Assessment of Research Data Management Awareness and Practice among University and Higher Education Students: A Systematic Literature Review

Aulia Puspaning Galih[1*], and Ágnes Hajdu Barát[2]
*auliapuspa@ub.ac.id
ORCID: 0000-0003-0264-6916

ELTE Eötvös Loránd University, Budapest, Hungary[1,2]
Universitas Brawijaya, Indonesia[1]

Abstract. Researchers are aware of their need for research data awareness, knowledge, and skills. Therefore, they demand research data management training. As early career researchers and future generations of great researchers, students need to have a complete understanding and skills to manage their research data. Therefore, this study is made to identify university students' research data management practices. This study uses a systematic literature review as the method. PRISMA framework is used to obtain the reviewed literature. Web of Science database results from specific keywords were downloaded, and 60 articles were found. After the screening, the literature to be reviewed is five articles. The results are that there is clear evidence that university students are not adequately prepared to manage research data. This is evident because they have only a rudimentary understanding of the concept, process, and service of research data management at their institutions. Students can only be expected to have advanced knowledge of research data management if they are aware of its importance. Therefore, to improve their awareness, understanding, and skills, it is necessary to conduct research data management training.

Keywords: data management, higher education, PRISMA, systematic literature review, research data management

1. Introduction

Over the past decade, the significance of research data management was reported to have experienced substantial growth due to the exponential increase in data [1] and the rapid advancement of technology. In this context, it is crucial to emphasize that the concept constitutes the fundamental basis for advancements in research. The data facilitates the development of new perspectives, insights, and ground-breaking solutions supported by the evidence necessary for rational decision-making [2]. The importance of efficient research data management is consistently underscored by funding agencies, governments, libraries, universities, and institutions all over the world. However, a significant proportion of research may not be well-versed in the recommended procedures for management [3].

Given the proliferation of research data, efficient data management skills must be cultivated. These skills include technical and non-technical skills [4] and their relevance extends across diverse fields. Meanwhile, non-technical skills constitute a wide array of abilities, including self-management, effective communication, interpersonal relationships, emotional intelligence, collaborative teamwork, adept relationship-building, leadership, managerial

AICOBPA 2023, November 01, Malang, Indonesia

DOI 10.4108/eai.1-11-2023.2355286

capabilities, analytical and critical thinking, problem-solving, integrated thinking, sound judgment and decision-making, and professional skepticism [5]. Technical skills constitute a comprehensive understanding of data types, metadata, and the relevant regulatory and legal frameworks. Proficiency in training and advocacy, as well as contextual knowledge and skills in information technologies, is also in high demand within this context [4].

In the efforts to develop skills, iSchools and Library and Information Science schools in the United States offer courses on research data management [6], but this practice is not followed by many universities due to different limitations. Therefore, several researchers lack understanding and awareness of research data management and do not practice the concept [7]. In this context, awareness and understanding have been analyzed, and the importance was recognized [8] to support the lack of knowledge and skill.

2. Objective

This study was conducted to identify research data management practices of university and higher education students, regardless of their course duration. Furthermore, student's research data management awareness, skills, and training were analyzed. A systematic literature review helped provide evidence through rational thinking, holistic synthesis, and evaluation of the relevant literature.

3. Literature Review

Data management encompasses a holistic strategy for efficiently distributing, managing, organizing, preserving, and repurposing the growing quantities of data produced through research [9]. Research data management refers to the continuous and assertive management of data, starting from its entry into the research process and extending to the sharing and preservation of significant findings. This covers various aspects, such as the long-term preservation and availability of research data, protecting data, establishing institutional data repositories, and facilitating data sharing [10].

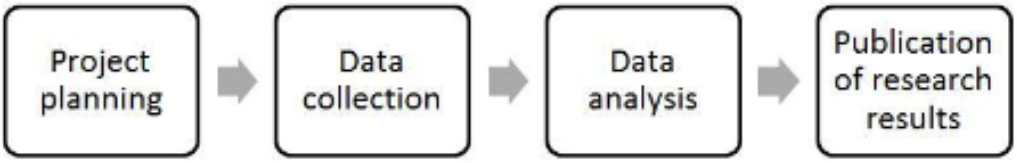

Figure 1: Traditional Data Life Cycle [11]

Traditionally, research projects publish results from collected and analyzed data. Except for excerpts, data did not dominate the research project's planning or publication of data analysis. Data was considered a byproduct of research, while publications were the project's output. As data becomes more digital and researchers can curate and preserve it, it can outlive the research study. In the 1990s and early 2000s, the data lifecycle concept was promoted to support data sharing, preservation, reuse, and repurposing beyond the original study.

4. Methods

4.1. Information Source

The present evaluation has been conducted in adherence to the guidelines outlined by the PRISMA approach for scoping review, commonly referred to as PRISMA. Web of Science database was searched on June 29, 2023, with the topic of research data management and students in higher education or university. This database was selected because of the ability to discover a list of the sources easily through links leading to articles. In this context, it used the key metrics for determining important resources. The planned protocol was performed without any significant modifications, and the methods reported were reflective of the performance.

4.2. Research Selection

A list of articles related to the topic was compiled and served as a validation tool for the search strategy. The scope was limited to articles written in English and covering the years 2014 through 2023. The results were exported to an Excel spreadsheet for manual screening. The terms of the query were ALL=(("research data management" OR "RDM") AND ("student" OR "students") AND ("higher education" OR "university" OR "universities")).

4.3. Exclusion Criteria

There are exclusion criteria compiled in this study, where the articles published before 2014 and written in non-English were eliminated. Non-English publications are eliminated because of the limitation of the language barrier. The remaining articles were screened independently by reading the abstract, and those outside the range of the topic were eliminated.

4.4. Inclusion Criteria

Following the exclusion criteria above, the remaining articles were screened manually and independently. In this study, students from all different types of higher education and university programs were included. Articles that constitute academics, librarians, and university staff members were considered for inclusion, provided specific information about students was obtained. The contexts were research data management skills, training, and awareness of higher education or university students. Another criterion was that the articles had to be open-access to be viewed and read. Furthermore, English should be used as language instruction and published between 2014-2023.

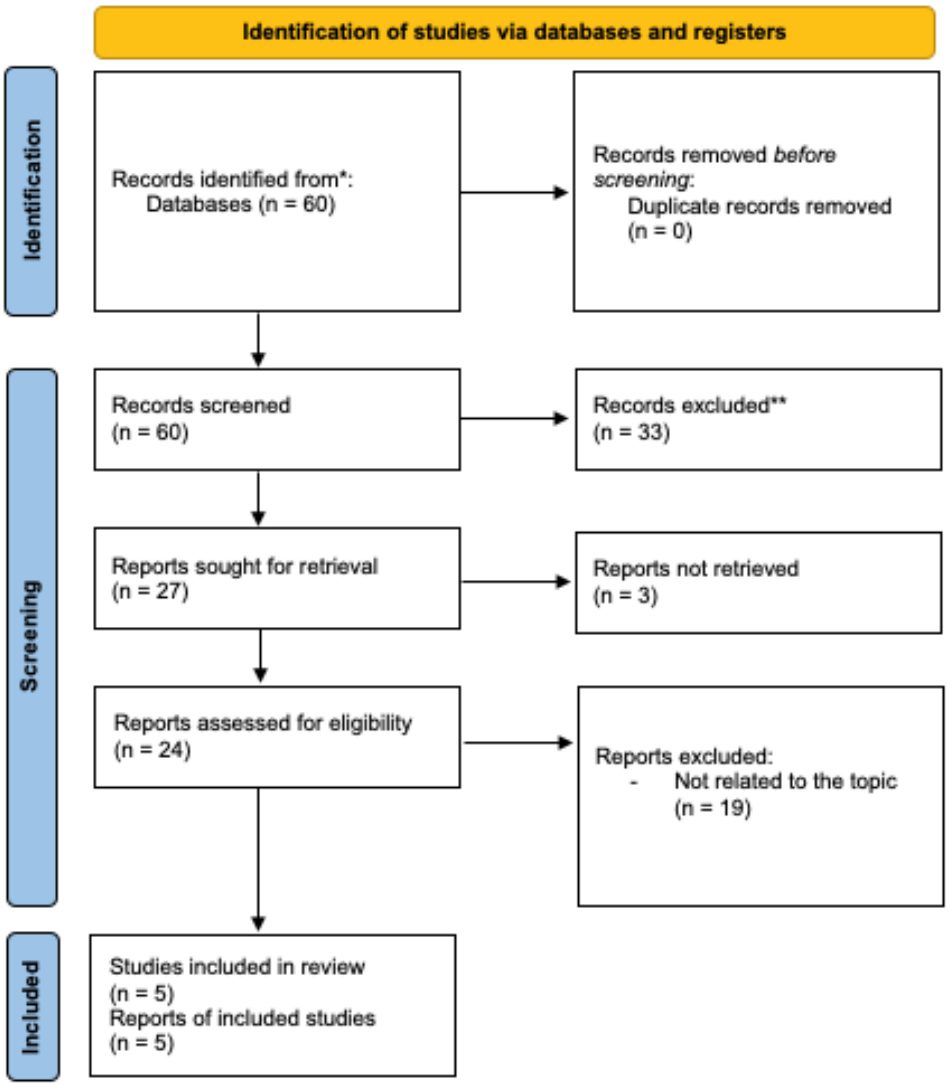

Figure 2: PRISMA Flow Diagram
Reprinted from Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71 [12]

5. Results

5.1 Overview of the Results

The initial search on the Web of Science database obtained a total of 60 articles. Following a screening process, only 5 articles have met the criteria for further discussion. Table 1 provides a comprehensive breakdown of these remaining articles, including details such as publication type, year, author names, research methods, article titles, and the respective journal names.

Table 1: Overview of the Articles

Publication Type	Year	Author Full Names	Methods	Article Title	Source Title
Journal article	2022	Xu, Zhihong; Zhou, Xuan; Kogut, Ashlynn; Clough, Michael	Quantitative	Effect of online research data management instruction on social science graduate students? RDM skills	Library & Information Science Research
Journal article	2023	Mosha, Neema Florence; Ngulube, Patrick	Mixed methods	Teaching research data management courses in higher learning institutions in Tanzania	Library Management
Journal article	2023	Zhou, Xuan; Xu, Zhihong; Kogut, Ashlynn	Mixed methods	Research data management needs assessment for social sciences graduate students: A mixed methods research	PLoS One
Journal article	2023	Xu, Zhihong; Zhou, Xuan; Watts, John; Kogut, Ashlynn	Quasi-experiment	The effect of students' engagement strategies in online instruction for data management skills	Education and Information Technologies
Journal article	2023	Kotze, Ruhan; Miller, Taryn	Quasi-experiment	Relational and decision-making skills development in South African accounting students	Industry and Higher Education

The entire articles were journal types published in 2022-2023, and only one was published in 2022 (Xu, Zhou, Kogut, & Clough, 2022). Meanwhile, the remaining four were published in 2023 (Kotze & Miller, 2023; Mosha & Ngulube, 2023; Xu et al., 2023; Zhou et al., 2023). Only one of the five articles used quantitative analysis (Xu, Zhou, Kogut, & Clough, 2022), while two used mixed methods (Mosha & Ngulube, 2023; Zhou et al., 2023). The remaining two articles used a quasi-experiment research design (Kotze & Miller, 2023; Xu et al., 2023).

5.2 Research Data Management Awareness

An understanding of research data management should be shown by students during the early part of their careers to understand the profound significance before conducting projects. A comprehensive review of the literature shows a common trend, where a significant portion of students admit to their limited awareness regarding research data management [13] the associated services offered by institutions [14]. This apparent lack of awareness can be attributed to the limitation of information. This suggests a problem in the marketing and promotion of available resources and research data management services.

According to Kotze & Miller [5], early-year students are aware that their knowledge concerning research data management needs to be developed. The necessity of being at an advanced level of practice is acknowledged. The awareness of students regarding the lack of skills and knowledge should also be discussed with the requirements of research data management. In this context, there is a requirement to raise the level of awareness regarding the resources, knowledge, and skills needed. The first step to practice effective research data management is making a plan. Students in university or higher education institutions have a proven lack of awareness of data management plan standard practice [14]. There are several processes, but there is no awareness of the standardized procedures.

5.3 Research Data Management Skills

A total of thirteen research data management non-technical skills need to be possessed. These skills are divided into two categories: relational (self-management, communication, people, emotional intelligence, teamwork, relationship-building, leadership, and managing others) and decision-making skills (analytical or critical thinking, problem-solving, integrated thinking, judgment and decision-making, professional skepticism) [5]. Students have self-awareness of their skills with self-management at the highest level, followed by analytical and problem-solving skills. The inability to manage others with a lack of leadership and relationship-building skills was also acknowledged. Therefore, students possessed better decision-making skills and awareness influenced their ability to manage the research data. In this context, problem-solving skill is needed in managing their research, as well as self-management and integrated thinking. It was believed that managing others and leadership skills were not quite necessary. These perceptions show a lack of understanding of practicing efficient research data management.

University students lacking proficiency tend to show lower levels of confidence in handling their research data [14]. This deficiency compels individuals to seek opportunities for enhancing knowledge and skills in practicing effective data management. The results are consistent with Kotze & Miller's research [5], which underscores the importance of cultivating and refining these skills to an advanced level. Furthermore, new students show a heightened concern regarding the enhancement, showing that novices possess more evident skill deficits than their counterparts at higher academic levels. This discrepancy implies that new students face deficiencies compared to those at advanced academic levels. For those who attain advanced skill levels, the acquisition of research data management literacy becomes feasible. This includes the proficiency to engage effectively in the various facets of research, spanning from data collection to curation [13].

5.4 Research Data Management Training

Research data management non-technical skills assessment is conducted by Kotze & Miller [5] and results in the training program. Students have a perception that the academic program should allocate greater efforts toward the enhancement of non-technical skills. The relational skills of managing others, leadership, interpersonal dynamics, and relationship

building were identified as the least prioritized areas to be developed. Based on the analysis of survey data, there is a correlation between the confidence of students' level and their expressed requirements for professional development [14]. Due to a lack of confidence in understanding the concept, individuals show a pronounced demand for comprehensive instruction and training in this domain.

Another research was conducted to prove the necessity of research data management training. [15] experimented on this topic and compared students who learned the teaching materials by themselves, as well as teacher-centered groups and students-centered learning groups. The results showed that the student-centered learning group achieved the highest score compared to the other groups. In general, it reported the most substantial influence on the research data management skills of students. Conversely, the group with a teacher-centered method showed a comparatively lesser but noteworthy impact on the variable. These results were in contrast to the control group, which received no instruction and did not experience any discernible effects. In this context, there was a need to teach in a research data management training program [13]. [16] found that there was no statistically significant difference in research data management skills of students based on the number of years enrolled in the current program. Therefore, the enrolment of students in a major in university does not necessarily improve their research data management skills.

6. Discussion

6.1 Research Data Management Awareness

The lack of awareness of students is consistent with prior studies that found early career researchers did not have sufficient awareness of the research data management and its services offered by their institutions [17]. In some cases, students are aware of the importance of the concept but possess insufficient knowledge regarding it [14]. The unfamiliarity with the commencement of the process leads to the creation of a personalized research data lifecycle that appears incomplete and imperfect [17].

The low level of awareness in the field of social science poses a significant concern in the library [14]. In this context, the research data management service is a comprehensive solution that necessitates the collaboration of various components to perform cohesively [18]. The variable covers various activities throughout the entire data lifecycle, including data planning, collection, processing and analysis, publishing and sharing, preservation, and reusability [19]. Furthermore, the implementation needs a thorough awareness of technical, ethical, and legal considerations, with adherence to governmental regulations on the dissemination of research data [18].

A shortcoming in awareness of students of effective research data management comes from a limited understanding of the collaborative nature of practice. The process cannot be conducted as an individual work since collaboration with other project stakeholders is important [1]. Engaging with other individuals is essential for successful research data management [5].

6.2 Research Data Management Skills

The result of the reviewed articles shows that students lack leadership skills [5]; meanwhile leadership concept is important for their professional advancement [20]. Leadership has the potential to assist in the formulation of a personal strategic plan to effectively guide and facilitate the attainment of career objectives. It is proven that when students do not possess leadership skills, their ability to manage others will be affected [21]. In the context of dynamic and rapidly evolving global and technological innovation environments, it has become increasingly crucial

to possess the capacity to strategically steer career trajectories. This includes the ability to assess and nurture fundamental leadership skills that play an important role in attaining success.

In research data training, students must practice management skills [22]. This can be achieved through efficient research data management practice and the ability to communicate and interact with others [23]. In this context, the academic curriculum should assume a more prominent role in fostering the development of these interpersonal abilities.

6.3 Research Data Management Training

In research data training, students must train their skills with case research to develop their soft and non-technical skills [22]. Case research can be a choice of conventional teaching and training methods. Developing soft skills can be accomplished through the use of collaborative learning environments. It was determined that the case research was beneficial to the development of soft skills [24]. The teaching and training should be conducted in multiple ways and platforms, such as face-to-face, online, or hybrid [16,25].

The majority of research data training provided by academic libraries or other institutions has been targeted toward faculty members and librarians [15]. Therefore, the lack of research data management skills of university students is understandable. Regarding the lack of awareness and low level of skills, the variable needs to be integrated into the core curriculum of the university [14] in every major program in the university.

7. Limitation

A key limitation of this research is the absence of prior literature (2014-2021) to enhance the complexity. Even though the results from the database are sufficient to range from 2014 to 2023, the suitable articles are only 5 from 2022 to 2013. Furthermore, only the Web of Science database was used in this literature.

8. Conclusion

In conclusion, students in higher education institutions and the university were proven to lack research data management awareness. This was reported from their minimum knowledge of the concept and its procedure and service in institutions. With the lack of awareness, students were not expected to possess high-level skills in research data management. A low level of confidence in practicing efficient management was shown because of the limitation in skills. Therefore, research data management training was required, preferably through online and case research methods, to improve their skills.

References

[1] Rod AB, Zhou B, Rousseau ME. There's no "I" in Research Data Management: Reshaping RDM Services Toward a Collaborative Multi-Stakeholder Model. J Escience Librariansh. 2023 Feb 15;12(1).

[2] Khalid AF, Lavis JN, El-Jardali F, Vanstone M. Supporting the use of research evidence in decision-making in crisis zones in low- and middle-income countries: a critical interpretive synthesis. Health Res Policy Syst. 2020 Dec 18;18(1):21.

[3] Clare C, Cruz M, Papadopoulou E, Savage J, Teperek M, Wang Y, et al., editors. Engaging Researchers with Data Management. Open Book Publishers; 2019.

[4] Kennan MA. Data management: Knowledge and skills required in research, scientific and technical organisations. In: IFLA World Library and Information Congress: 82nd IFLA General Conference and Assembly. Columbus, USA: IFLA; 2016. p. 1–10.
[5] Kotze R, Miller T. Relational and decision-making skills development in South African accounting students. Industry and Higher Education. 2023 Apr 9;095042222311620.
[6] Harris-Pierce RL, Quan Liu Y. Is data curation education at library and information science schools in North America adequate? New Library World. 2012 Nov 16;113(11/12):598–613.
[7] Tenopir C, Rice NM, Allard S, Baird L, Borycz J, Christian L, et al. Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. PLoS One. 2020 Mar 11;15(3):e0229003.
[8] Tang R, Hu Z. Providing Research Data Management (RDM) Services in Libraries: Preparedness, Roles, Challenges, and Training for RDM Practice. Data Inf Manag. 2019 Jun;3(2):84–101.
[9] Lyon L. The Informatics Transform: Re-Engineering Libraries for the Data Decade. Int J Digit Curation. 2012 Mar 11;7(1):126–38.
[10] Wilms KL, Stieglitz S, Ross B, Meske C. A value-based perspective on supporting and hindering factors for research data management. Int J Inf Manage. 2020 Oct;54:102174.
[11] Aranha S, Wigham CR. Virtual exchanges as complex research environments: facing the data management challenge. A case study of Teletandem Brasil. Journal of Virtual Exchange. 2020 Aug 17;3:13–38.
[12] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021 Mar 29;n71.
[13] Mosha NF, Ngulube P. Teaching research data management courses in higher learning institutions in Tanzania. Library Management. 2023 Mar 7;44(1/2):166–79.
[14] Zhou X, Xu Z, Kogut A. Research data management needs assessment for social sciences graduate students: A mixed methods study. PLoS One. 2023 Feb 23;18(2):e0282152.
[15] Xu Z, Zhou X, Kogut A, Watts J. A Scoping Review: Synthesizing Evidence on Data Management Instruction in Academic Libraries. The Journal of Academic Librarianship. 2022 May;48(3):102508.
[16] Xu Z, Zhou X, Kogut A, Clough M. Effect of online research data management instruction on social science graduate students' RDM skills. Libr Inf Sci Res. 2022 Oct;44(4):101190.
[17] Mattern E, Jeng W, He D, Lyon L, Brenner A. Using participatory design and visual narrative inquiry to investigate researchers' data challenges and recommendations for library research data services. Program: electronic library and information systems. 2015 Sep 1;49(4):408–23.
[18] Pinfield S, Cox AM, Smith J. Research Data Management and Libraries: Relationships, Activities, Drivers and Influences. PLoS One. 2014 Dec 8;9(12):e114734.
[19] Mosconi G, Li Q, Randall D, Karasti H, Tolmie P, Barutzky J, et al. Three Gaps in Opening Science. Computer Supported Cooperative Work (CSCW). 2019 Jun 29;28(3–4):749–89.
[20] Rubens A, Schoenfeld GA, Schaffer BS, Leah JS. Self-awareness and leadership: Developing an individual strategic professional development plan in an MBA leadership course. The International Journal of Management Education. 2018 Mar;16(1):1–13.

[21] Brown A, Stedman N, Sowcik M, Bunch J. EXPLORING STOICISM IN LEADERSHIP: A Comparison with Emotional Intelligence in Undergraduate Leadership Students. The Journal of Leadership Education. 2023 Jan;22(1).
[22] Keevy M. Using case studies to transfer soft skills (also known as pervasive skills). Meditari Accountancy Research. 2016 Aug 8;24(3):458–74.
[23] Ghani MA, Suryani AW. Professional Skills Requirements for Accountants: Analysis of Accounting Job Advertisements. Jurnal ASET (Akuntansi Riset). 2020 Dec 29;12(2):212–26.
[24] Samkin G, Keevy M. Using a stakeholder developed case study to develop soft skills. Meditari Accountancy Research. 2019 Feb 12;27(6):862–82.
[25] Oo CZ, Chew AW, Wong ALH, Gladding J, Stenstrom C. Delineating the successful features of research data management training: a systematic review. International Journal for Academic Development. 2022 Jul 3;27(3):249–64.