Learned Publishing

LEARNED PUBLISHING Open Access

ORIGINAL ARTICLE OPEN ACCESS

Publish-Review-Curate Modelling for Data Paper and Dataset: A Collaborative Approach

Youngim Jung[1,2] | Sungsoo Robert Ahn[3,4]

[1]Korea Institute of Science and Technology Information, Seoul, South Korea | [2]University of Science and Technology, Seoul, South Korea | [3]Department of Software Engineering, Gyeongsang National University, Jinju, South Korea | [4]Gyeongnam Aerospace & Defence Institute of Science and Technology, Gyeongsang National University, Jinju, South Korea

Correspondence: Sungsoo Robert Ahn (sungsoo.ahn@gnu.ac.kr)

Received: 29 October 2024 | **Revised:** 10 July 2025 | **Accepted:** 18 August 2025

Funding: This work was supported by the Korea Institute of Science and Technology Information (K25L3M1C2; J25JR040-25), the New Faculty Research Support Grant from Gyeongsang National University, GNU-NFRSG-0121, in 2023, and the Glocal University 30 Project Fund of Gyeongsang National University, GNU-GlocalUniv-30, in 2025.

ABSTRACT

Research datasets—capturing natural, societal, or artificial phenomena—are critical in generating new scientific insights, validating research models, and supporting data-intensive discovery. Data papers that describe and contextualise these datasets aim to ensure their findability, accessibility, interoperability, and reusability (FAIR) while providing academic credit to data creators. However, the peer review of data papers and associated datasets presents considerable challenges, requiring reviewers to assess both the syntactic and semantic integrity of the data, metadata quality, and domain-specific scientific relevance. Furthermore, the coordination between journal editors, reviewers, and curators demands substantial effort, often leading to publication delays in the conventional review and then publishing framework. This study proposes a novel Publish-Review-Curate (PRC) model tailored to the synchronised publication and review of data papers and their underlying datasets. Building on preprint and open science practices, the model defines a collaborative, multi-stakeholder workflow involving authors, peer reviewers, data experts, and journal editors. The PRC model integrates open feedback, transparent peer review, and structured curation to improve research data's quality, discoverability, and impact. By articulating conceptual and operational workflows, this study contributes a practical framework for modernising data publishing infrastructures and supporting the co-evaluation of narrative and data artefacts.

1 | Introduction

Research on natural, societal, or artificial phenomena increasingly uses datasets to show the validity of new research ideas, analysis, experimentation, or models (Simon 2019). The more researchers understand and access trusted, well-curated, publicly available, and high-quality datasets in the research, the more researchers can apply the research dataset to their study. In this regard, data papers describing essential characteristics of research datasets are valuable sources to understand natural, social, or artificial phenomena, advance research, and validate scientific knowledge (Callaghan et al. 2012; Clifton-Ross et al. 2019). The data paper also gives academic incentives to dataset creators. The current method of reviewing and then publishing data papers has been helpful in making the research dataset findable, accessible, interoperable, and reusable (FAIR) for researchers in the community of data publishing.

However, the traditional reviewing and publishing model for data papers and associated datasets remains underdeveloped and inconsistent across academic journals. As Chavan et al. (2013) point out, the peer-review process for data publishing lacks shared standards and maturity, underscoring reviewer expertise and consistency challenges. While peer

Summary

- Publishing data papers and associated datasets presents considerable challenges, requiring reviewers to assess both the syntactic and semantic integrity of the data and substantial effort in the coordination between journal editors, reviewers, and curators.
- This study presents a conceptual model for the synchronized publication and review of data papers and datasets.
- This study models a detailed workflow outlining the roles of authors, peer reviewers, repository curators, and journal editors.
- We conduct a comparative analysis between the presented PRC model and several influential frameworks.
- This study contributes a practical framework for modernizing data publishing infrastructures and supporting the co-evaluation of narrative and data artifacts.

reviewing data papers and datasets, reviewers need to assess the metadata quality, data collection method, and potential value of the described dataset in the research domain (Mayernik et al. 2015). Bozada et al. (2021) emphasise that reviewers are responsible for assessing the dataset's syntactic structure and compliance with FAIR principles to ensure long-term integrity and reusability. Reviewers are expected to critically assess the scientific rigour of the methodologies employed in dataset collection, management, and processing. This includes evaluating the overall study design, the accuracy of citations, and other relevant scholarly standards. Following peer review, editors are responsible for verifying the integrity of the data paper and its associated dataset and determining the manuscript's suitability for publication. Upon acceptance, the data paper undergoes a final production process before it is formally published and made publicly accessible. The intensive intellectual effort and extended duration required to evaluate data papers and their associated datasets can delay publication, dissemination, access, and reuse. Moreover, the complexity of this process may contribute to suboptimal review practices—such as omitting the dataset review, neglecting to assess the integrity of both the data paper and dataset, and ultimately approving the publication of flawed or inaccurate data. These issues compromise the reliability of published outputs, impede dataset reusability, and hinder the broader progress of scientific research.

Although the current review and publish framework in research journals include closed, open, or hybrid peer review methods, a new publishing model, called Publish-Review-Curate (PRC), has recently emerged to resolve some issues previously described (Corker et al. 2024; Jeong et al. 2022; Eisen et al. 2020a). Some research journals have adopted the PRC model to rapidly review research papers, share the research work, and make the review process more transparent.

This study investigates data publishing workflows informed by the PRC framework, aiming to enhance the visibility of research data and expedite scholarly communication surrounding it. Specifically, this paper presents essential concepts of the workflows, a novel PRC workflow in reviewing data papers and datasets, and participants' roles in the proposed model. The model also shows a detailed collaboration process and data exchange between the reviewers of the data journals and the technical staff of the data repository.

The remainder of this paper is organised as follows. In Section 2, we review prior research on data publishing to contextualise current practices and supporting infrastructures. Section 3 introduces a novel workflow for the collaborative dissemination, quality control, and curation of data papers and associated datasets, grounded in the Publish–Review–Curate (PRC) framework. The Discussion section presents a comparative analysis between the proposed model and existing frameworks. It highlights its contributions and limitations and outlines key challenges that warrant further investigation for its practical implementation. Finally, the Conclusion summarises the main findings and proposes future research and development directions.

2 | Related Work

2.1 | Motivation of Data Publishing and Peer Review

Research datasets are foundational to scientific discovery across natural, social, and artificial domains (Smith 2009). As the Open Science movement expands, so too does the demand for trusted, high-quality datasets that are openly available and reusable (Callaghan et al. 2012; Conrad et al. 2024). Data publishing validates datasets through peer review and provides formal citation mechanisms such as DOIs, enhancing dataset traceability and version control (Callaghan et al. 2013).

Peer review serves as a cornerstone of scholarly communication, ensuring rigour, originality, and credibility. Integrating datasets into this process—called data peer review—increases trust and promotes transparency (Mayernik et al. 2015). Borgman (2015) highlighted the importance of legitimising datasets as scholarly outputs and raised questions about when peer review should occur in the data lifecycle. Mayernik et al. (2015) categorised data peer review across four venues: (1) datasets in traditional articles, (2) standalone data articles in journals, (3) datasets submitted to open-access repositories, and (4) articles in data journals.

Conferences on data science and artificial intelligence, such as NeurIPS and ACM Multimedia (2025), now include dataset-specific tracks because datasets play critical roles in machine learning. However, this remains uncommon, and such conferences are not treated as general data publication venues in this study.

2.2 | Data Repository and Their Role in Data Publishing

Repositories are central in preserving research data and enabling FAIR (Findable, Accessible, Interoperable, Reusable) principles. Upon deposition, datasets are assigned persistent identifiers like DOIs, supporting reproducibility and data reuse

(Asok, Dandpat, et al. 2024). Public repositories such as Zenodo, Harvard Dataverse, Figshare, and Dryad offer services, including metadata checks, version management, and access controls (Hansson and Dahlgren 2022). While most offer free services, some impose fees for large datasets.

Domain-specific repositories like Global Biodiversity Information Facility (GBIF), Ocean Biodiversity Information System (OBIS), and PANGAEA support specialised publication workflows and often collaborate with journals (Costello and Wieczorek 2014). Assante et al. (2016) evaluated repositories using eight publishing criteria, noting their utility for data sharing but recommending enhancements to address evolving research needs.

2.3 | Journals as a Data Publishing Venue

Depositing datasets alone does not guarantee academic credit or quality assurance. In contrast, data papers—scholarly articles describing datasets—are indexed, cited, and reviewed; thus, they integrate into the scholarly communication ecosystem (Wang and Xu 2025). These papers support dataset credibility and reuse while enhancing visibility and citation potential (Callaghan et al. 2012; Costello et al. 2013; Marsolek et al. 2023; Poline 2019; Chavan and Penev 2011).

Publishers increasingly require or encourage data sharing alongside research articles (Jackson 2021; Asok, Gupta, and Shrivastava 2024), and many now offer dedicated data journals. For example, Elsevier and Ubiquity Press publish journals focused on datasets, software, and protocols, while Springer Nature and Wiley also support peer-reviewed data publications (Elsevier 2024, Scientific Data 2024, Jung et al. 2020; ESA 2024).

However, traditional peer review of data papers remains time-consuming and under-resourced. Editors and reviewers often lack training or clear guidelines for assessing datasets (Mayernik et al. 2015); leading to delays in publication and inconsistent review quality.

2.4 | Cross-Linking Between Data Repositories and Journals

Efforts to bridge repositories and journals aim to enhance data discoverability and integrity. Lawrence et al. (2011) proposed a cross-linking model for integrating datasets and related articles, identifying best practices and system-level bottlenecks in geoscience publishing. Callaghan et al. (2013) also examined workflows that align repositories and data journals. This study documents the workflow used in the British Atmospheric Data Centre (BADC) and geoscience journals, highlighting procedures to formalise data publication in collaboration with data centres and scientific journals.

Ahn et al. (2023) identified challenges for academic societies in understanding data paper structure, conducting peer review, and selecting trusted repositories. The PREPARDE project further explored repository requirements and peer review models via multi-stakeholder workshops (Callaghan et al. 2014). Despite growing awareness, implementation remains limited due to unclear roles and reliance on traditional publishing workflows.

The RDA-WDS Data Publishing Workflows group develops a comprehensive reference model for data publishing workflows. It examines various existing workflows across disciplines and institutions to identify common components and standard practices in data publishing (Bloom et al. 2015). Its revised model for data publishing was synthesised from best practices across disciplines, focusing on interoperability between systems and stakeholders (Austin et al. 2017). Their model includes modular components such as submission, quality assurance, peer review, registration, preservation, and dissemination and describes stakeholder roles spanning data producers, repositories, journals, publishers, and aggregators. Metadata enrichment, persistent identification, and provenance tracking are emphasised to reflect the complexity and flexibility of real-world data publishing workflows.

This work responds to that gap by proposing a collaborative data publishing workflow. The discussion section presents detailed comparisons with existing modes suggested in the above studies.

2.5 | Rise of Preprints and the PRC Model

The proliferation of preprints has driven new scholarly communication models. At eLife, 70% of submissions are already available on preprint servers such as bioRxiv, arXiv, or medRxiv (Eisen et al. 2020a; eLife 2022). The Publish-Review-Curate (PRC) model has emerged: research is shared as a preprint, reviewed post-publication, and then curated (cOAlition S 2023). Platforms such as Rapid Review, Biophysics Colab, and GigaByte have adopted this approach (Edmunds 2023).

The PRC model accelerates dissemination, promotes transparency, and supports community-led peer review. It aligns with FAIR principles and allows researchers to claim priority while receiving feedback (Tennant 2019; Kraker 2011; Chiarelli 2019; Moher et al. 2018). Ross-Hellauer (2017) argues that open peer review reduces bias; Bender and Friedman (2018) note that post-publication feedback enhances rigour.

Despite its benefits, the PRC model faces critiques around quality control, reviewer accountability, and compatibility with academic incentives. Immediate dissemination of preprints can lead to the spread of unvetted claims, particularly in sensitive fields like biomedicine (Berg et al. 2024). Sokolova (2016) warns that emphasising novelty and popularity may undermine review integrity.

3 | Methods

This study builds on recent developments in scholarly communication by proposing a *Publish–Review–Curate (PRC)*-based workflow tailored specifically for data papers and datasets. Our goal is to design and evaluate a workflow that enables rapid dissemination, transparent peer review, and structured post-publication curation—core principles that support the goals of Open Science. We create high-level and operational PRC

workflow models using Business Process Model and Notation (BPMN) standards. The high-level workflow outlines the sequential stages: initial preprint publishing, open peer review, and post-review curation. The detailed operational workflow maps inter-role communications, tool integrations (e.g., preprint servers and repositories), and review data flow.

3.1 | Conceptual Publish-Review-Curate Model of Data Paper and Dataset

We describe essential concepts in our proposed workflow. Figure 1 shows essential concepts and their relationships to other concepts in our peer review approach with PRC for data papers and datasets.

The author prepares a data paper and a dataset for publication in data journals during the data paper publishing. The author then deposits a data paper and a corresponding dataset as a preprint to the data repository. The general public or interested researchers may comment on the preprint in a data repository. After the author submits the data paper to the data journal, the editor decides to review the submitted paper and dataset. The editor then invites reviewers to review the preprinted data paper, providing the review criteria managed at the data journal. In collaboration with the data repository, the editor also invites technical data experts to review the deposited dataset using the review criteria. After both the data paper and dataset review, the data journal editor decides on the publication of the data paper and curates the data paper as a version of the record in the Journal issue and data repository. Other researchers interested in the dataset may use or refer to the published paper and the dataset for their own research.

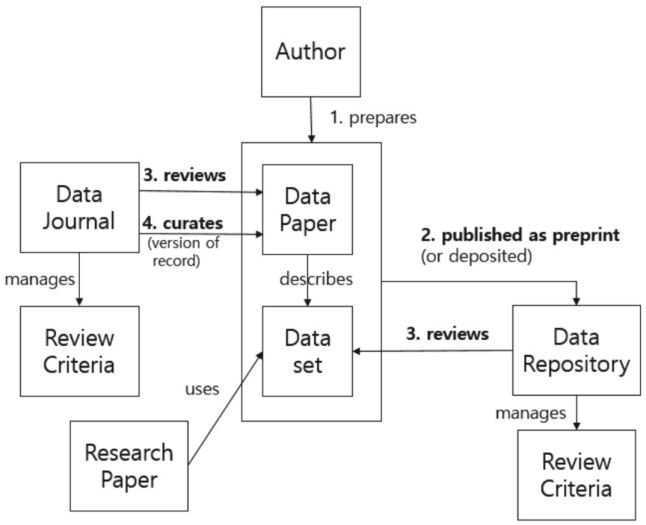

FIGURE 1 | Conceptual model of applying PRC.

3.2 | High-Level Publish-Review-Curate Workflow for Data Paper and Dataset

This section shows the high-level workflow of our approach to adopting the PRC model. While Figure 2a shows the PRC model for research paper view at the high level, Figure 2b shows our proposed data publishing workflow for the data paper and dataset incorporating the PRC model. The proposed workflow deals with both data papers and corresponding datasets. Both submitted data papers and datasets are deposited to the data repository and reviewed by different experts.

In Figure 2b, the author first publishes the data paper as a preprint and deposits the dataset in the data repository. When the author submits the preprint to the data journal, the editor decides the editorial review for the submitted preprint and dataset. If the journal editor decides to review the preprint and dataset, invited reviewers will review the data paper and dataset from both scientific and technical perspectives. They will then write a review comment and make recommendations for editorial decisions. The data journal editor then collects and evaluates the review reports and decides on publication. If the preprint is accepted for publication, the version of the

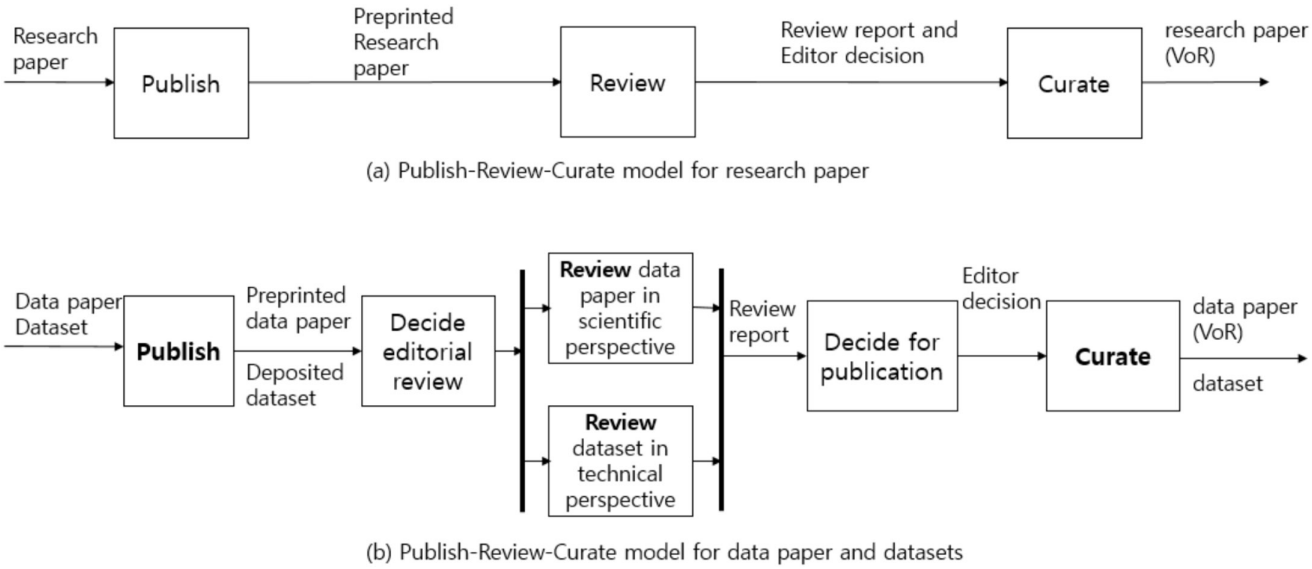

FIGURE 2 | Publish-review-curate model comparison for research and data paper.

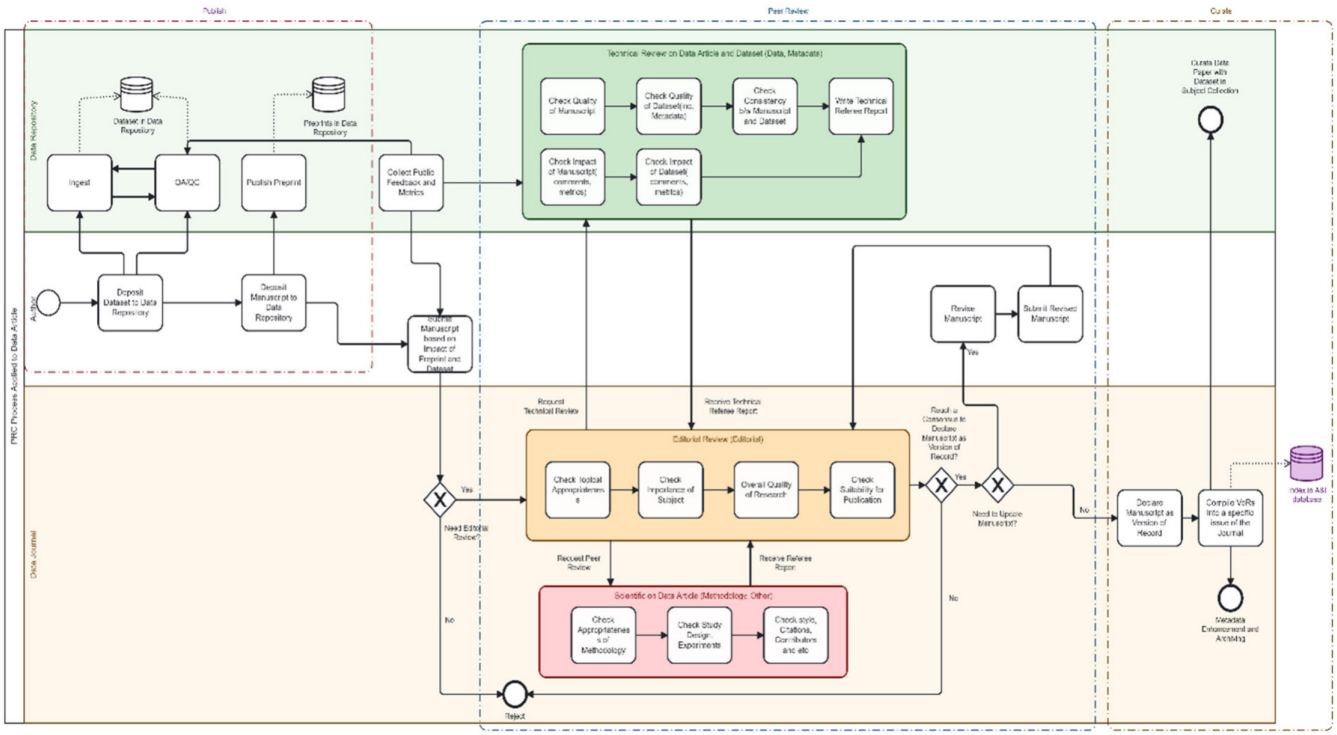

FIGURE 3 | Peer review workflow of data paper using the PRC model.

record (VoR) of the data paper is compiled in the Journal Issue and curated in the data repository. The dataset remains in the data repository but may be updated to incorporate the review comments.

In the following sections, we describe each publish, review, and curate step in more detail.

3.3 | Data Publishing Workflow Based on the Publish-Review-Curate Model

In this section, we delineate the step-by-step process of publishing data papers and datasets in the PRC model, which serves as the foundation for the collaborative data paper peer review and publication. By breaking down the model into discrete stages, we aim to provide a comprehensive overview of the methodologies and a set of criteria, ensuring that the process is transparent and replicable. This structured approach facilitates a deeper understanding of the PRC model of data paper. It enables its application in a wide range of scenarios within the collaborative approach between the data repository and other parties in scholarly communication. Figure 3 illustrates an overview of the workflow using the PRC model in data paper peer review, publishing, and curation.

3.3.1 | Publish Stage

The *publish stage* is the initial phase, where the dataset and its accompanying data paper are made publicly available for the first time, as shown in Figure 4. This stage is designed to facilitate the rapid dissemination of research data while ensuring that both the dataset and the data paper are accessible and open for early feedback. The key components of the *publish stage* are as follows:

- **Submission of Dataset and Preprint**: Authors submit both the dataset and the preprint of the data paper to a data repository. The repository is responsible for conducting an initial quality control check, which includes verifying the integrity of the dataset, the completeness of its metadata, and its alignment with repository guidelines.
- **Quality Control and Preprint Publication**: After passing the initial checks, the dataset is assigned a persistent identifier, such as a Digital Object Identifier or DOI, and the manuscript is published as a preprint. This early publication allows the research community to access the dataset and paper quickly, fostering immediate feedback and engagement from peers.
- **Public Feedback and Metrics Collection**: Once the preprint is published, it becomes open to public feedback. Researchers in the field can comment on the dataset's quality, relevance, and potential applications by referring to the preprint of the data paper describing the dataset. Additionally, usage metrics such as downloads, citations, and altmetrics are tracked to gauge the impact and reach of the dataset and the preprint.

The publish stage is critical for enabling rapid access to research datasets and data papers, promoting transparency, and allowing for the early identification of potential issues or improvements. By making datasets and preprints of data papers available early in the process, the PRC model encourages broader participation from the research community,

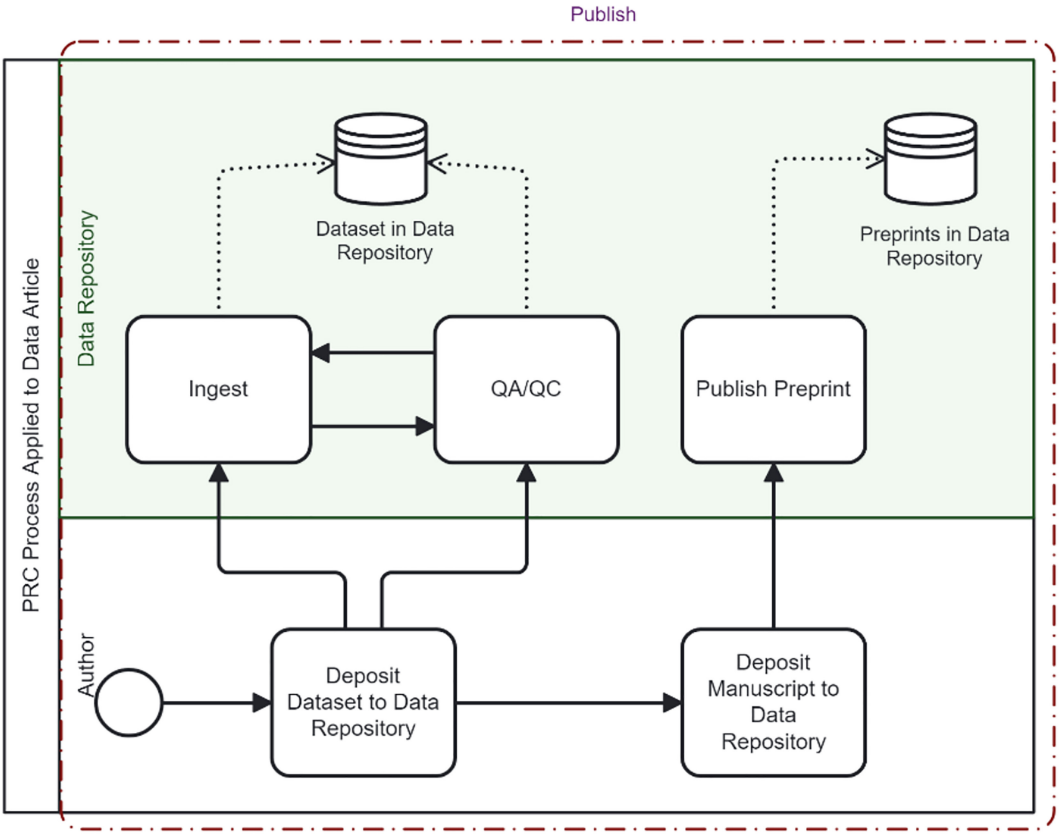

FIGURE 4 | Workflow of publishing the preprint of the data paper and corresponding dataset.

ultimately leading to higher-quality datasets and more robust scientific outcomes.

3.3.2 | Review Stage

Authors submit the preprint to a data journal for formal peer review based on the feedback received and the metrics collected. If the dataset and the preprint have garnered sufficient positive attention, this can expedite the peer review process and enhance the chances of acceptance by a data journal. The review of the data paper and dataset to which the data paper refers is performed in the data journal in collaboration with the data repository where the dataset is located. Formal peer review consists of an editorial, technical, and scientific review. During this review, journal editors do the editorial review, data experts evaluate the technical features and quality of the dataset in the data repository, and reviewers from the subject field review the data paper.

Figure 5 illustrates the workflow of the collaborative peer review between the data journal and data repository. Once a manuscript is submitted, it first undergoes the editorial review stage before progressing to technical and scientific review or being rejected.

Carpenter (2017) examined five categories of common review criteria for reviewing data papers and datasets across 39 data journals and examined whether they were included in the peer review policy. Based on Carpenter's research, our study reused Carpenter's review criteria and categorised them into factors that should be considered at the editorial review, technical review, and scientific review stages.

1. Editorial review

The editorial review process within the PRC model for data papers, illustrated in the third lane of Figure 5, is a critical step that ensures the manuscript's suitability for publication. This process involves a thorough evaluation by journal editors, who assess the manuscript against a set of criteria outlined in Table 1, which include topical appropriateness, importance of the subject, originality, and overall research quality. The editor-in-chief evaluates whether the manuscript fits the journal's scope and meets the required standards for publication. If the manuscript passes this stage, the editor may either move it to technical and scientific review, request revisions, or reject it outright.

The components listed in Table 1 guide the editorial review by providing specific evaluation points. Through these criteria, the editorial review determines the manuscript's readiness for further evaluation or its need for improvement, making it a crucial gatekeeper in the PRC model.

2. Technical Review of dataset and data paper

The technical review within the PRC model plays a crucial role in ensuring the integrity and quality of the dataset associated with a data paper. This review process, depicted in the first lane of Figure 5, involves a comprehensive examination of both the dataset and its accompanying metadata by data experts

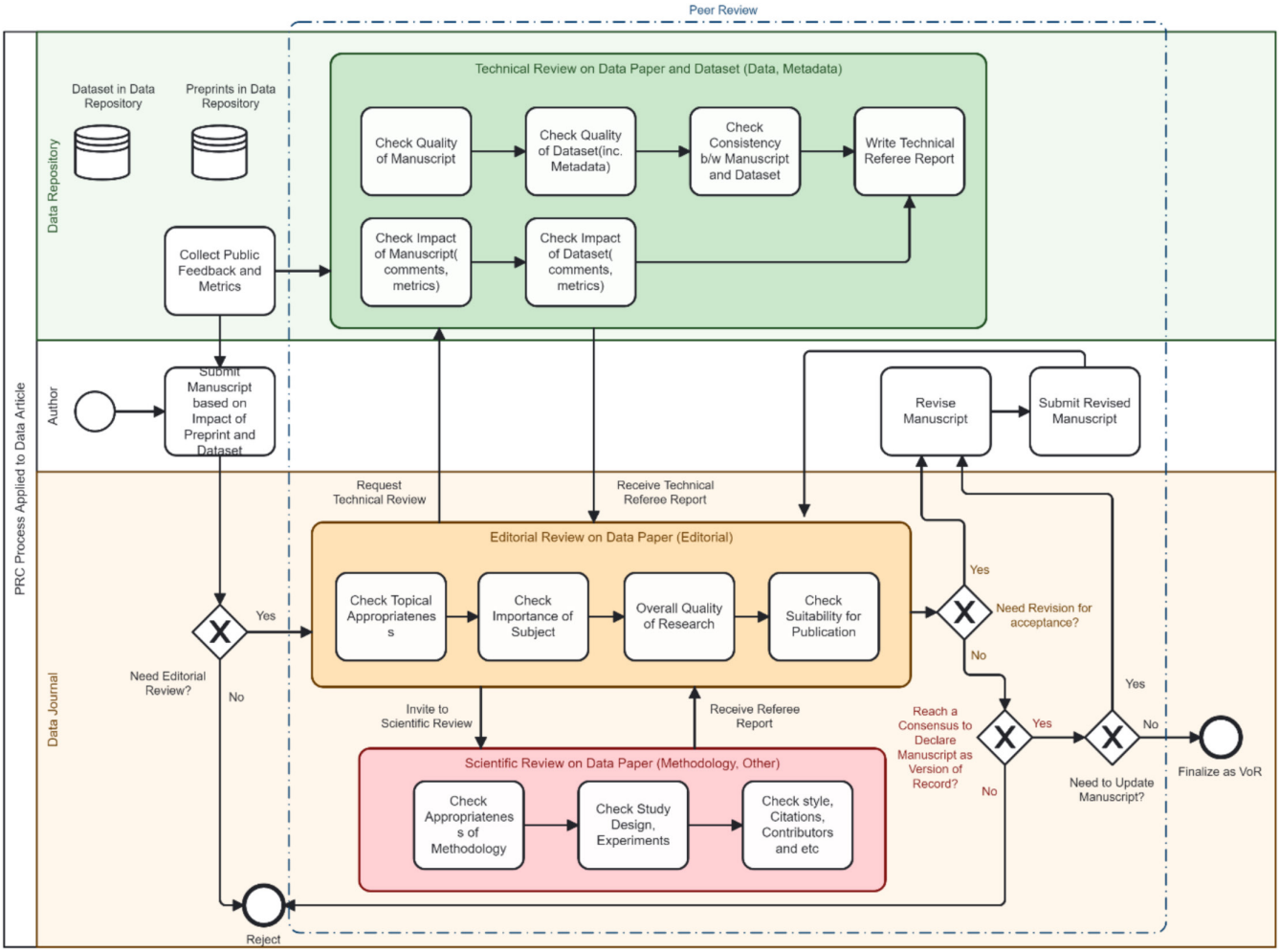

FIGURE 5 | Workflow of collaborative peer review of data paper.

TABLE 1 | Constitutions of editorial review.

Constitutions of editorial review	Note
Open Peer Review	Involves a review process where the identities of both reviewers and authors are known to each other, promoting transparency and accountability in the review process.
Conflict of Interest Policy	Ensures that there are no conflicts of interest among the authors that could bias the research or its presentation. It checks for any potential financial, personal, or professional conflicts that might affect the integrity of the work.
Topical Appropriateness in Title	Ensures that the manuscript's title accurately reflects the content of the paper and aligns with the journal's focus and thematic scope.
Suitability for Publication in Title	Assesses whether the manuscript's title indicates that the content is appropriate for publication within the specific journal, considering its audience and scope.
Importance of the Subject	Evaluates the relevance and significance of the research topic within its field, considering the potential impact and contribution to existing knowledge
Overall Quality of Research	Reviews the methodological rigour, clarity, and overall quality of the research, ensuring it meets the standards required for scholarly publication
Unpublished	Confirms that the research has not been published elsewhere and is original work submitted exclusively to the journal
Originality/Novelty	Evaluates whether the research offers new insights, concepts, or advances in the field, ensuring that the work contributes something novel to the existing body of knowledge

typically within a repository. These data experts focus on the technical aspects of the data to ensure the dataset is robust, well-documented, and suitable for broader use. The technical review includes several key stages:

- **Submission of Technical Review Request**: After the initial editorial review, the manuscript and the information of its dataset are forwarded to the technical reviewers for detailed technical assessment.
- **Evaluation of Metadata Quality**: Reviewers assess the metadata for completeness, accuracy, and adherence to standards. This step ensures that the dataset is well described and accessible for future users. In addition, the consistency between the metadata of the dataset and the manuscript is reviewed in this step.
- **Assessment of Data Logic and Consistency**: Reviewers check the dataset for logical coherence, ensuring that data points align correctly and that no inconsistencies could undermine the data's validity.
- **Verification of Data Format and Non-Proprietary Formats**: Reviewers confirm that the data is available in widely recognised, non-proprietary formats, facilitating reuse and long-term preservation.
- **Composition of Technical Referee Report**: The technical review concludes with preparing a detailed report, which is sent to the journal editor. This report summarises the findings of the technical review, highlighting any issues with the dataset or metadata that need addressing before publication. The editor then uses this report to make informed decisions about the manuscript's progression, including whether further revisions are required.

Table 2 provides a detailed breakdown of the specific criteria used during the technical review of a data paper, ensuring both the metadata and data meet high quality and usability standards.

By combining these rigorous assessments of the manuscript, the dataset, and its metadata, the technical review process ensures that the dataset is not only scientifically valuable but also technically sound, accessible, and reusable, which is essential for assessing the suitability of the manuscript to the publication in the journal.

3. Scientific review

The data journal's editor invites experts in the field to be reviewers. The reviewers may review the methodological adequacy of the data collection described in the submitted manuscript, the accuracy of the methodological design, the experimental design, the adherence to the research data collection method and publication ethics, and the preparation of the manuscript by the journal's publication guidelines; and send the results of the scientific review to the journal editor.

The scientific review process within the PRC model is a critical stage that ensures the data paper's methodological rigour and scholarly value. This process, represented in the third lane of Figure 5, is conducted by reviewers who thoroughly evaluate the manuscript's scientific content, including the dataset's methodological accuracy, experimental design, and adherence to research ethics. The scientific review includes several key stages:

- **Invitation to Review:** Journal editors invite experts in the field to review the data paper. These experts focus on the scientific integrity of the research described in the manuscript.
- **Methodological Assessment:** The reviewers assess the adequacy of the research methods used to collect the data, ensuring that the study design is appropriate for the research questions.
- **Ethics and Compliance Review:** The review also includes an evaluation of the study's adherence to ethical guidelines and compliance with open science principles, such as public data sharing under open licences.

Table 3 outlines the key criteria used during the scientific review.

Through these criteria, the scientific review ensures that the data paper not only presents high-quality data but also adheres to the best practices in scientific research. When combined with the technical review, the peer review process provides a robust evaluation that enhances the credibility and impact of the data paper.

3.3.3 | Curate Stage

The curate stage in the PRC model, as shown in Figure 6, is the final step in the data paper publication process. It focuses on organising, preserving, and making the dataset and its associated data paper accessible for future use. This stage is critical for ensuring that the published data and the corresponding article are both discoverable and reusable, fulfilling the principles of FAIR data management.

During the curate stage, the following activities typically occur:

- **Finalisation and Indexing:** Once the manuscript has passed both technical and scientific reviews, it is declared as the version of the record. The version of the record is compiled to a specific issue and volume of the data journal. Then, the data paper is indexed in Abstracting and Indexing (A&I) databases, ensuring that it is easily discoverable by other researchers.
- **Metadata Enhancement and Linking:** The metadata associated with the dataset and the data paper is enhanced and linked. This ensures that users can easily navigate between the data and the publication, facilitating better understanding and reuse of the research. The metadata may also be curated to ensure it meets the standards required for long-term preservation.
- **Repository Curation:** The data and the data paper might be curated within the data repository, which may involve organising the data into relevant collections or subject areas. This helps in contextualising the dataset within broader research themes and disciplines, making it easier for researchers to find and reuse the data.

TABLE 2 | Constitutions of technical review.

Constitutions		Note
Metadata	Metadata Quality	Focuses on the accuracy, clarity, and completeness of the metadata, ensuring it fully represents the dataset
	Metadata Presentation	Ensures that the metadata is presented in a clear, organised, and readable manner.
	Metadata Standards Conformance	Checks if the metadata conforms to recognised standards, which aids in interoperability and understanding across systems.
	Title/Abstract/Writing Clarity	Assesses the clarity and comprehensibility of the title, abstract, and overall writing, ensuring that they effectively communicate the content and purpose of the dataset.
	Key Selection	Ensures that key information is appropriately selected and highlighted within the metadata, making it easier to find and understand.
	Dataset DOI Assignment	Verifies that a DOI (Digital Object Identifier) is correctly assigned to the dataset, ensuring it can be easily cited and accessed.
	Metadata Rights information	Reviews the information on rights and usage terms provided in the metadata, ensuring it is clear and accurate.
	Provenance	Evaluates the documentation of the dataset's origins, its history, and any transformations it has undergone.
	Metadata Completeness	Ensures that the metadata is thorough and includes all necessary details to fully describe the dataset.
Data	Data Logic & Consistency	Ensures the dataset is logically structured and free of errors or inconsistencies
	Data Format & Consistency	Verifies that the dataset is stored in formats that are accessible and reusable across different platforms and applications
	Data – Non-Proprietary Formats	Ensures that the dataset is stored in formats that are not proprietary, promoting long-term accessibility and usability
	Data plausibility	Assesses the dataset to ensure the data appears reasonable and credible, considering the context and expected patterns
	Data of High Quality	Reviews the overall quality of the data, including accuracy, completeness, and reliability
	Data-Worthy of sub selection/broad enough/Appropriate selection	Evaluates whether the dataset covers an appropriate scope, ensuring it is comprehensive yet focused
	Data Units of Measure	Checks that units of measure are consistent, correctly used, and clearly documented throughout the dataset
	Data Quality Methods	Examines the methods used to ensure data quality, including processes for validation, cleaning, and verification
	Data Anonymization	Reviews the steps taken to anonymize data, ensuring that privacy and confidentiality are maintained where necessary
	Data Anomalies/outliers identified/treated/documented	Ensures that any anomalies or outliers in the data are identified, treated appropriately, and documented.
	Any data errors introduced technique, fact, calculation	Assesses whether any errors were introduced during the data collection, processing, or calculation stages, and how they were handled

- **Ongoing Access and Updates:** The curate stage also includes setting up mechanisms for ongoing access to the dataset and data paper. This could involve assigning persistent identifiers (e.g., DOIs) and ensuring that the dataset remains accessible over time. Additionally, any updates or corrections to the dataset are managed through this stage to maintain its relevance and accuracy.

The curate stage ensures that the dataset and the data paper continue contributing to the scientific community long after

TABLE 3 | Constitutions of scientific review.

Constitutions		Note
Methodology	Appropriate	Ensures that the methodology used in the study is suitable and well-suited to achieve the research objectives.
	Current	Confirms that the methods and techniques used are up-to-date with the latest standards and practices in the field.
	Data Collection Methods	Reviews the methods employed for data collection, ensuring they are robust, accurate, and relevant to the study's goals.
	High Technical Standard	Assesses whether the research meets high technical standards in terms of accuracy, precision, and reliability.
	Equipment Description	Ensures that all equipment and tools used in the study are described in detail, including specifications and settings.
	Replicable	Verifies that the methodology is described in sufficient detail to allow other researchers to replicate the study.
	Quality Control	Evaluates the processes in place to ensure the quality and integrity of the data collected and analysed.
	Supporting Experiments	Assesses whether additional experiments or tests were conducted to support the main findings and how they were integrated into the study.
	Study design—Overall, Replications Identified	Reviews the overall design of the study, ensuring that replications (if any) are clearly identified and justified.
	Study design—Dependency, Data preprocessing described	Examines the study design to ensure that dependencies and data preprocessing steps are well-described and appropriate for the research.
Other	Citations to Other Relevant Materials	Ensures that the manuscript includes appropriate citations to other relevant studies and materials that support the research.
	Fairness	Evaluates whether the study was conducted and reported in an unbiased and fair manner, without undue influence or partiality.
	Anonymity of Reviewers (if desired)	Considers the anonymity of the reviewers, ensuring that the review process remains impartial if the reviewers choose to remain anonymous.
	Ethics of Experimentation	Reviews the ethical considerations and approvals related to the study, ensuring that the research complies with ethical standards and guidelines.
	Public Data Sharing, Open Licence Requirement	Confirms that data is shared publicly and that it is done under an open licence, allowing others to freely use and build upon the data.
	Data Repository with a Sustainability Model	Ensures that the data is stored in a repository that has a sustainability model, ensuring long-term access and preservation.
	Data Sharing—Platform Agnostic	Verifies that data can be accessed and used across different platforms and systems without compatibility issues.
	Link to Data Repository	Confirms that there is a direct link provided to the data repository where the data is stored, ensuring easy access for others.
	Descriptions of How to Access Data	Ensures that clear instructions are provided on how to access the data, making it user-friendly for others who wish to use it.
	Abbreviations Noted/Defined	Checks that all abbreviations used in the manuscript are clearly noted and defined to avoid confusion.
	All Contributors/Authors Credited	Ensures that all individuals who contributed to the research are properly credited and acknowledged in the manuscript.

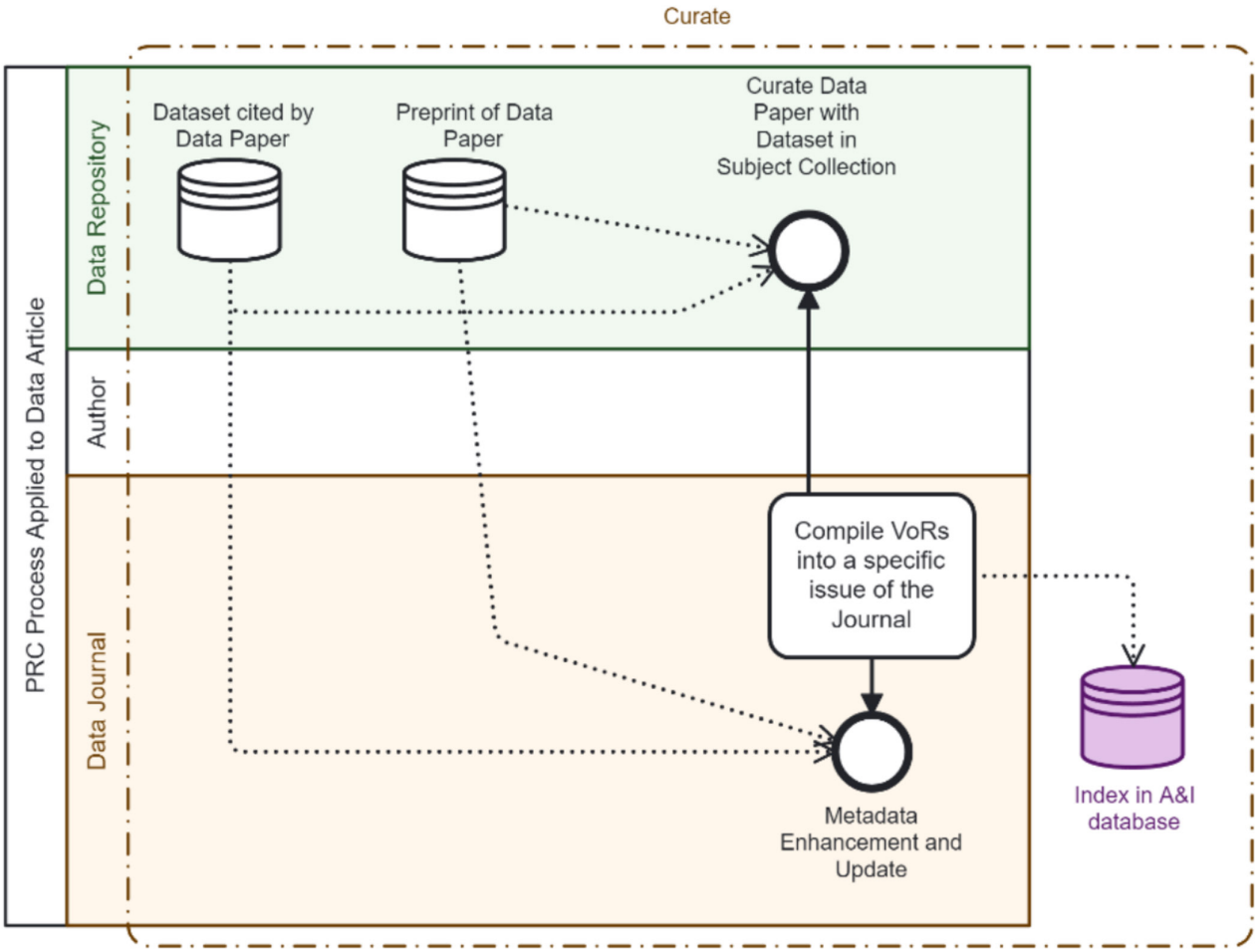

FIGURE 6 | Curation process in PRC model.

their initial publication, supporting ongoing research and innovation.

3.4 | Roles and Functions of Contributors in the PRC Model

The PRC model for data papers involves multiple contributors, each playing distinct roles to ensure the effective publication, review, and curation of the dataset and its corresponding data paper. The contributors include authors, data experts, journal editors, peer reviewers, and the broader research community. Below is a detailed outline of their roles and functions:

- **Authors/Researchers**
  - Data Submission: Authors are responsible for creating, managing, and submitting both the dataset and the data paper to the repository. They must ensure that the dataset adheres to FAIR principles.
  - Preprint Submission: After initial repository checks, authors may submit the manuscript as a preprint of a data paper, enabling early public access and feedback.
  - Manuscript Preparation and submission to a journal: Authors describe the dataset, including its context, methodology, and potential applications, in a manuscript according to the author's guidelines of a data journal. This manuscript is submitted with the link to the dataset and its preprint.
  - Responding to Reviews: Authors address feedback and make necessary revisions based on the reviews from technical reviewers, scientific reviewers, and editors.
- **Data experts**
  - Initial Quality Control: Upon receiving the dataset and manuscript, repository staff perform a preliminary review of the dataset's quality, including its metadata, consistency, and adherence to repository guidelines.
  - Data Management: They ensure the dataset is adequately archived, assign a persistent identifier (e.g., DOI), and manage version control. The repository also facilitates public access to the preprints and datasets, ensuring they are available for reuse.
  - Facilitation of Public Feedback: Repository staff monitor and manage public feedback on the dataset and preprint, aggregating metrics such as usage, citations, and altmetrics for further assessment.
  - Dataset Evaluation: Technical reviewers, typically data experts within the repository, review the dataset for metadata quality, logical consistency, adherence to standards, and overall integrity. Their focus is on the technical aspects of the dataset, ensuring it is accurate, complete, and properly described.

- Preparation of Technical Reports: After reviewing the dataset and its associated metadata, technical reviewers prepare a report detailing their findings. Public feedback and aggregated metrics on the dataset and preprint may be included in the technical report. This report is submitted to the journal editors to aid in the publication decision.
- Final Curation: After successful peer review, the curation team within the repository or journal organises the final publication of the data paper and dataset. They ensure that the dataset is properly indexed and accessible and that it is linked to related datasets or collections for broader discovery and reuse.

- **Journal Editors**
  - Editorial Review: Journal editors assess the suitability of the data paper for publication, considering factors such as relevance, originality, and methodological soundness. They decide whether the manuscript should proceed to the peer review stage, be rejected, or be published as a final version.
  - Coordination of Peer Review: Editors coordinate with technical and scientific reviewers to ensure a comprehensive evaluation of the dataset and the data paper. They also mediate between authors and reviewers, facilitating revisions and final publication decisions.

- **Experts in the Field**
  - Scientific Review: Experts in the field evaluate the data paper's scientific content, including the methodology used in data collection, the relevance of the data to the research field, and the overall quality of the manuscript. They assess whether the data paper meets the scholarly standards of the field.
  - Feedback on Research Impact: These experts also provide insights into the potential impact of the dataset and the data paper on future research, helping to determine its significance and value.

- **Research Community**
  - Public Feedback and Usage Metrics: The broader research community can provide feedback once the dataset and data paper preprint are publicly available. This feedback and usage metrics (downloads, citations) contribute to the overall evaluation of the data paper's impact and usability.
  - Curation and Reuse: Community members may reuse the dataset for further research, thereby contributing to its citation and expanding its academic value.

The PRC model thus clearly delineates the roles of various contributors, ensuring that each step—from publication to review and curation—enhances the quality, transparency, and accessibility of data papers and their associated datasets.

4 | Discussion

We discuss the presented workflow's strengths and weaknesses by comparing it with other models suggested in the related studies.

4.1 | Comparative Analysis With Existing Models

Callaghan et al. (2013) helped establish foundational procedures for dataset publication within specific domains (e.g., geosciences), prioritising archival quality, metadata rigour, and traceability. However, it lacked integration between review and curation, treated the dataset as a secondary object, and offered limited transparency. In contrast, the presented PRC model advances this legacy in the following respects: (1) treating the dataset and data paper as co-equal publishable units, (2) embedding both into a shared peer review and curation workflow, and (3) enabling preprint openness and public evaluation, in line with FAIR and open science principles. Callaghan's model illustrates a split editorial–curation structure, leading to siloed responsibilities. In contrast, our model introduces a coordinated editorial ecosystem, where data repositories and journals share synchronised tasks across the publication lifecycle. Callaghan's workflow assumes closed peer review with no early sharing. PRC explicitly introduces preprint publishing and open peer feedback, reflecting modern open science values—Callaghan's model results in a citable dataset and a narrative article linked via DOI. The PRC model results in a jointly reviewed, curated unit—the data paper and dataset as linked scholarly outputs, traceable and versioned from preprint to publication.

Austin et al. (2017) developed a comprehensive, modular reference model for data publishing that was synthesised from best practices across disciplines, focusing on interoperability between systems and stakeholders. Their approach reflects the complexity and flexibility of real-world data publishing workflows—but does not enforce a unified editorial process across data and article review. Austin et al.'s model is agnostic to workflow order and flexible across disciplines. Our model offers a structured editorial pipeline, which may be more applicable for scholarly communication where accountability, versioning, and traceability are paramount. The reference model highlights peer review as a possible component but does not specify how or by whom. The PRC model for data papers and datasets defines it explicitly, assigning review responsibilities by domain (scientific) and function (technical), addressing long-standing gaps in dataset validation. The suggested model also introduces open peer review and early community engagement via preprints, aligning with trends in open science. While both models recognise curation, our model elevates it as a final collaborative phase involving curators, repositories, and journal editors to maintain the quality and accessibility of the final record.

While Bloom et al. (2015) offer a crucial early-stage blueprint for recognising the interdependence of journals and repositories, their framework remains disjointed and minimally collaborative. The suggested model improves this by proposing a seamless, collaborative pipeline that treats the dataset and data paper as co-equal scholarly outputs. The PRC model strengthens the transparency, accountability, and reusability of data-centric research by embedding repository staff and curators directly into the peer review and curation process.

Our model contributes a structured, transparent framework for integrating peer review and curation into data paper publishing. Unlike traditional workflows that often emphasise only the narrative components of data papers, it incorporates a parallel review of both the manuscript and its underlying dataset—a practice currently not widely implemented by most data journals (Candela et al. 2015; Costello et al. 2013). This dual-review approach bridges a critical gap by ensuring scientific and technical scrutiny and enhancing data quality and trustworthiness. Our conceptual and operational workflow reflects current innovations in scholarly communication, particularly those informed by preprint culture (Eisen et al. 2020b; Corker et al. 2024). By allowing pre-submission public feedback and structured technical reviews from repository experts, the PRC model facilitates rapid scholarly exchange and distributes the review burden more efficiently among specialised evaluators. These features align with emerging practices in open peer review and collaborative data publishing.

We acknowledge concerns about the risks of accelerated peer review, especially the potential for disseminating unverified or erroneous information via preprints (Berg et al. 2024; Sokolova 2016). While this risk is more pronounced in clinical or policy-sensitive domains, the primary goal of data papers—to describe datasets rather than interpret findings—somewhat mitigates these dangers. Nevertheless, we emphasise that rigorous technical and scientific reviews remain central to our model. To address reviewer engagement, journals and repositories may explore hybrid incentive mechanisms, such as micro-credits, citable review reports, or formal reviewer acknowledgments, as suggested in prior studies (Moher et al. 2018; Tennant 2019).

4.2 | Limitations of the Suggested Model

While the suggested model in this work offers a comprehensive and transparent framework for data paper publishing, several technical, organisational, and cultural limitations must be acknowledged. Despite its conceptual rigour, our model requires substantial coordination between data journals and repositories—a level of integration that is currently rare. Most journals operate independently of repositories and lack the workflows, APIs, or policy frameworks to support cross-platform peer review. The model assumes the availability of domain experts and data specialists to conduct rigorous reviews within a limited timeframe. However, the current academic reward system does not consistently incentivise participation in data-specific peer review, particularly from repository professionals or technical metadata experts.

The PRC model is designed to be discipline-agnostic, but the implementation may be more feasible in data-intensive fields (e.g., genomics, geoscience, machine learning) where FAIR principles and repository ecosystems are mature. Adopting may face cultural resistance or practical barriers in less structured domains or where community practices for data sharing are still evolving. The suggested model is grounded in design science and literature synthesis but has not yet undergone systematic empirical testing in operational journal–repository environments. Pilot implementations and longitudinal studies are needed to assess its effectiveness, stakeholder satisfaction, and impact on data quality and reuse.

In summary, though this work sets a high standard for transparent, rigorous, and collaborative data publishing, practical challenges—especially around reviewer participation, technical integration, policy harmonisation, and sustainability—must be addressed to ensure scalable and equitable adoption across the scientific ecosystem.

4.3 | Other Issues to Be Discussed

Though our work focuses more on infrastructure, there are crucial technical considerations to improve interoperability. There is a need for a standardised protocol that allows for the transfer of manuscripts, technical review reports, public feedback, and metrics data between repositories and journal platforms. Existing standards like the Manuscript Exchange Common Approach (MECA) can be adapted for this purpose. In addition, established metrics standards such as COUNTER (Fenner et al. 2018) and altmetrics can be leveraged to measure the impact of datasets and data paper preprints.

On the policy front, a critical decision revolves around whether repositories should manage preprints of data papers and how these preprints should differ from the metadata or data management plans associated with the datasets. Questions remain regarding who owns the preprint, how intellectual property is shared between repositories and journals, and how to handle version control of data and manuscripts. These issues require harmonised policies on data licensing, embargo periods, and open peer review disclosures, none of which are yet standardised across platforms. The decision will significantly affect how data is curated and reviewed within repositories.

Finally, the business model for peer review of data papers requires careful consideration. Addressing the question of who bears the cost of peer review is paramount. Potential models include an author-pays system or a journal curation service model. Both options present different implications for the sustainability and scalability of the PRC model in the context of data paper peer review. While the PRC model presents a promising framework for enhancing the quality and efficiency of data paper reviews, its successful implementation hinges on addressing further technical, policy, and business challenges.

5 | Conclusion

We presented a novel editorial and technical framework for enhancing the publication and evaluation of research datasets and data papers through the Publish–Review–Curate (PRC) model. The presented PRC model aims to advance data publishing by integrating preprint workflows, open peer review, and structured data curation, addressing several limitations of current data journal practices. Its contributions span conceptual, methodological, and operational dimensions.

First, this work contributes a conceptual model for the synchronised publication and review of data papers and datasets. Unlike conventional data publishing frameworks that often prioritise the manuscript over the dataset, the PRC model emphasises their interdependence and proposes coordinated

evaluation criteria. This integrated approach supports data quality assurance, enhances transparency in scholarly communication, and aligns with emerging principles of open science.

Second, the study provides a detailed workflow outlining the roles of authors, peer reviewers, repository curators, and journal editors. The model facilitates open-access dissemination via preprint servers while introducing structured touchpoints for data validation and feedback across multiple stages of the editorial process. This includes early dissemination, iterative review, metadata checks, and post-publication curation—each mapped onto defined responsibilities.

Third, this study conducts a comparative analysis between the presented PRC model and several influential frameworks. These comparisons demonstrate that while existing models offer modular views of data publishing, they often lack integration between repositories and journals or omit the peer review of datasets entirely. The PRC model addresses these gaps by establishing a collaborative workflow that explicitly links trusted data repositories (as preprint platforms and data service platforms) and (data) journal systems.

Fourth, the paper presents challenges and limitations of this work, including discipline-specific adoption barriers, reviewer incentivisation gaps, interoperability concerns, and the absence of shared technical standards for versioning. These limitations are acknowledged and framed as opportunities for future implementation studies and community-driven refinement of the model.

In sum, this work provides a forward-looking blueprint for modernising data publishing infrastructure, potentially enhancing reproducibility, data reuse, and academic recognition of data authorship. Future research should include pilot deployments in domain-specific journals and repositories to empirically evaluate the PRC model's effectiveness and scalability within real-world scholarly communication ecosystems.

Author Contributions

Youngim Jung and Sungsoo Robert Ahn contributed to the conceptualisation, methodology, investigation, and validation of the study. The initial manuscript draft was prepared by Youngim Jung and Sungsoo Robert Ahn. All authors provided feedback on earlier versions of the manuscript. All authors reviewed and approved the final manuscript.

Acknowledgements

This work was supported by the Korea Institute of Science and Technology Information (K25L3M1C2; J25JR040-25), the New Faculty Research Support Grant from Gyeongsang National University, GNU-NFRSG-0121, in 2023, and the Glocal University 30 Project Fund of Gyeongsang National University, GNU-GlocalUniv-30, in 2025.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

ACM Multimedia. 2025. "Call for Dataset Paper Submissions." https://acmmm2025.org/datasets/.

Ahn, S., S. N. Cho, and Y. Jung. 2023. "Analysis and Modeling of Essential Concepts and Process for Peer-Reviewing Data Paper." *Journal of the Korean Library and Information Science Society* 54, no. 3: 321–346.

Asok, K., S. S. Dandpat, D. K. Gupta, and P. Shrivastava. 2024. "Common Metadata Framework for Research Data Repository: Necessity to Support Open Science." *Journal of Library Metadata* 24, no. 2: 133–145.

Asok, K., D. K. Gupta, and P. Shrivastava. 2024. "Research Data Policy: A Library and Information Science Publishers' Perspective." *Quality and Quantity* 59: 995–1016. https://doi.org/10.1007/s11135-024-01994-8.

Assante, M., L. Candela, D. Castelli, and A. Tani. 2016. "Are Scientific Data Repositories Coping With Research Data Publishing?" *Data Science Journal* 15, no. 6: 1–24. https://doi.org/10.5334/dsj-2016-006.

Austin, C. C., T. Bloom, S. Dallmeier-Tiessen, et al. 2017. "Key Components of Data Publishing: Using Current Best Practices to Develop a Reference Model for Data Publishing." *International Journal on Digital Libraries* 18: 77–92.

Bender, E. M., and B. Friedman. 2018. "Data Statements for NLP: Toward Mitigating System Bias and Enabling Better Science." *Transactions of the Association for Computational Linguistics* 6: 587–604. https://doi.org/10.1162/tacl_a_00041.

Berg, R. M. G., K. L. Hamilton, J. F. Murray, and P. Fong. 2024. "Peer Review: The Imprimatur of Scientific Publication." *Journal of Physiology* 602: 4079–4083. https://doi.org/10.1113/jp286273.

Bloom, T., S. Dallmeier-Tiessen, F. Murphy, et al. 2015. "Workflows for Research Data Publishing: Models and Key Components (Submitted Version)." https://doi.org/10.5281/zenodo.20308.

Borgman, C. L. 2015. "Big Data, Little Data, No Data: Scholarship in the Networked World."

Bozada, T., J. Borden, J. Workman, M. Del Cid, J. Malinowski, and T. Luechtefeld. 2021. "Sysrev: A FAIR Platform for Data Curation and Systematic Evidence Review." *Frontiers in Artificial Intelligence* 4: 685298. https://doi.org/10.3389/frai.2021.685298.

Callaghan, S., S. Donegan, S. Pepler, et al. 2012. "Making Data a First Class Scientific Output: Data Citation and Publication by NERC'S Environmental Data Centres." *International Journal of Digital Curation* 7, no. 1: 107–113.

Callaghan, S., F. Murphy, J. Tedds, et al. 2013. "Processes and Procedures for Data Publication: A Case Study in the Geosciences." *International Journal of Digital Curation* 8, no. 1: 193–203. https://doi.org/10.2218/ijdc.v8i1.253.

Callaghan, S., J. Tedds, J. Kunze, et al. 2014. "Guidelines on Recommending Data Repositories as Partners in Publishing Research Data." *International Journal of Digital Curation* 9, no. 1: 152–163. https://doi.org/10.2218/ijdc.v9i1.309.

Candela, L., D. Castelli, P. Manghi, and A. Tani. 2015. "Data Journals: A Survey." *Journal of the Association for Information Science and Technology* 66, no. 9: 1747–1762.

Carpenter, T. A. 2017. "What Constitutes Peer Review of Data: A Survey of Published Peer Review Guidelines." ArXiv Preprint ArXiv:1704.02236.

Chavan, V., and L. Penev. 2011. "The Data Paper: A Mechanism to Incentivize Data Publishing in Biodiversity Science." *BMC Bioinformatics* 12: 1–12.

Chavan, V., L. Penev, and D. Hobern. 2013. "Cultural Change in Data Publishing Is Essential." *Bioscience* 63, no. 6: 419–420. https://doi.org/10.1525/bio.2013.63.6.3.

Chiarelli, A. 2019. "Preprints and Open Peer Review: The Case of the Life Sciences." Research on Research Institute (RoRI) Report. Zenodo. https://doi.org/10.5281/zenodo.3357727.

Clifton-Ross, J., A. Dale, and R. Newell. 2019. "Frameworks and Models for Disseminating Curated Research Outcomes to the Public." *SAGE Open* 9, no. 2: 1–13. https://doi.org/10.1177/2158244019840112. (Original work published 2019).

cOAlition S. 2023. "Towards Responsible Publishing: A Proposal From COAlition S." https://doi.org/10.5281/zenodo.8398480.

Conrad, T. O. F., E. Ferrer, D. Mietchen, L. Pusch, J. Stegmüller, and M. Schubotz. 2024. "Making Mathematical Research Data FAIR: Pathways to Improved Data Sharing." *Scientific Data* 11: 676. https://doi.org/10.1038/s41597-024-03480-0.

Corker, K. S., L. Waltman, and J. A. Coates. 2024. "Understanding the Publish-Review-Curate (PRC) Model of Scholarly Communication." https://doi.org/10.31222/osf.io/h7swt.

Costello, M., and J. Wieczorek. 2014. "Best Practice for Biodiversity Data Management and Publication." *Biological Conservation* 173: 68–73. https://doi.org/10.1016/j.biocon.2013.10.018.

Costello, M. J., W. K. Michener, M. Gahegan, Z.-Q. Zhang, and P. E. Bourne. 2013. "Biodiversity Data Should Be Published, Cited, and Peer Reviewed." *Trends in Ecology & Evolution* 28, no. 8: 454–461.

Edmunds, S. C. 2023. "Peer Review Week 2023: GigaByte Joins the 'Publish, Review, Curate' Revolution Using Society." GigaScience. http://gigasciencejournal.com/blog/peer-review-week-2023/.

Eisen, M. B., A. Akhmanova, T. E. Behrens, D. M. Harper, D. Weigel, and M. Zaidi. 2020a. "Implementing A Publish, Then Review Model of Publishing." *eLife* 9: e64910.

Eisen, M. B., A. Akhmanova, T. E. Behrens, D. M. Harper, D. Weigel, and M. Zaidi. 2020b. "Peer Review: Implementing a Publish, Then Review Model of Publishing." *eLife* 9: e64910. https://doi.org/10.7554/eLife.64910.

eLife. 2022. "Elife's New Model: Changing the Way You Share Your Research." *eLife*. https://elifesciences.org/inside-elife/54d63486/elife-s-new-model-changing-the-way-you-share-your-research.

Elsevier. 2024. "Research Elements Journals." https://www.elsevier.com/researcher/author/tools-and-resources/research-elements-journals.

ESA. 2024. "Ecology Author Guidelines." https://esajournals.onlinelibrary.wiley.com/hub/journal/19399170/author-guidelines#data-paper-submission-type.

Fenner, M., D. Lowenberg, M. Jones, et al. 2018. "Code of Practice for Research Data Usage Metrics Release 1 (No. e26505v1)." *PeerJ Preprints*.

Hansson, K., and A. Dahlgren. 2022. "Open Research Data Repositories: Practices, Norms, and Metadata for Sharing Images." *Journal of the Association for Information Science and Technology* 73, no. 2: 303–316.

Jackson, B. 2021. "Open Data Policies Among Library and Information Science Journals." *Publica* 9, no. 2: 25. https://doi.org/10.3390/publications9020025.

Jeong, Y. I., J. Y. Ro, S. N. Cho, and S. Ahn. 2022. "A Study on Science and Technology Scholarly Societies' Understanding on Open Peer Review." *Journal of the Korea Contents Association* 22, no. 10: 59–73.

Jung, Y., O. Kwon, K. Kim, S. Kim, T. Seo, and S. Kim. 2020. "A Study on the Strategies for Publishing Data Journals in the Field of Ecology: Focused on K Institution." *Journal of Korean Library and Information Science Society* 51, no. 4: 83–100.

Kraker, P. 2011. "The Publish-Review-Curate Model: A Proposal for a More Open Scholarly Publishing System." ArXiv Preprint. ArXiv:1105.0532.

Lawrence, B., C. Jones, B. Matthews, S. Pepler, and S. Callaghan. 2011. "Citation and Peer Review of Data: Moving Towards Formal Data Publication." *International Journal of Digital Curation* 6, no. 2: 4–37.

Marsolek, W., S. J. Wright, H. Luong, S. M. Braxton, J. Carlson, and S. Lafferty-Hess. 2023. "Understanding the Value of Curation: A Survey of Researcher Perspectives of Data Curation Services From Six US Institutions." *PLoS One* 18, no. 11: e0293534. https://doi.org/10.1371/journal.pone.0293534.

Mayernik, M. S., S. Callaghan, R. Leigh, J. Tedds, and S. Worley. 2015. "Peer Review of Datasets: When, Why, and How." *Bulletin of the American Meteorological Society* 96, no. 2: 191–201.

Moher, D., F. Naudet, I. A. Cristea, F. Miedema, J. P. A. Ioannidis, and S. N. Goodman. 2018. "Assessing Scientists for Hiring, Promotion, and Tenure." *PLoS Biology* 16, no. 3: e2004089. https://doi.org/10.1371/journal.pbio.2004089.

Poline, J. B. 2019. "From Data Sharing to Data Publishing." *MNI Open Research* 2: 1. https://doi.org/10.12688/mniopenres.12772.2.

Ross-Hellauer, T. 2017. "What Is Open Peer Review? A Systematic Review." *F1000Research* 6: 588. https://doi.org/10.12688/f1000research.11369.2.

Scientific Data. 2024. "Aims and Scope." https://www.nature.com/sdata/aims-and-scope.

Simon, H. A. 2019. *The Sciences of the Artificial, Reissue of the Third Edition With a New Introduction by JohnLaird*. MIT press.

Smith, V. S. 2009. "Data Publication: Towards a Database of Everything." *BMC Research Notes* 2: 1–3.

Sokolova, I. M. 2016. "Toss the Water, Keep the Child." *EMBO Reports* 17: 931. https://doi.org/10.15252/embr.201642699.

Tennant, J. P. 2019. "Navigating the Evolving Landscape of Open Science." *F1000Research* 8: 2143. https://doi.org/10.12688/f1000research.21066.1.

Wang, X., and L. Xu. 2025. "What Are Journals and Reviewers Concerned About in Data Papers? Evidence From Journal Guidelines and Review Reports." *Learned Publishing* 38, no. 2. https://doi.org/10.1002/leap.2001.