri Lankan enterprises. *Proceedings of the International Conference on Industrial Engineering and Operations Management Monterrey*, 920–930.

Redkina, N. S. (2019). Current trends in research data management. *Scientific and Technical Information Processing*, *46*(2), 53–58. https://doi.org/10.3103/S0147688219020035

Reis, J. R., Viterbo, J., & Bernardini, F. (2018). A rationale for data governance as an approach to tackle recurrent drawbacks in open data portals. *Proceedings of the*

19th Annual International Conference on Digital Government Research: Governance in the Data Age, 1–9. https://doi.org/10.1145/3209281.3209354

Rosenbaum, S. (2020). Data governance and stewardship: Designing data stewardship entities and advancing data access. *Health Services Research*, *45*(Special issue), 1442–1455. https://doi.org/10.1111/j.1475-6773.2010.01140.x

Rousi, A. M. (2022). Using current research information systems to investigate data acquisition and data sharing practices of computer scientists. *Journal of Librarianship and Information Science*, 1–13. https://doi.org/10.1177/09610006221093049

Rousi, A. M., & Laakso, M. (2020). Journal research data sharing policies: A study of highly-cited journals in neuroscience, physics, and operations research. *Scientometrics*, *124*(1), 131–152. Scopus.

Rowe, G., & Wright, G. (2011). The Delphi technique: Past, present, and future prospects: Introduction to the special issue. *Technological Forecasting and Social Change*, *78*(9), 1487–1490. https://doi.org/10.1016/j.techfore.2011.09.002

Royal Society. (2012). *Science as an open enterprise*. The Royal Society. https://royalsociety.org/~/media/Royal_Society_Content/policy/projects/sape/2012-06-20-SAOE.pdf

Ryan, A. B. (2006). Post-positivist approaches to research. In *Researching and writing your thesis: A guide for postgraduate students* (pp. 12–26). MACE: Maynooth Adult and Community Education.

Sadler, P. M., & Hammerman, J. K. (1999). Employing quantitative models of a qualitative admissions process: Uncovering hidden rules, saving time, and reducing bias. *College and University*, *74*(3), 8–24.

Saeed, S., & Ali, P. M. N. (2019). Research data management and data sharing among research scholars of Life Sciences and Social Sciences. *DESIDOC Journal of Library & Information Technology*, *39*(06), 290–299. https://doi.org/10.14429/djlit.39.06.14997

Saihi, A., Ben-Daya, M., & As'ad, R. (2023). Underpinning success factors of maintenance digital transformation: A hybrid reactive Delphi approach. *International Journal of Production Economics*, *255*, 108701. https://doi.org/10.1016/j.ijpe.2022.108701

Schöpfel, J. (2019). *Towards a culture of data: Ten proposals for a research data policy in Social Sciences and Humanities* (hal-02088501; p. 14). HAL.

Schöpfel, J., & Azeroual, O. (2021). Rewarding research data management. *Companion Proceedings of the Web Conference 2021*, 446–450. https://doi.org/10.1145/3442442.3451367

Schöpfel, J., Prost, H., & Rebouillat, V. (2017). Research data in current research information systems. *Procedia Computer Science*, *106*, 305–320. https://doi.org/10.1016/j.procs.2017.03.030

Schropp, L., Cats, R. B., de Kleijn, R. J. C. M. F., van Hattum, E. S., Middeldorp, S., Nijkeuter, M., Westerink, J., Petri, B. J., de Borst, G. J.; Upper Extremity PTS group. (2023). The upper extremity postthrombotic syndrome score: An international Delphi consensus study to determine the score's functional disability

component. *Research and Practice in Thrombosis and Haemostasis*, *7*(1), 100051. https://doi.org/10.1016/j.rpth.2023.100051

Scopus. (2022). *Research and innovation research rankings 2022*. https://www.scimagoir.com/rankings.php?ranking=Research

Seuring, S., Brandenburg, M., Sauer, P. C., Schünemann, D. S., Warasthe, R., Aman, S., Qian, C., Petljak, K., Neutzling, D. M., Land, A., & Khalid, R. U. (2022). Comparing regions globally: Impacts of COVID-19 on supply chains – a Delphi study. *International Journal of Operations & Production Management*, *42*(8), 1077–1108. https://doi.org/10.1108/IJOPM-10-2021-0675

Sforzini, L., Worrell, C., Kose, M., Anderson, I. M., Aouizerate, B., Arolt, V., Bauer, M., Baune, B. T., Blier, P., Cleare, A. J., Cowen, P. J., Dinan, T. G., Fagiolini, A., Ferrier, I. N., Hegerl, U., Krystal, A. D., Leboyer, M., McAllister-Williams, R. H., McIntyre, R. S., … Pariante, C. M. (2022). A Delphi-method-based consensus guideline for definition of treatment-resistant depression for clinical trials. *Molecular Psychiatry*, *27*(3), 1286–1299. https://doi.org/10.1038/s41380-021-01381-x

Sharif, N., Ritter, W., Davidson, R. L., & Edmunds, S. C. (2018). An open science "state of the art" for Hong Kong: Making open research data available to support Hong Kong Innovation Policy. *Journal of Contemporary Eastern Asia*, *17*(2), 200–221. https://doi.org/10.17477/JCEA.2018.17.2.200

Shitole, A., Dhapte, O., Perzade, E., Gaikwad, S., & Patil, A. (2019). REDSC: Reliability of data sharing in cloud. *International Research Journal of Engineering and Technology (IRJET)*, *6*(3), 6956–6960.

Shmagun, H., Shim, J., Kim, J., Choi, K.-N., & Oppenheim, C. (2023). Identifying key factors and actions: Initial steps in the open science policy design and implementation process. *Journal of Information Science*, 01655515231205496. https://doi.org/10.1177/01655515231205496

Singeh, F. W., Abrizah, A., & Karim, N. H. A. (2013a). Malaysian authors' acceptance to self-archive in institutional repositories: Towards a unified view. *The Electronic Library*, *31*(2), 188–207. https://doi.org/10.1108/02640471311312375

Singeh, F. W., Abrizah, A., & Karim, N. H. A. (2013b). What inhibits authors to self-archive in open access repositories? A Malaysian case. *Information Development*, *29*(1), 24–35. https://doi.org/10.1177/0266666912450450

Singh, N. K., Monu, H., & Dhingra, N. (2018). Research data management policy and institutional framework. *2018 5th International Symposium on Emerging Trends and Technologies in Libraries and Information Services (ETTLIS)*, 111–115. https://doi.org/10.1109/ETTLIS.2018.8485259

Singh, R. K., Agrawal, S., Sahu, A., & Kazancoglu, Y. (2021). Strategic issues of big data analytics applications for managing health-care sector: A systematic literature review and future research agenda. *The TQM Journal*, *35*(1), 262-291. https://doi.org/10.1108/TQM-02-2021-0051

Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The Delphi method for graduate research. *Journal of Information Technology Education: Research*, *6*, 001–021. https://doi.org/10.28945/199

Solomonides, A. (2019). Research data governance, roles, and infrastructure. In R. Richesson & J. Andrews (Eds.), *Clinical research informatics* (pp. 291–310). Springer. https://doi.org/10.1007/978-3-319-98779-8_14

Stanford Professionals. (2022). *Data practitioners*. Communities of Practice. https://cop.stanford.edu/community/data-practitioners

Starkbaum, J., & Felt, U. (2019). Negotiating the reuse of health-data: Research, big data, and the European General Data Protection Regulation. *Big Data & Society*, *6*(2), 1–12. https://doi.org/10.1177/2053951719862594

Steinerová, J., & Ondrišová, M. (2019). Research data literacy perception and practices in the information environment. In S. Kurbanoğlu, S. Špiranec, Y. Ünal, J. Boustany, M. L. Huotari, E. Grassian, D. Mizrachi, & L. Roy (Eds.), *Information literacy in everyday life* (Vol. 989, pp. 545–555). Springer International Publishing. https://doi.org/10.1007/978-3-030-13472-3_51

Sukmasetya, P., Santoso, H. B., & Sensuse, D. I. (2018). Current e-government public service on user experience perspective in Indonesia. *2018 International Conference on Information Technology Systems and Innovation (ICITSI)*, 159–164. https://doi.org/10.1109/ICITSI.2018.8695962

Sung, U. L., Liming, Z., & Jeffery, R. (2019). Data governance decisions for platform ecosystems. *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 6377–6386. https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1770&context=hicss-52

Suzuki, M. Y., Ohnuki, Y., & Takeshita, K. (2023). Genetic data governance in Japanese hospitals. *Asian Bioethics Review*, *15*(4), 377–395. https://doi.org/10.1007/s41649-023-00242-9

Szkuta, K., & Osimo, D. (2016). Rebooting science? Implications of science 2.0 main trends for scientific method and research institutions. *Foresight*, *18*(3), 204–223. https://doi.org/10.1108/FS-06-2014-0040

Taitingfong, R., Martinez, A, Hudson, M., Lovett, R., Maher, B., Prehn, J., Rowee, R. K., Boileau, K., Franks, A., Khan, S., Walker, J. D., & Carroll, S. R. (2024). Aligning policy and practice to implement CARE with FAIR through Indigenous Peoples' protocols. *Acta Borealia*, *41*(2), 80–90. https://doi.org/10.1080/08003831.2024.2410112

Tallon, P. P., Short, J. E., & Harkins, M. W. (2013). The evolution of information governance at Intel. *MIS Quarterly Executive*, *12*(4), 189–198.

Tamošaitienė, J., Khosravi, M., Cristofaro, M., Chan, D. W. M., & Sarvari, H. (2021). Identification and prioritization of critical risk factors of commercial and recreational complex building projects: A Delphi study using the TOPSIS method. *Applied Sciences*, *11*(17), 7906. https://doi.org/10.3390/app11177906

Tan, G. K. S., & Lim, S. S. (2022). Communicative strategies for building public confidence in data governance: Analyzing Singapore's COVID-19 contact-tracing initiatives. *Big Data & Society*, *9*(January–June), 1–5. https://doi.org/10.1177/20539517221104086

Tang, R., & Hu, Z. (2019). Providing research data management (RDM) services in libraries: Preparedness, roles, challenges, and training for RDM practice. *Data*

and Information Management, *3*(2), 84–101. https://doi.org/10.2478/dim-2019-0009

Tangalakis, K., Lexis, L., Hryciw, D. H., Towstoless, M., Bakker, A. J., Beckett, E., Brown, D., Cameron, M., Choate, J., Chopin, L., Cooke, M. B., Douglas, T., Estaphan, S., Etherington, S., Gaganis, V., Moorhouse, A., Moro, C., Paravicini, T., Perry, B., … Hayes, A. (2023). Establishing consensus for the core concepts of physiology in the Australian higher education context using the Delphi method. *Advances in Physiology Education*, *47*(3), 419–426. https://doi.org/10.1152/advan.00140.2022

Tauber, A. (2022). Chapter 9. Outline of a Post-Positivist Philosophy of Science. In *The triumph of uncertainty: Science and self in the postmodern age* (pp. 193-210). Budapest, Hungary: Central European University Press. https://doi.org/10.1515/9789633865828-012

Taylor, A., Tallon, D., Kessler, D., Peters, T. J., Shafran, R., Williams, C., & Wiles, N. (2020). An expert consensus on the most effective components of cognitive behavioural therapy for adults with depression: A modified Delphi study. *Cognitive Behaviour Therapy*, *49*(3), 242–255. https://doi.org/10.1080/16506073.2019.1641146

Tenopir, C., Rice, N. M., Allard, S., Baird, L., Borycz, J., Christian, L., Grant, B., Olendorf, R., & Sandusky, R. J. (2020). Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. *PLOS ONE*, *15*(3), 1–26. https://doi.org/10.1371/journal.pone.0229003

The Concordat Working Group. (2016). *Concordat on open research data*. Retrieved March 5, 2024 at https://www.ukri.org/wp-content/uploads/2020/10/UKRI-020920-ConcordatonOpenResearchData.pdf

Thompson, N., Ravindran, R., & Nicosia, S. (2015). Government data does not mean data governance: Lessons learned from a public sector application audit. *Government Information Quarterly*, *32*(3), 316–322. https://doi.org/10.1016/j.giq.2015.05.001

Thorat, R., & Van Brakel, R. (2019). Brief communication: The need for a national-level working group for higher education research data in the Netherlands. *Knowledge Organization*, *46*(5), 380–386. Scopus. https://doi.org/10.5771/0943-7444-2019-5-380

Thursby, J. G., Haeussler, C., Thursby, M. C., & Jiang, L. (2018). Prepublication disclosure of scientific results: Norms, competition, and commercial orientation. *Science Advances*, *4*(5), 1–14. https://doi.org/10.1126/sciadv.aar2133

Tiffin, N., George, A., & LeFevre, A. E. (2019). How to use relevant data for maximal benefit with minimal risk: Digital health data governance to protect vulnerable populations in low-income and middle-income countries. *BMJ Global Health*, *4*, 1–9. https://doi.org/10.1136/bmjgh-2019-001395

Tilly, R., Posegga, O., Fischbach, K., & Schoder , D. (2017). Towards a conceptualization of data and information quality in Social Information Systems. *Business & Information Systems Engineering*, 59, 3–21. https://doi.org/10.1007/s12599-016-0459-8

Tripathi, M., Chand, M., Sonkar, S. K., & Jeevan, V. K. J. (2017a). A brief assessment of researchers' perceptions towards research data in India. *IFLA Journal*, 1–18. https://doi.org/10.1177/0340035216686984

Tripathi, M., Shukla, A., & Sonkar, S. K. (2017b). Research data management practices in university libraries: A study. *DESIDOC Journal of Library & Information Technology*, *37*(6), 417. https://doi.org/10.14429/djlit.37.11336

Tu, Z., & Shen, J. (2023). Value of open research data: A systematic evaluation framework based on multi-stakeholder survey. *Library & Information Science Research*, *45*(4), 101269. https://doi.org/10.1016/j.lisr.2023.101269

Tu, Z., & Yang, Z. (2020). "3W" model in research data services: A case study of national research data centers in China. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, 495-496. https://doi.org/10.1145/3383583.3398580

Tuesta, E. F., Garcia-Zorita, C., Ayllon, R. R., & Sanz-Casado, E. (2019). Does a country/region's economic status affect its universities' presence in international rankings? *Journal of Data and Information Science*, *4*(2), 56–78. https://doi.org/10.2478/jdis-2019-0009

Turkyilmaz-van der Velden, Y., Dintzner, N., & Teperek, M. (2020). Reproducibility starts from you today. *Patterns*, *1*(6), 100099. https://doi.org/10.1016/j.patter.2020.100099

UK Data Archive. (2018). Research data lifecycle. Retrieved from http://data-archive.ac.uk/create-manage/life-cycle

Ulster University. (n.d.). *Research Data Management Policy*. Ulster University. Research Data Management Policy (ulster.ac.uk)

University College Cork. (2016). *Research Data Management Policy*. University College Cork. ResearchDataManagementPolicy.pdf (ucc.ie)

University College Dublin. (2020). *Research Data Management Policy*. University College Dublin.

University Council Committee. (2018). *Scientific Research Data Policy*. Boston University. Scientific Research Data Policy | Policies (bu.edu)

University of Birmingham. (2018). *Research Data Management Policy*. University of Birmingham. Research Data Management Policy (birmingham.ac.uk)

University of Cambridge. (2021). *Research Data Management Policy Framework.* University of Cambridge. University of Cambridge Research Data Management Policy Framework | Research Data Management

University of Cape Town. (2018). *Research Data Management Policy*. University of Cape Town. TGO_Policy_Research_Data_Management_2018.pdf (uct.ac.za)

University of Edinburgh. (2022). *Research Data Management Policy*. University of Edinburgh. Research Data Management Policy | The University of Edinburgh

University of Exeter. (2017). *Open Access Research and Research Data Management Policy*. University of Exeter. OA_RDM_Policy_Final.pdf (exeter.ac.uk)

University of Leiden. (2016). *Research Data Management Regulations*. University of Leiden. (universiteitleiden.nl)

University of Liverpool. (2019). *Research Data Management Policy*. University of Liverpool. researchdatamanagementpolicy.pdf (liverpool.ac.uk)

University of Melbourne. (2022). *Research Data Management Policy*. University of Melbourne. Research Data Management Policy (unimelb.edu.au)

University of Notre Dame. (2019). *Research Data Management Policy*. University of Notre Dame. Policy-Research-Data-Management.pdf (notredame.edu.au)

University of Oxford. (n.d.). *Policy on the Management of Data Supporting Research Outputs*. University of Oxford. Policy on the Management of Data Supporting Research Outputs | Research Data Oxford

University of Portsmouth. (2021). *Research Data Management Policy*. University of Portsmouth. (port.ac.uk)

University of Pretoria. (2017). *Research Data Management Policy*. University of Pretoria. (up.ac.za)

University of Southampton. (2019). *Research Data Management Policy*. University of Southampton. Research Data Management Policy.pdf (southampton.ac.uk)

University of Surrey. (2021). *Research Data Management Policy*. University of Surrey. Research Data Management Policy (surrey.ac.uk)

University of Twente. (2018). *Research Data Management Policy*. University of Twente. Research Data Management policy (utwente.nl)

University of Wollongong. (2020). *Research Data Management Policy*. University of Wollongong. <NAME> POLICY (uow.edu.au)

Vaismoradi, M., Turunen, H., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing & Health Sciences*, *15*(3), 398–405. https://doi.org/10.1111/nhs.12048

Van Mil, J. W. F., & Henman, M. (2016). Terminology, the importance of defining. *International Journal of Clinical Pharmacy, 38*, 709–713. https://doi.org/10.1007/s11096-016-0294-5

Van Zeeland, H., & Ringersma, J. (2017). The development of a research data policy at Wageningen University & Research: Best practices as a framework. *LIBER Quarterly*, *27*(1), 153–170. Scopus. https://doi.org/10.18352/lq.10215

Vanclay, F. (2018). *Research Data Management Policy and Procedure*. University of Groningen.

Vanderbilt University. (2021). *Research Data Governance Framework*. Vanderbilt University. Research-Data-Governance-Framework-Final-June-13-2021.pdf (vanderbilt.edu)

Vasilevsky, N. A., Minnier, J., Haendel, M. A., & Champieux, R. E. (2017). Reproducible and reusable research: Are journal data sharing policies meeting the mark? *PeerJ*, *5*, e3208. https://doi.org/10.7717/peerj.3208

Verbaan, E., & Cox, A. M. (2014). Occupational sub-cultures, jurisdictional struggle and third space: Theorising professional service responses to research data management. *Journal of Academic Librarianship, 40*(3-4), 211-219. https://doi.org/10.1016/j.acalib.20

Vericon. (2023). *2023 Data Breach Investigations Report*. https://www.verizon.com/business/resources/T68a/reports/2023-data-breach-investigations-report-dbir.pdf

Vernon, M. M., Danley, C. M., & Yang, F. M. (2021). Developing a measure of innovation from research in higher education data. *Scientometrics*, *126*(5), 3919–3928. https://doi.org/10.1007/s11192-021-03916-z

Vicente-Saez, R., & Martinez-Fuentes, C. (2018). Open science now: A systematic literature review for an integrated definition. *Journal of Business Research*, *88*(July), 428–436. https://doi.org/10.1016/j.jbusres.2017.12.043

Vilar, P., & Zabukovec, V. (2019). Research data management and research data literacy in Slovenian science. *Journal of Documentation*, *75*(1), 24–43. https://doi.org/10.1108/JD-03-2018-0042

Vilminko-Heikkinen, R., & Pekkola, S. (2019). Changes in roles, responsibilities and ownership in organizing master data management. *International Journal of Information Management*, *47*, 76–87. https://doi.org/10.1016/j.ijinfomgt.2018.12.017

Vreugdenhil, J., Döpp, D., Custers, E. J. F. M., Reinders, M. E., Dobber, J., & Kusukar, R. A. (2022). Illness scripts in nursing: Directed content analysis. *Journal of Advanced Nursing*, *78*(1), 201–210. https://doi.org/10.1111/jan.15011

Wagner, M., Henzen, C., & Müller-Pfefferkorn, R. (2021). A research data infrastructure component for the automated metadata and data quality extraction to foster the provision of fair data in Earth System Sciences. *AGILE: GIScience Series*, *2*, 1–7. https://doi.org/10.5194/agile-giss-2-41-2021

Wang, J., Li, Y., Song, W., & Li, A. (2018). Research on the theory and method of grid data asset management. *Procedia Computer Science*, *139*, 440–447. https://doi.org/10.1016/j.procs.2018.10.258

Wang, W., Zhang, X., Shi, Y., Gao, F., & Xu, W. (2021). Exploration and analysis of university scientific research data management strategies under big data environment. *Journal of Physics: Conference Series, 1881*, 1–5. https://doi.org/10.1088/1742-6596/1881/3/032059

Wardlow, L., Leff, B., Biese, K., Roberts, C., Archbald-Pannone, L., Ritchie, C., DeCherrie, L. V., Sikka, N., Gillespie, S. M., & The Collaborative for Telehealth and Aging. (2023). Development of telehealth principles and guidelines for older adults: A modified Delphi approach. *Journal of the American Geriatrics Society*, *71*(2), 371–382. https://doi.org/10.1111/jgs.18123

Weber, K., Otto, B., & Österle, H. (2009). One size does not fit all: A contingency approach to data governance. *ACM Journal of Data and Information Quality*, *1*(1), 1–27. http://doi.acm.org/10.1145/1515693.1515696

Wende, K., & Otto, B. (2007). A contingency approach to data governance. *Proceedings of 12th International Conference on Information Quality*, 15.

White, M. D., & Marsh, E. E. (2006). Content analysis: A flexible methodology. *Library Trends*, *55*(1), 22–45. https://doi.org/10.1353/lib.2006.0053

Whitehead, D., & Day, J. (2016). Mixed-methods research. In Z. Schneider, D. Whitehead, G. LoBiondo-Wood, & J. Haber (Eds.), *Nursing & midwifery research: Methods and appraisal for evidence-based practice* (pp. 237–251). Elsevier.

Wibisono, A., Sammon, D., & Heavin, C. (2022). Plausible pictures for data governance: A narrative network approach. *ECIS 2022 Research Papers*, 1–15. https://aisel.aisnet.org/ecis2022_rp/110

Wiley, C., & Kerby, E. E. (2018). Managing research data: Graduate student and postdoctoral researcher perspectives. *Issues in Science and Technology Librarianship*, *89*. https://doi.org/10.29173/istl1725

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, *3*(1), 1–9. https://doi.org/10.1038/sdata.2016.18

Willaert, T., Cottyn, J., Kenens, U., Vandendriessche, T., Verbeke, D., & Wyns, R. (2019). Research data management and the evolutions of scholarship: Policy, infrastructure and data literacy at KU Leuven. *LIBER Quarterly: The Journal of the Association of European Research Libraries*, *29*(1), 1–19. https://doi.org/10.18352/lq.10272

Wolski, M., Howard, L., & Richardson, J. (2017). The importance of tools in the data lifecycle. *Digital Library Perspectives*, *33*(3), 235–252. https://doi.org/10.1108/DLP-11-2016-0042

Wong, D. H. T., Maarop, N., & Samy, G. N. (2020). Data governance and data stewardship: A success procedure. *2020 8th International Conference on Information Technology and Multimedia (ICIMU)*, 54–61. https://doi.org/10.1109/ICIMU49871.2020.9243574

Wong, S. (2024). Equity, diversity, and inclusion in institutional research data management strategies: A contributive justice approach. *International Journal of Librarianship*, *8*(4), 97–113. https://doi.org/10.23974/ijol.2024.vol8.4.339

Woolley, M. (1992). Using statistics for desk research. *Aslib Proceedings*, *44*(5), 227–233. https://doi.org/10.1108/eb051276

World Bank. (2022). *Research and development expenditure (% of GDP)—World | Data*. https://data.worldbank.org/indicator/GB.XPD.RSDV.GD.ZS?locations=1W

Wylie, M., Sculthorp, M., Gagnon-Turcotte, S., & Chatwin, M. (2021). A promising path to developing data governance in the performing arts sector: exploring charters and principles for data governance. *Open North and Canadian Association for the Performing Arts (CAPACOA)*.

Yebenes, J., & Zorrilla, M. (2019). Towards a data governance framework for third generation platforms. *Procedia Computer Science*, *151*(2019), 614–621. https://doi.org/10.1016/j.procs.2019.04.082

Zain, W. M., Ahmad, Z. S., Lackie, P., Kamarudin, N., & Nordin, S. K. (2023). Academic library opportunity: Examining the research data management challenges facing Malaysian information science researchers. *Journal of Information and Knowledge Management*, *2*(Special Issue), 94–106.

Zartha Sossa, J. W., Halal, W., & Hernandez Zarta, R. (2019). Delphi method: Analysis of rounds, stakeholder and statistical indicators. *Foresight*, *21*(5), 525–544. https://doi.org/10.1108/FS-11-2018-0095

Zhang, Y. (2021). The experience of international street vendor economy's policy and enlightenment through desk research. *Frontiers in Business, Economics and Management*, *2*(3), 82–92. https://doi.org/10.54097/fbem.v2i3.204

Zhou, Q. (2018). Academic libraries in research data management service: Perceptions and practices. *Open Access Library Journal*, *5*(6), 1–12. https://doi.org/10.4236/oalib.1104693

Ziglio, E. (1996). The Delphi method and its contribution to decision-making. In M. Adler, & E. Ziglio (Eds.), *Gazing into the oracle: The Delphi method and its application to social policy and public health* (pp. 3-33). Jessica Kingsley Publishers.

Zuiderwijk, A., Shinde, R., & Jeng, W. (2020). What drives and inhibits researchers to share and use open research data? A systematic literature review to analyze factors influencing open research data adoption. *PLOS ONE*, *15*(9), 1–49. https://doi.org/10.1371/journal.pone.0239283