

Strategy and Motivation of Institutional Research Data Management: Integration of SWOT Analysis and Convergence Matrix

Fitroh^{1,2}, Irman Hermadi^{1*}, Kudang Boro Seminar³, Rina Trisminingsih¹

Abstract—Research data management (RDM) in institutional repositories requires alignment between stakeholder needs, organizational conditions, and technical readiness for the findable, accessible, interoperable, and reusable (FAIR) principles. This study developed and applied an integrated institutional RDM evaluation framework combining needs analysis, SWOT analysis, and the GO FAIR Convergence Matrix. Design science research guided the framework development. Data were collected through a purposive survey of 105 stakeholders, expert interviews, repository observations, and independent FAIR ratings by two expert evaluators. The survey showed that 86.7% of respondents selected the two highest response categories for easy repository data access and use, while 87.6% did so for improving the current data management system. SWOT analysis identified institutional commitment as the principal strength (89.5%), additional RDM training as the leading weakness-related need (65.7%), emerging-technology adoption as the principal opportunity (81.0%), and government policy changes as the leading threat (73.3%). The FAIR assessment revealed a substantial gap between articles and datasets. Articles achieved normalized scores of 100.0% for Findable, 100.0% for Accessible, 62.5% for Interoperable, and 45.8% for Reusable, whereas datasets achieved 3.1% for Findable and 0.0% for the other domains. The study contributes an integrated framework that connects stakeholder priorities, organizational SWOT factors, and indicator-level FAIR evidence to support phased, evidence-based institutional RDM interventions.

Index Terms—Convergence matrix, FAIR principles, institutional repository, research data management, SWOT analysis.

I. INTRODUCTION

Institutional repositories are increasingly expected to support not only the storage and preservation of scholarly outputs but also the curation, discovery, and responsible reuse of research data. Institutional research data management (RDM) extends traditional repository functions by organizing research data throughout their lifecycle and supporting open-science practices [1], [2]. The FAIR principles provide an operational basis for this role by requiring research data and metadata to be discoverable, accessible under clearly defined conditions, interoperable through shared standards, and reusable with sufficient contextual information [3], [4].

FAIR-based RDM cannot be achieved through repository technology alone. Its implementation also depends on institutional policy, infrastructure, staff and researcher capabilities, incentives, stakeholder coordination, and data-sharing practices [5]–[8]. Previous studies have therefore characterized RDM readiness as a multidimensional institutional issue involving technological, organizational, people-related, and environmental factors [7], [8].

Previous research has provided valuable but fragmented perspectives on institutional RDM readiness. Organizational readiness and maturity studies assess dimensions such as awareness, data policy, funding, services, user needs, and information-technology infrastructure [7]–[9]. These approaches support institutional benchmarking and the identification of service-development priorities, but they do not directly indicate which FAIR requirements have been implemented at the repository level. Conversely, FAIR-oriented repository assessments examine observable technical features and indicator fulfillment [3], but they provide limited guidance on how stakeholder needs and institutional conditions should shape strategic priorities. Consequently, universities still lack an integrated basis for translating stakeholder demand, organizational context, and technical repository evidence into prioritized RDM interventions.

This study addresses this gap by integrating three complementary analytical components. Needs analysis captures stakeholder priorities and expectations regarding institutional RDM. SWOT analysis examines the internal and external organizational factors that may enable or constrain

Received: 8 July 2026; Revised: 13 July 2026; Accepted: 23 July 2026.

*Corresponding author

¹Fitroh, Department of Computer Science, IPB University (e-mail: 88salamunfitroh@apps.ipb.ac.id).

²Fitroh, Information Systems Study Program, UIN Syarif Hidayatullah Jakarta Indonesia (e-mail: fitroh@uinjkt.ac.id).

¹Irman Hermadi, Department of Computer Science, IPB University (e-mail: irmanhermadi@apps.ipb.ac.id).

³Kudang Boro Seminar, Mechanical Engineering and Biosystems Study Program, IPB University (e-mail: kseminar@apps.ipb.ac.id).

¹Rina Trisminingsih, Department of Computer Science, IPB University (e-mail: rina.trisminingsih@apps.ipb.ac.id).

implementation. The GO FAIR convergence matrix provides an indicator-level assessment of observable repository features for research articles and datasets. This integration is necessary because stakeholder needs alone do not demonstrate technical FAIR readiness, SWOT analysis does not directly measure the fulfillment of FAIR indicators, and the Convergence Matrix does not explain the institutional conditions underlying the observed repository gaps. By synthesizing the three components, the proposed framework connects stakeholder demand, organizational context, and technical repository evidence, thereby supporting the formulation and prioritization of institutionally relevant FAIR-based RDM strategies.

At the time of the study, IPB University maintained an institutional repository for collecting, preserving, and providing access to publications and other scholarly outputs produced by its academic community. A dataverse-based research data service was also available at <https://dataverse.ipb.ac.id/>. Preliminary institutional observations indicated that the development of RDM involved not only repository infrastructure but also data-sharing practices, governance arrangements, institutional policy, and the readiness of the academic community. Accordingly, this study focuses on institutional evaluation and strategy prioritization rather than on the development or technical performance testing of a repository application.

Based on this context, the study addresses the following research questions:

RQ1: How do relevant stakeholders assess the importance of data accessibility and improved institutional RDM services at IPB University?

RQ2: Which internal and external organizational factors shape the implementation of FAIR-based RDM at IPB University?

RQ3: What is the indicator-level FAIR readiness of research articles and datasets in the IPB repository, as assessed using the GO FAIR Convergence Matrix?

The study aims to develop and apply an integrated institutional RDM evaluation framework that captures stakeholder needs, maps relevant internal and external organizational factors, and assesses repository readiness at the FAIR-indicator level. The findings from these components are synthesized to identify and prioritize institutionally relevant strategies for the phased implementation of FAIR-based RDM at IPB University.

II. RESEARCH METHOD

A. Research Design

Design science research (DSR) guided the development of the institutional RDM evaluation framework, which was treated as the main research artifact. DSR was selected because it provides a systematic approach to designing and evaluating artifacts that address relevant information-system problems, [10], [11]. Its use in this study is also supported by recent research applying DSR to institutional repository architecture [12] and FAIR-oriented data management systems [13].

The work proceeded through three connected stages. First, the institutional problem was identified by examining stakeholder needs and the existing condition of the IPB repository through surveys, expert interviews, and system

observations. Second, the framework was designed by combining three complementary components: needs analysis to capture user priorities, SWOT analysis to identify relevant organizational factors, and the GO FAIR Convergence Matrix to assess repository readiness at the indicator level. The findings from these components were then synthesized to identify strategic priorities for FAIR-based RDM. Finally, the framework was applied to the IPB repository and evaluated through triangulation of survey findings, expert assessments, and observed repository features. This process provided a structured link between the identified institutional problems, the repository readiness assessment, and the resulting strategic priorities. The research stages are shown in Fig. 1.

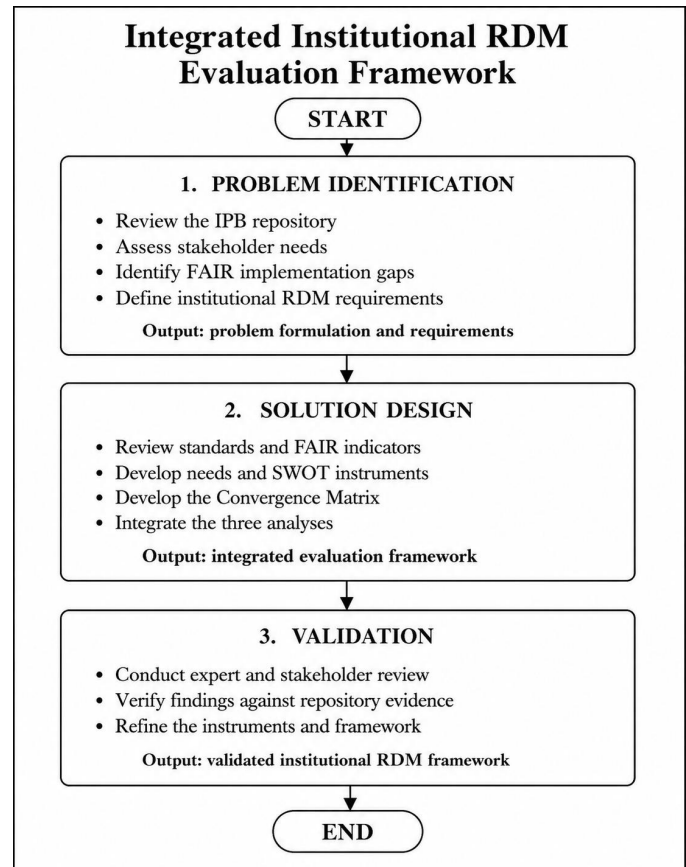


Fig. 1. Research stages of the integrated institutional RDM evaluation framework.

B. Research Population and Sample

A purposive, stakeholder-based sampling strategy was used to recruit 105 respondents across seven predefined groups: academics or researchers, administrative or technical staff, students, policymakers and managers, external users, librarians and curators, and industry or community stakeholders. Students formed the largest group because they constituted the largest stakeholder population within the institution. The sample was considered adequate for the exploratory and descriptive objectives of the needs and SWOT analyses because it included all stakeholder groups relevant to institutional RDM and captured a range of user, managerial, technical, and external perspectives. The study does not claim that the sample statistically represents the entire IPB University population. (Table 1). The 105 survey respondents participated only in the

needs and SWOT analyses. Repository readiness was assessed separately by two purposively selected expert evaluators: one representative from the Information Management and Digital Technology Institute and one representative from the IPB University Library. The evaluators were selected because of their direct technical and operational knowledge of the institutional repository. Their composition is presented in Table 2.

Table 1.
Respondents of Needs Analysis and SWOT Analysis

No	Respondent Classification	Total	Percentage (%)
1	Academics or researchers	23	21.9
2	Administrative staff/technical staff	5	4.8
3	Students	47	44.8
4	Policy makers and managers	5	4.8
5	External users	10	9.5
6	Librarians and curators	3	2.9
7	Stakeholders/industry/community	12	11.4
	Total	105	100

Table 2.
Expert Evaluators for The GO FAIR Convergence Matrix

No	Expert evaluator affiliation	Number of evaluators	Percentage (%)
1	Information Management and Digital Technology Institute	1	50
2	IPB University Library	1	50
	Total	2	100

C. Research Instrument

The needs and SWOT questionnaire was developed in three stages. First, the needs items were formulated from the research objectives concerning research-data accessibility and repository improvement (Table 3). Second, the initial questionnaire domains were informed by the RDM readiness factors proposed by [7], which cover technological, organizational, people-related, and environmental dimensions. Relevant factors were then mapped into the SWOT categories and reformulated to reflect the institutional RDM context at IPB University. Third, the draft questionnaire was reviewed by language, domain, and repository experts to assess its clarity, readability, and content relevance before being administered using a five-point Likert scale. The second instrument, the GO FAIR convergence matrix, was used to evaluate FAIR-indicator fulfillment for two entities: research articles and datasets (Table 4). The indicator structure was informed by published interpretations and implementation considerations for the FAIR principles [14]. Each indicator was rated using the five-point evidence-based scale described in the Data Analysis Techniques subsection.

Table 3.
Instrument of Needs Analysis and SWOT Analysis

No	Variable	Operational Definition
1	Research data availability needs and expectations at IPB	Ease of access to research data and improved data access
2	Strengths	a. Academic and research expertise b. Rich data resources

		c. Institutional support d. Active academic community (academic, professional, environmental, online, and research & development communities)
3	Weaknesses	a. Limited IT infrastructure b. Lack of awareness about big data management c. Financial resources d. Immature data policies
4	Opportunities	a. Interdisciplinary collaboration b. Support from government and NGOs c. Adoption of the latest technology d. Increasing research visibility and impact
5	Threats	a. Government policy changes b. Academic competition c. Rapid technological evolution d. Data security risks

Table 4.
Instrument of Convergence Matrix

No	Variable	Operational Definition	Indicator
1	Findable	Ease of discovery of article and dataset metadata	- F1: A unique and persistent identifier is available (DOI, Handle, URI) - F2: Metadata uses a standard schema (DublinCore) - F3: Metadata links to the described object - F4: Metadata is indexed by search engines (Google Scholar, aggregator)
2	Accessible	Availability of long-term metadata with open protocols and authorization mechanisms	- A1.1: article and dataset Available via open protocols - A1.2: article and dataset Protected through authentication and authorization mechanisms if required - A2: article and dataset remains available in the long term, even if the main data are no longer available.
3	Interoperable	Ability of article and dataset metadata to be integrated across systems	- I1: Metadata uses a formal machine-readable format (RDF, JSON-LD) - I2: Metadata uses a structured vocabulary (DublinCore, ontology) - I3: Metadata follows cross-repository standard schemas
4	Reusable	Ease of reusing article and dataset metadata	- R1.1: Metadata includes a clear license - R1.2: Metadata includes provenance (origin, change history) - R1.3: Metadata follows community/domain ontology standards
5	Constraints and Recommendations	Barriers and opportunities for developing FAIR-based repositories	- Main barriers to FAIR implementation - Recommendations for IPB repository readiness

D. Instrument Quality Assessment

The instruments underwent qualitative content review and quantitative face-validity and readability assessment. Following established expert-review procedures for instrument validation [15], domain and repository experts examined the conceptual alignment of the items with the intended needs, SWOT, and FAIR constructs. The experts subsequently rated six criteria related to language appropriateness, clarity, conciseness, comprehensibility of response options, layout, and font readability using a four-point scale. Ratings of 3 or 4 were classified as agreement. The results are presented in Tables 5 and 6.

For each criterion, the I-CVI was calculated by dividing the number of experts assigning a rating of 3 or 4 by the total number of experts, following established CVI calculation and reporting procedures [16], [17]. An I-CVI of at least 0.78 was considered acceptable for panels comprising three or more experts [17]. All six criteria in both instruments achieved an I-CVI of 1.00. The S-CVI/Ave was calculated as the mean of the six I-CVI values, whereas the S-CVI/UA was calculated as the proportion of criteria receiving universal agreement from all experts [16]. Because all criteria received universal agreement, the S-CVI/Ave and S-CVI/UA values were both 1.00 for each instrument. These results indicate unanimous expert agreement regarding the clarity, readability, and comprehensibility of the instruments.

The needs and SWOT questionnaire was designed as a diagnostic instrument comprising distinct RDM and SWOT

indicators rather than as a unidimensional psychometric scale. Each item was analyzed separately using response distributions, and no composite scale score was calculated. Cronbach's alpha evaluates the internal consistency of related items intended to form a common scale and does not, by itself, establish unidimensionality [18]. Internal-consistency coefficients are also not necessarily appropriate when indicators represent distinct and non-interchangeable aspects of a construct [19]. Cronbach's alpha was therefore not calculated, and no claim of internal consistency or composite-scale reliability is made in this study. The consistency of the GO FAIR Convergence Matrix assessment was addressed procedurally by reviewing the two evaluators' ratings, clarifying discrepant scores through follow-up discussion, and verifying the assessments against repository observations and supporting evidence.

Repository readiness was assessed by two purposively selected expert evaluators with direct technical and operational knowledge of the institutional repository. Each evaluator independently rated the 26 entity-specific FAIR assessment items using the five-point evidence-based scale. After independent scoring, the ratings were compared item by item. When ratings differed, each evaluator explained the evidence underlying the assigned score, and the indicator definition was reviewed against observable repository features and available supporting documentation. The process sought agreement on the interpretation of the evidence rather than forcing identical numerical ratings. If a one-point difference remained after clarification, both ratings were retained and their arithmetic mean was used as the indicator score.

Table 5.
Face Validity and Readability Assessment of the Needs and SWOT Questionnaire

Number	Criterion	Expert 1	Expert 2	Expert 3	Expert 4	Experts Rating 3 or 4	Total Number of Experts	I-CVI	Inter-pretati on
1	The language used is appropriate for the target evaluators understanding	4	3	4	3	4	4	1	Acceptable
2	Questions in the assessment item or matrix are clear and unambiguous	3	4	3	4	4	4	1	Acceptable
3	Statements are short and to the point	4	3	3	3	4	4	1	Acceptable
4	Answer choices are easy to understand and not confusing	3	3	4	4	4	4	1	Acceptable
5	The matrix layout is neat and easy to follow	4	4	4	4	4	4	1	Acceptable
6	Font size is adequate and the typeface is easy to read	3	3	3	3	4	4	1	Acceptable

Table 6.
Face Validity and Readability Assessment of the GO FAIR Convergence Matrix

Number	Criterion	Expert 1	Expert 2	Expert 3	Total Score 3-4	Number of experts	I-CVI	Inter-pretatio n
1	The language used is appropriate for the target evaluators understanding	4	4	4	3	3	1	Acceptable
2	Questions in the assessment item or matrix are clear and unambiguous	3	4	4	3	3	1	Acceptable
3	Statements are short and to the point	4	4	4	3	3	1	Acceptable
4	Answer choices are easy to understand and not confusing	4	4	3	3	3	1	Acceptable
5	The matrix layout is neat and easy to follow	4	4	4	3	3	1	Acceptable
6	Font size is adequate and the typeface is easy to read	4	3	4	3	3	1	Acceptable

E. Data Collection Procedure

The three data-collection methods served distinct but complementary roles in the triangulation process. The survey provided quantitative evidence of stakeholder needs and perceptions of the internal and external factors affecting institutional RDM. Expert interviews were used to interpret the organizational, operational, and technical context of the repository and to clarify the ratings assigned in the GO FAIR Convergence Matrix. System observations examined observable repository features, metadata practices, access mechanisms, and other supporting evidence. Triangulation was conducted by comparing findings across these three sources. Convergent evidence strengthened the resulting interpretation, whereas discrepancies were re-examined through follow-up clarification with the evaluators and verification against repository observations and supporting documentation.

Before completing the questionnaire, respondents received written information describing the researcher's identity, the study purpose, the expected completion time, and the intended use of the responses. The questionnaire stated that the information provided would be used solely for research purposes and treated confidentially. Because no separate consent form or consent checkbox was used, completion and return of the questionnaire were treated as implied consent. Respondent names were not requested. Profile information was used only to describe and classify the stakeholder groups, and the findings were analyzed and reported in aggregate form to reduce the risk of identifying individual participants.

F. Data Analysis Techniques

Responses to each needs and SWOT item were coded according to their ordinal positions on the five-point response scale. For each response category, the frequency and percentage were calculated using $P=(n/105) \times 100$, where n is the number of respondents selecting that category. Because the questionnaire comprised distinct diagnostic indicators, no composite scores were calculated for the Strength, Weakness, Opportunity, or Threat dimensions. All percentages reported in Table 7 and the related narrative represent the combined frequencies of response categories 4 and 5, calculated as $P4+5 = [(n4 + n5)/105] \times 100$ and rounded to one decimal place. Key SWOT factors were prioritized based on the magnitude of respondent endorsement or concern and their relevance to institutional RDM. The SO, WO, ST, and WT strategies were subsequently formulated by matching the prioritized internal factors with the relevant external factors; these strategies were interpretive outputs rather than additional quantitative scores.

Each evaluator independently rated the FAIR indicators for research articles and datasets using a five-point evidence-based scale: 1 = not fulfilled, 2 = minimally fulfilled, 3 = partially fulfilled, 4 = largely fulfilled, and 5 = fully fulfilled. The two evaluator ratings were compared, and their arithmetic mean was used as the indicator score. The ratings were interpreted with reference to observable repository features and available supporting documentation. For each indicator and entity, the indicator score was calculated as the arithmetic mean of the two

verified evaluator ratings. The domain score was then calculated separately for research articles and datasets by averaging all indicator scores within the corresponding Findable, Accessible, Interoperable, or Reusable domain. For presentation on a 0–100 scale, each domain mean was normalized using $R=[(x^- - 1)/4] \times 100$, where x^- is the mean domain score. This transformation assigns 0% to a score of 1 and 100% to a score of 5. Scores close to 1 indicate nonfulfillment or very limited implementation, values around 3 indicate partial fulfillment, and scores close to 5 indicate substantial or full fulfillment. A combined domain score was calculated as the mean of the article and dataset domain scores only to provide an overall descriptive summary. These scores represent an evidence-informed expert assessment of indicator fulfillment and should not be interpreted as a population estimate or formal FAIR maturity certification.

G. Data Triangulation

Triangulation integrated the survey, expert interview, and repository observation findings. Survey results identified stakeholder needs and organizational perceptions; interviews provided contextual explanations of governance, operational, and technical conditions; and system observations verified the repository features used in the FAIR assessment. Convergent patterns across the three sources informed the institutional priorities, whereas discrepancies were re-examined through clarification with the evaluators and verification against supporting repository evidence.

III. RESULT

A. User Needs and Expectations

The survey identified two dominant user needs. Based on the raw response frequencies, 91 of 105 respondents (86.7%) selected the two highest response categories regarding the importance of easy repository data access and use. Similarly, 92 respondents (87.6%) selected the two highest response categories regarding the need to improve the current repository data management system. These findings are visually compared in Fig. 2.

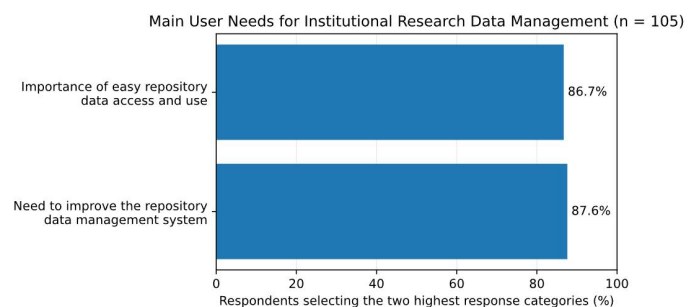


Fig. 2. Main user needs for institutional research data management.
Note: Percentages represent the combined frequencies of the two highest response categories based on 105 valid responses

The high demand for accessible data and an improved data

management system has direct implications for institutional RDM design. Repository development should prioritize user-centered discovery through structured metadata, persistent identifiers, effective search and filtering functions, and explicit links between research publications and their supporting datasets. Accessibility should be implemented through transparent access conditions rather than interpreted as unrestricted openness. The system should therefore distinguish among open, restricted, and sensitive data and clearly communicate the relevant request, authentication, and authorization procedures. To support reuse, dataset records should include machine-readable licenses, provenance information, standard metadata, and clear citation and download information. These technical requirements should be

accompanied by institutional support, including deposit guidance, metadata assistance, FAIR and RDM training, and data-stewardship services. Together, these priorities translate the identified user needs into functional and governance requirements for the phased development of RDM at IPB University.

B. SWOT Analysis

A SWOT analysis was used to organize the item-level survey findings into internal and external factors relevant to the implementation of FAIR-based RDM at IPB University. Table 7 presents the standardized percentages for the identified strengths, weaknesses, opportunities, and threats, together with the resulting SO, WO, ST, and WT strategies and the priority ranking of the most critical factors.

Table 7
Results of the SWOT analysis at IPB University

Internal Factors		
Strengths	Weaknesses	
1. Academic and research capacity. A total of 73.3% of respondents were confident or very confident in IPB's academic and research capacity, while 67.6% rated IPB's research contribution as significant or highly significant. 2. Rich research-data resources. A total of 64.8% of respondents considered IPB's research data diverse or highly diverse and sufficiently or highly in-depth. In addition, 49.5% rated the accessible data coverage as broad or very broad. 3. Institutional support. A total of 89.5% of respondents agreed or strongly agreed that IPB had a strong commitment to innovation and research. 4. Active academic community. A total of 71.4% of respondents considered the academic community open or very open to new research and development ideas.	1. IT infrastructure performance. Only 39.0% of respondents rated the speed and efficiency of the repository infrastructure as fast or very fast. Moreover, 9.5% reported frequent or continuous difficulties in handling large data volumes, and 7.6% frequently or continuously experienced access or storage problems. 2. Limited practical involvement and training needs. Only 29.5% of respondents reported being frequently or continuously involved in big-data management, while 65.7% expressed a high or very high need for additional information and training. 3. Financial constraints. A total of 61.9% of respondents believed that limited funding had a strong or very strong impact on repository development. 4. Data-policy maturity. Only 31.4% considered repository data management sufficiently or highly developed, 38.1% rated data standardization and management as optimal or highly optimal, and 41.0% considered the standardization guidelines detailed or very detailed.	
External Opportunities and Response Strategies		
Opportunities	SO Strategies	WO Strategies
1. Interdisciplinary collaboration. A total of 76.2% of respondents perceived substantial or strong opportunities for interdepartmental collaboration, while 70.5% believed that interinstitutional collaboration would have a significant or highly significant effect on repository management. 2. Government and non-governmental support. A total of 56.2% of respondents were optimistic or very optimistic about government support, while 52.4% were somewhat or highly confident that nonprofit organizations could provide potential funding. 3. Adoption of emerging technologies. A total of 81.0% of respondents perceived substantial or strong opportunities for IPB to adopt emerging technologies, including AI, and 84.8% considered such adoption important or very important. 4. Research visibility and impact. A total of 83.8% of respondents believed that more efficient data use would improve research quality, while 84.8% believed that it would have a large or very large positive effect on IPB's institutional image and reputation.	1. Develop an AI-assisted FAIR metadata curation and validation service by combining IPB's institutional commitment and academic expertise with the opportunity to adopt emerging technologies. The service should initially support persistent identifiers, standardized metadata, controlled vocabularies, licensing information, and provenance records for research datasets. 2. Develop interdisciplinary FAIR data packages through collaboration among academic units, librarians, repository managers, and researchers. These packages can support joint research, grant proposals, and the reuse of research data across disciplines. 3. Promote curated and citable featured datasets by leveraging IPB's research reputation, diverse data resources, and the opportunity to improve research visibility and institutional impact.	1. Use emerging technologies to address repository infrastructure and metadata gaps by developing dataset services that support persistent identifiers, standard metadata schemas, machine-readable formats, licensing, provenance, and long-term preservation. 2. Establish a data-steward and FAIR-literacy program through interdisciplinary collaboration to address limited researcher involvement and the strong demand for additional RDM training. 3. Use government, nonprofit, and collaborative funding opportunities to support infrastructure improvement, staff capacity development, and the formulation of an institutional RDM policy.

External Threats and Response Strategies

Threats	ST Strategies	WT Strategies
1. Government policy changes. A total of 73.3% of respondents believed that changes in government policy could affect or strongly affect research funding. 2. Academic competition. A total of 62.9% believed that competition with other institutions could affect or strongly affect IPB's ability to attract research and funding. 3. Rapid technological evolution. A total of 72.4% considered rapid technological evolution a significant or highly significant challenge. 4. Data-security risks. A total of 69.5% considered data-security risks a significant or highly significant concern in research-data collection and storage.	1. Use strong institutional commitment and academic expertise to establish adaptive RDM governance, including periodic reviews of policies, metadata standards, and repository procedures in response to changes in government policy and technology. 2. Use the active academic community and institutional support to implement security-aware data-sharing procedures, including data classification, tiered access, authentication, and clear reuse conditions. 3. Use IPB's academic reputation and research-data resources to differentiate the institutional repository through high-quality FAIR datasets, persistent citation, and measurable research reuse, thereby responding to competition among institutional repositories.	1. Establish minimum institutional requirements for dataset deposit, including metadata completeness, persistent identifiers, licenses, provenance documentation, access classification, and data-management plans. This strategy addresses immature data policies while reducing policy and security risks. 2. Implement a phased infrastructure-strengthening plan covering storage, backup, long-term preservation, access control, and interoperability. Prioritization should be based on risk and available funding to reduce vulnerability to rapid technological change. 3. Conduct periodic FAIR-readiness monitoring using clearly defined indicators and follow-up actions. Annual reassessment can prevent the repository from falling behind changing standards and competing institutions.

Priority Ranking of the Most Critical SWOT Factors

Priority rank	Strengths	Weakness-related findings	Opportunities	Threats
1	Strong institutional commitment to innovation and research - 89.5%	High need for additional RDM information and training - 65.7%	Adoption of emerging technologies - 81.0% perceived substantial or strong opportunities; 84.8% rated adoption important or very important	Government policy changes affecting research funding - 73.3%
2	Strong academic and research capacity - 73.3%	Perceived strong or very strong impact of limited funding on repository development - 61.9%	Potential to improve research visibility and institutional impact - 83.8% associated efficient data use with improved research quality; 84.8% associated it with institutional reputation	Rapid technological evolution as a significant or highly significant challenge - 72.4%

Note: The percentages represent the combined frequencies of response categories 4 and 5 and were recalculated from the raw frequencies using $P_{4+5} = [(n_4 + n_5) / 105] \times 100$, rounded to one decimal place. No composite scores were calculated for the Strength, Weakness, Opportunity, or Threat dimensions. For positively worded items classified as weaknesses, the term *only* indicates limited endorsement of the two highest response categories. The priority ranking identifies the two most critical factors within each SWOT dimension based on the magnitude of the combined category 4-5 responses and their direct relevance to FAIR-based RDM implementation. The ranking is ordinal and does not represent a composite SWOT score.

The interpretation prioritizes the factors receiving the strongest respondent endorsement and having the most direct implications for FAIR-based RDM implementation. Institutional commitment was identified as the principal strength, with 89.5% of respondents agreeing/strongly agreeing that IPB had a strong commitment to innovation and research. The adoption of emerging technologies presents a significant opportunity, with 81.0% recognizing substantial chances for adoption and 84.8% viewing it as important. Conversely, 65.7% identified a high demand for additional Research Data Management (RDM) training, and 61.9% noted limited funding's negative impact on repository development. Government policy changes (73.3%) and rapid technological evolution (72.4%) were prioritized as the leading threats because it may affect funding continuity, technical standards, and long-term repository sustainability. The remaining factors presented in Table 7 provide supporting institutional context. The explicit priority ranking was used to ensure the resulting strategies focused on the most strongly supported and

institutionally relevant factors rather than treating all SWOT factors as equally.

C. Evaluation of Repository Readiness using FAIR Principles through the Convergence Matrix

The aggregated FAIR readiness results for research articles and datasets are presented in Table 8. Indicator scores were derived from the mean ratings of the two evaluators, while domain scores were calculated separately for the findable, accessible, interoperable, and reusable dimensions.

Across the 26 entity-specific ratings reported in Table 8, the two evaluators provided identical scores for 23 ratings, corresponding to an exact agreement of 88.5%. Three one-point differences remained after clarification: external indexing of dataset metadata under F4, interoperability of article metadata schemas under I3, and compliance of article metadata with community-specific standards under R1.3. Both reviewed ratings were retained transparently, and the arithmetic mean was used as the indicator score.

Table 8
Indicator-level FAIR readiness assessment for research articles and datasets

FAIR domain	Indicator	Entity	Assessment criterion	Evaluator 1 (Library Manager)	Evaluator 2 (Digital Library Assistant)	Mean indicator score
Findable	F1	Research articles	Persistent and unique identifier	5	5	5.00
	F1	Datasets	Persistent and unique identifier	1	1	1.00
	F2	Research articles	Standard metadata schema	5	5	5.00
	F2	Datasets	Standard metadata schema	1	1	1.00
	F3	Research articles	Metadata linked to the described object	5	5	5.00
	F3	Datasets	Metadata linked to the corresponding dataset	1	1	1.00
	F4	Research articles	Metadata indexed by external search services	5	5	5.00
	F4	Datasets	Metadata indexed by external search services	2	1	1.50
Accessible	A1.1	Research articles	Access through open communication protocols	5	5	5.00
	A1.1	Datasets	Access through open communication protocols	1	1	1.00
	A1.2	Research articles	Authentication and authorization where required	5	5	5.00
	A1.2	Datasets	Authentication and authorization where required	1	1	1.00
	A2	Research articles	Metadata remain accessible if the object is unavailable	5	5	5.00
	A2	Datasets	Metadata remain accessible if the dataset is unavailable	1	1	1.00
Interoperable	I1	Research articles	Machine-readable formal representation	3	3	3.00
	I1	Datasets	Machine-readable formal representation	1	1	1.00
	I2	Research articles	Structured or controlled vocabulary	4	4	4.00
	I2	Datasets	Structured or controlled vocabulary	1	1	1.00
	I3	Research articles	Metadata schema interoperable with other repositories	3	4	3.50
	I3	Datasets	Metadata schema interoperable with other repositories	1	1	1.00
Reusable	R1.1	Research articles	Explicit reuse license	3	3	3.00
	R1.1	Datasets	Explicit reuse license	1	1	1.00
	R1.2	Research articles	Provenance information	3	3	3.00
	R1.2	Datasets	Provenance information	1	1	1.00
	R1.3	Research articles	Compliance with community-specific metadata standards	3	2	2.50
	R1.3	Datasets	Compliance with community-specific metadata standards	1	1	1.00

Note: Each indicator score represents the arithmetic mean of the two evaluators' ratings. Ratings ranged from 1 (not fulfilled) to 5 (fully fulfilled). The assessment criteria were condensed for readability without changing the meaning of the corresponding FAIR indicators. Domain-level mean and normalized scores are presented separately in Table 9.

The resulting domain mean and normalized scores are presented in Table 9.

Table 9
Aggregated FAIR Readiness Scores for Research Articles and Datasets

FAIR domain	Research articles: mean score	Research articles: normalized score (%)	Datasets: mean score	Datasets: normalized score (%)
Findable	5.00	100.0	1.13	3.1
Accessible	5.00	100.0	1.00	0.0
Interoperable	3.50	62.5	1.00	0.0
Reusable	2.83	45.8	1.00	0.0

Note: Domain mean scores were calculated by averaging the indicator scores within each FAIR domain separately for research articles and datasets. The normalized scores were calculated using $R = [(\bar{x} - 1)/4] \times 100$, where $\bar{x}$ represents the domain mean score. A score of 1 corresponds to 0%, whereas a score of 5 corresponds to 100%. The scores represent a descriptive expert assessment and should not be interpreted as formal FAIR maturity certification.

Table 9 reveals a substantial FAIR readiness gap between research articles and datasets. Research articles achieved normalized scores of 100.0% for Findable, 100.0% for

Accessible, 62.5% for Interoperable, and 45.8% for Reusable. In contrast, datasets achieved only 3.1% for Findable and 0.0% for Accessible, Interoperable, and Reusable. These results indicate that discovery and access mechanisms for research articles are well established, whereas interoperability and reuse remain partially fulfilled and institutional dataset management remains at an early stage across all four FAIR domains.

D. Research Articles

At the indicator level, research articles received the highest ratings for Findable and Accessible because their metadata included persistent identifiers, used the Dublin Core schema, linked to the described objects, and were indexed by external search services. Article metadata were also available through open protocols, particularly OAI-PMH, and remained accessible through the repository infrastructure. These features explain the full normalized scores obtained for the Findable and Accessible domains.

The lower Interoperable score reflects the partial

implementation of machine-readable and semantically structured formats. Although the repository uses standardized metadata and controlled vocabularies, support for richer semantic representations, such as RDF or JSON-LD, remains limited. The Reusable score was also lower because information on machine-readable licenses, provenance, and community-specific metadata standards was not consistently available.

E. Research Datasets

At the indicator level, research datasets received very low ratings across all four FAIR domains. The Findable domain achieved a normalized score of 3.1%, reflecting only limited evidence that dataset metadata were indexed by external search services. The assessment did not identify consistent implementation of persistent identifiers, standardized metadata schemas, or explicit links between metadata records and the corresponding datasets. The Accessible, Interoperable, and Reusable domains each achieved a normalized score of 0.0%. No documented open or controlled-access protocol for datasets was identified, and there was insufficient evidence that dataset metadata would remain accessible if the underlying data became unavailable. The assessment also found no consistent use of machine-readable formats, structured vocabularies, interoperable metadata schemas, explicit reuse licenses, provenance information, or community-specific metadata standards. These indicator-level gaps explain the very low dataset scores and indicate that institutional dataset management remains at an early stage. Priority improvements should therefore focus on persistent identifiers, standardized metadata, repository indexing, transparent access protocols, long-term metadata preservation, machine-readable formats, licensing, and provenance documentation.

The comparison confirms a substantial readiness gap between research articles and datasets. While article-level discovery and access mechanisms are well established, dataset services require fundamental improvements across all FAIR domains. The indicator-level differences are visualized in Fig. 3.

IV. DISCUSSION

A. Empirical Findings and Their Interpretation

- Empirical findings

The survey showed that 86.7% of respondents selected the two highest response categories for the importance of easy repository data access and use, while 87.6% selected the two highest categories for the need to improve the current data management system. The prioritized SWOT findings included institutional commitment as the principal strength, the adoption of emerging technologies as the principal opportunity, the need for additional RDM training and limited funding as the most actionable weakness-related findings, and government policy changes and rapid technological evolution as the leading threats.

The Convergence Matrix assessment showed that research articles achieved normalized scores of 100.0% for Findable, 100.0% for Accessible, 62.5% for Interoperable, and 45.8% for Reusable. Datasets achieved 3.1% for Findable and 0.0% for the other three domains. These values constitute the directly observed survey and expert-assessment results.

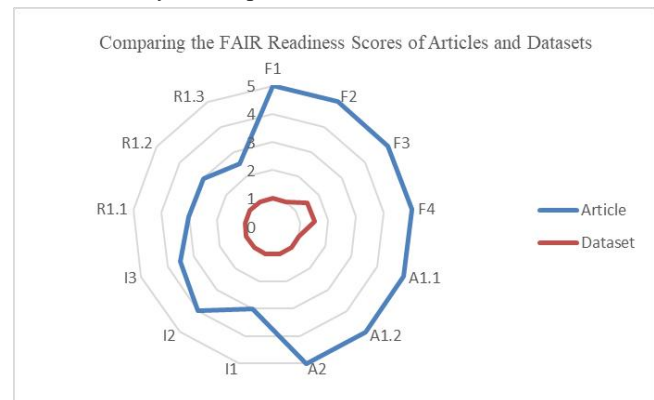


Fig. 3. Indicator-Level FAIR assessment scores for research articles and datasets.

Note: The plotted values represent the arithmetic mean of the two evaluators' ratings on a scale from 1 (*not fulfilled*) to 5 (*fully fulfilled*).

- Interpretation

Taken together, the findings suggest an asymmetry between organizational readiness and dataset-level FAIR implementation. Strong stakeholder demand and institutional commitment may provide an enabling basis for RDM development, but they should not be interpreted as evidence that the repository already complies fully with the FAIR principles. Similarly, the low dataset scores describe the repository features and supporting evidence examined during the assessment and should not be generalized to every dataset produced at IPB University. The convergence of the three analyses indicates that priority should be given to formal RDM governance, persistent identifiers, standardized metadata, transparent access protocols, licensing, provenance documentation, long-term preservation, and capacity development. The absence of clearly documented preservation mechanisms may create longer-term repository-sustainability risks [20], while the partial Interoperable and Reusable results are consistent with the importance of metadata harmonization for cross-system data use [21]. These interpretations are based on the observed patterns but do not establish causal relationships.

- Relevance of SWOT Analysis to FAIR Implementation

The SWOT results suggest that institutional commitment and technology adoption provide favorable conditions for FAIR-based RDM, whereas training needs, funding constraints, policy changes, and rapid technological evolution may limit implementation. Related studies have shown that RDM adoption depends on policy, researcher capabilities, incentives, and data-sharing practices, not only on repository technology [5], [6], [22]. The SWOT analysis therefore translates stakeholder perceptions into organizational priorities that complement the indicator-level FAIR assessment.

B. Comparison with Previous Studies

The findings are consistent with [22], who identified policy and FAIR literacy as important barriers to FAIR data sharing. Reference [23] likewise highlighted the importance of community participation in research data management, while [21] showed that metadata harmonization is central to effective cross-domain data integration and reuse. In the IPB context, the low dataset interoperability and reusability scores indicate related gaps in semantic representation, licensing, provenance, and community-specific standards. These comparisons reinforce the interpretation that technical repository improvements must be accompanied by clear governance and active stakeholder engagement.

C. Practical Implications for Universities

The proposed strategy has several practical implications for universities seeking to develop FAIR-based RDM. First, implementation should be treated as a phased institutional program rather than as a single repository-technology project. Universities can begin by assessing their organizational and technical gaps, establishing an institutional RDM policy, defining responsibilities, and selecting priority research units or data domains for initial implementation. Second, a minimum dataset service should include persistent identifiers, standardized metadata, explicit links between publications and datasets, transparent access classifications, reuse licenses, provenance information, and long-term preservation arrangements. Third, implementation requires coordinated responsibilities. University leaders should provide policy authority and sustainable resources; research offices should incorporate data-management requirements into research procedures; libraries should lead metadata curation, data stewardship, and researcher support; and information-technology units should maintain repository infrastructure, authentication, interoperability, backup, and security mechanisms. Researchers should be supported through practical deposit guidance, FAIR literacy programs, and incentives for responsible data sharing. For universities with limited resources, the SWOT–Convergence Matrix framework can be used to identify the most urgent organizational and FAIR-indicator gaps, allowing investments to be prioritized rather than attempting to implement all components simultaneously. The strategy should therefore be adapted to each university’s research profile, data sensitivity, technical capacity, governance structure, and funding conditions.

D. Integration of SWOT and Convergence Matrix as an Evaluation Approach

The integration of SWOT analysis and the GO FAIR Convergence Matrix provides complementary organizational and technical perspectives on institutional RDM readiness. SWOT analysis identifies internal and external factors that may influence implementation, whereas the Convergence Matrix provides an indicator-level assessment of observable repository features. Their integration supports a structured transition from diagnosis to strategy prioritization without treating organizational commitment as equivalent to technical FAIR fulfillment. The framework therefore represents a context-sensitive evaluation approach that may be adapted and

tested in other universities rather than a universally validated maturity model.

E. Visualization of Findings and Their Interconnections

Figure 3 complements Tables 8 and 9 by visualizing the indicator-level contrast between research articles and datasets. The radar chart shows consistently higher article scores, particularly for Findable and Accessible indicators, while dataset scores remain close to the minimum across the assessed indicators. The visualization supports interpretation of the readiness gap but does not replace the indicator-level evidence and domain-level calculations reported in the tables.

V. CONCLUSION

This study identified strong stakeholder demand for accessible repository data and improved institutional data management, with 86.7% and 87.6% of respondents, respectively, selecting the two highest response categories. The SWOT analysis identified institutional commitment as the principal strength and emerging-technology adoption as the principal opportunity, while additional RDM training, limited funding, government policy changes, and rapid technological evolution emerged as priority concerns. The FAIR assessment revealed a substantial gap between research articles and datasets. Articles achieved normalized scores of 100.0% for Findable, 100.0% for Accessible, 62.5% for Interoperable, and 45.8% for Reusable, whereas datasets achieved 3.1% for Findable and 0.0% for the other three domains. By integrating needs analysis, SWOT analysis, and the GO FAIR Convergence Matrix, the study links institutional conditions with indicator-level FAIR readiness and supports the prioritization of FAIR-based RDM interventions.

The theoretical contribution of this study lies in conceptually linking stakeholder needs, organizational conditions, and indicator-level FAIR readiness within a single institutional RDM evaluation framework. This integration distinguishes organizational willingness and capacity from the actual implementation of FAIR requirements and shows how internal and external factors may shape the prioritization of repository interventions. Methodologically, the framework combines needs analysis, SWOT analysis, and the GO FAIR Convergence Matrix as complementary stages of diagnosis and strategy formulation. The practical contribution is a structured assessment and prioritization approach that universities can use to identify urgent governance, capacity, metadata, access, interoperability, licensing, provenance, and preservation needs. The framework can support phased RDM planning and resource allocation, although it should be adapted to the institutional context rather than interpreted as a universally validated FAIR maturity model.

Based on the identified organizational and FAIR-readiness gaps, the immediate institutional priorities are to strengthen formal RDM governance, FAIR literacy and data-stewardship support, dataset-deposit procedures, metadata standardization, transparent access mechanisms, licensing, provenance documentation, interoperability, and long-term preservation.

This study has several limitations. First, it was conducted

within a single university using a purposive stakeholder sample; therefore, the survey findings should not be interpreted as statistically representative of all IPB University stakeholders or directly generalized to other institutions. Second, the Convergence Matrix assessment relied on two expert evaluators with direct operational and technical knowledge of the repository. Although their ratings were compared with observable repository features and available supporting evidence, the small evaluator panel may not capture the full range of institutional perspectives. Third, the survey indicators were analyzed descriptively at the item level, and the study did not test causal relationships between organizational factors and FAIR readiness. Finally, the repository assessment represents conditions observed during a specific study period and does not evaluate longitudinal changes or the effects of subsequent policy and system interventions. The findings should therefore be interpreted as a diagnostic institutional assessment and a basis for prioritization rather than as a universally generalizable maturity measurement.

Future research should validate the integrated framework across multiple universities with different research profiles, governance structures, and levels of repository maturity. Larger and more diverse stakeholder samples should be used to examine whether the identified organizational priorities remain consistent across institutional contexts. The Convergence Matrix assessment should also involve a broader panel of technical, library, research-management, and disciplinary experts, accompanied by formal inter-rater agreement or consensus procedures. Longitudinal studies are needed to apply the proposed priorities as institutional interventions and compare FAIR readiness before and after the introduction of RDM policies, training, metadata services, and repository improvements. Future studies should also test the framework using deposited datasets from different disciplines and complement expert assessment with automated FAIR-evaluation tools and operational indicators, such as metadata completeness, deposit rates, access-request completion, dataset reuse, citation, and long-term availability. These studies would help determine the framework's transferability, sensitivity to institutional change, and usefulness for evaluating actual RDM implementation outcomes.

ACKNOWLEDGMENT

The authors thank Prof. Dr. Yeni Herdiyeni, S.Kom., M.Kom., for her ideas, critical input, methodological guidance, and substantive perspectives, which strengthened the problem formulation, research design, data collection, interpretation of results, and refinement of the manuscript.

REFERENCES

- [1] E. E. Baro and A. U. Nwabueze-Echedom, "An evaluation of institutional repository development in African universities," *IFLA J.*, vol. 49, no. 1, pp. 18–38, 2023, doi: 10.1177/03400352221089672.
- [2] M. K. Singh and G. T. K. Kumar, "Research data curation in academic institutions: Challenges and expectations," *DESIDOC J. Libr. Inf. Technol.*, vol. 43, no. 1, pp. 39–44, 2023, doi: 10.14429/djlit.43.01.18624.
- [3] F. A. Gómez and I. Bernal, "FAIR EVA: Bringing institutional multidisciplinary repositories into the FAIR picture," *Sci. Data*, vol. 10, Art. no. 764, 2023, doi: 10.1038/s41597-023-02652-8.
- [4] R. David *et al.*, "FAIRness literacy: The Achilles' heel of applying FAIR principles," *Data Sci. J.*, vol. 19, Art. no. 32, 2020, doi: 10.5334/dsj-2020-032.
- [5] N. F. Mosha and P. Ngulube, "The utilisation of open research data repositories for storing and sharing research data in higher learning institutions in Tanzania," *Libr. Manag.*, vol. 44, nos. 8–9, pp. 566–580, 2023, doi: 10.1108/LM-05-2023-0042.
- [6] R. Vrana, "Research data management in the Croatian academic community: A research study," *Publications*, vol. 12, no. 2, Art. no. 16, 2024, doi: 10.3390/publications12020016.
- [7] E. Marlina, A. N. Hidayanto, and B. Purwandari, "Towards a model of research data management readiness in Indonesian context: An investigation of factors and indicators through the fuzzy Delphi method," *Libr. Inf. Sci. Res.*, vol. 44, no. 1, Art. no. 101141, 2022, doi: 10.1016/j.lisr.2022.101141.
- [8] E. K. Donner, "Research data management systems and the organization of universities and research institutes: A systematic literature review," *J. Librariansh. Inf. Sci.*, vol. 55, no. 2, pp. 261–281, 2023, doi: 10.1177/09610006211070282.
- [9] Y. Minamiyama, U. Ikeuchi, K. Tanabe, K. Yuki, K. Hayashi, and T. Aoki, "Maturity model for organizational research data management services," *Data Sci. J.*, vol. 24, Art. no. 18, pp. 1–18, 2025, doi: 10.5334/dsj-2025-018.
- [10] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," *MIS Q.*, vol. 28, no. 1, pp. 75–105, 2004, doi: 10.2307/25148625.
- [11] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *J. Manag. Inf. Syst.*, vol. 24, no. 3, pp. 45–77, 2007, doi: 10.2753/MIS0742-1222240302.
- [12] Z. He and W. Fang, "Research data management in institutional repositories: An architectural approach using data lakehouses," *Digit. Libr. Perspect.*, vol. 41, no. 1, pp. 145–178, 2025, doi: 10.1108/DLP-02-2024-0022.
- [13] N. N. K. Krisnawijaya, B. Tekinerdogan, C. Catal, R. van der Tol, and Y. Herdiyeni, "Implementing FAIR principles in data management systems: A multi-case study in precision farming," *Comput. Electron. Agric.*, vol. 230, Art. no. 109855, 2025, doi: 10.1016/j.compag.2024.109855.
- [14] A. Jacobsen *et al.*, "FAIR principles: Interpretations and implementation considerations," *Data Intell.*, vol. 2, nos. 1–2, pp. 10–29, 2020, doi: 10.1162/dint_r_00024.
- [15] K. W. Lam, A. Hassan, T. Sulaiman, and N. Kamarudin, "Evaluating the face and content validity of an instructional technology competency instrument for university lecturers in Malaysia," *Int. J. Acad. Res. Bus. Soc. Sci.*, vol. 8, no. 5, pp. 367–385, 2018, doi: 10.6007/IJARBS/v8-i5/4108.
- [16] D. F. Polit and C. T. Beck, "The content validity index: Are you sure you know what's being reported? Critique and recommendations," *Res. Nurs. Health*, vol. 29, no. 5, pp. 489–497, 2006, doi: 10.1002/nur.20147.
- [17] D. F. Polit, C. T. Beck, and S. V. Owen, "Is the CVI an acceptable indicator of content validity? Appraisal and recommendations," *Res. Nurs. Health*, vol. 30, no. 4, pp. 459–467, 2007, doi: 10.1002/nur.20199.
- [18] M. Tavakol and R. Dennick, "Making sense of Cronbach's alpha," *Int. J. Med. Educ.*, vol. 2, pp. 53–55, 2011, doi: 10.5116/ijme.4dfb.8dfd.
- [19] K. A. Bollen and R. Lennox, "Conventional wisdom on measurement: A structural equation perspective," *Psychol. Bull.*, vol. 110, no. 2, pp. 305–314, 1991, doi: 10.1037/0033-2909.110.2.305.

-
- [20] D. Strecker, H. Pampel, R. Schabinger, and N. L. Weisweiler, "Disappearing repositories: Taking an infrastructure perspective on the long-term availability of research data," *Quant. Sci. Stud.*, vol. 4, no. 4, pp. 839–856, 2023, doi: 10.1162/qss_a_00277.
- [21] T. Cernava *et al.*, "Metadata harmonization—Standards are the key for a better usage of omics data for integrative microbiome analysis," *Environ. Microbiome*, vol. 17, no. 1, Art. no. 33, 2022, doi: 10.1186/s40793-022-00425-1.
- [22] L. Rainey, J. E. Lutonski, and M. J. Broeders, "FAIR data sharing: An international perspective on why medical researchers are lagging behind," *Big Data Soc.*, vol. 10, no. 1, 2023, doi: 10.1177/20539517231171052.
- [23] C. Schmidt, J. Hanne, J. Moore, C. Meesters, E. Ferrando-May, and S. Weidtkamp-Peters, "Research data management for bioimaging: The 2021 NFDI4BIOIMAGE community survey," *F1000Research*, vol. 11, Art. no. 638, 2022, doi: 10.12688/f1000research.121714.2.