

©2023,



This work is licensed under the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/legalcode>), which permits unrestricted use, distribution, and reproduction, provided the original work is properly credited. Cette œuvre est mise à disposition selon les termes de la licence Creative Commons Attribution (<https://creativecommons.org/licenses/by/4.0/legalcode>), qui permet l'utilisation, la distribution et la reproduction sans restriction, pourvu que le mérite de la création originale soit adéquatement reconnu.

IDRC GRANT / SUBVENTION DU CRDI : - GENDERING WATER AND CLIMATE SCIENCE RESEARCH IN SOUTH ASIA

Stage 2-Data Management Plan (DMP)

1. Project Title

The DMP title must be the same as your project title. The title needs to be clear, concise, and in plain language.

GENDERING WATER AND CLIMATE SCIENCE RESEARCH IN SOUTH ASIA

2. Project Description

The description should outline the aims, objectives, and methodology. This should read like your project abstract but highlight your data collection, management, and publishing plans.

Guidance resources:

UK Data Service, *Data management planning for ESRC researchers*. Available online:

<https://ukdataservice.ac.uk/manage-data/plan/dmp-esrc.aspx>

This resource provides a brief but thorough overview of the kinds of information that should be included in a data description.

UK Data Service, *Study-level documentation*. Available online:

<https://www.ukdataservice.ac.uk/manage-data/document/study-level.aspx>

The overarching objective of this project is to improve integration of gender analysis and larger interdisciplinary research and practice in WCS in South Asia.

The specific objectives are to:

1. Understand barriers and catalysts for effective practice of interdisciplinarity and gender analysis in WCS research with particular focus on whether gender inclusivity among WCS scholars improves this practice
2. Improve frameworks of methods, approaches, and pedagogical tools to effectively translate social and gender theory and analyses across disciplinary boundaries between social and natural sciences, and embody these frameworks in a common teaching curriculum and capacity building of 40 fellows towards their postgraduate research.
3. Strengthen a network and platform of multiple stakeholders for knowledge exchange, advocacy, and mentorship of young water and climate scientists for fostering and developing interdisciplinary research within and beyond higher education and academia.

Based on these objectives the action research project will be implemented in 3 countries of South Asia – Sri Lanka, Bangladesh, and Nepal by a consortium of four partner organizations. Objective 1 will involve collection of qualitative narratives from a sample of 100 past fellows of capacity building programmes conducted by the consortium over 15 years. These will be carried out using semi-

structured questionnaires. Data will be collected by postdoc researchers and new fellows/research assistants selected at the 3 Universities over the online Zoom platform. Through formal consent, these interviews will be recorded and transcribed using NVivo transcription software. Data will comprise video recordings and transcription documents. These will then be coded and analysed using NVivo subscribed software. A 2-user subscription (Academic one-time) for NVivo will be purchased and a Zoom subscription will be purchased for conducting online interviews with the participants. Data coding and analysis will be the final data outputs. Publications will emerge from this phase of the project – 2 open-access scientific peer-reviewed papers, 3 interim country reports, open-access blogs on the project website, and a project brochure available for download on the project website.

Phase 2 will analyse the same qualitative narratives for preparing a curriculum and training module. This will be in the form of a report. This report/training module will be opened for public access on the website at the end of the project period once the module has gone through multiple modifications based on the feedback from the assessment of the project fellows' research outputs. This phase will also lead to 40 postgraduate (master's and doctoral) theses across the 3 Universities. These theses are owned by the Universities and stored in institutional repositories. A digital copy of these theses shall be maintained by the Lead PI as project outputs. Fellows' research outputs and data shall be presented in the project through working papers, poster presentations at a virtual graduate seminar, and international conference presentations and conference proceedings.

Phase 3 of the project will not comprise any research data collection. However, research outputs will be disseminated through presentations, the circulation of research reports, case studies, and policy briefs. Workshop reports would be key data outputs from these sessions. Further data points of feedback from multiple stakeholders would be collected in the form of rapporteur reports.

3. Describe the datasets which will be created in the project.

Guidance:

Using this table, detail what data the project will create. Use different rows if you plan to collect multiple datasets. State what file formats you will use and your rationale for selecting the format. Bear in mind that the format may influence the ability of others to utilize and preserve your data. It can be useful to capture your data in (or convert it to) community-accepted data formats. Using standard or widely adopted formats will make your data interoperable. Open or non-proprietary formats are preferable, as you and others will have less trouble processing these later. If your data are to be deposited into a repository, particular formats may be required, so familiarize yourself with the guidance provided by the intended repository.

Guidance resources:

UK Data Service, *Format your data*. Available online:

<https://www.ukdataservice.ac.uk/manage-data/format.aspx>. This resource describes how to format data to maximize its sharing potential and long-term usability.

Description	Format	Volume	Will it be shared/open? Y/N
Past fellows and non-fellow's professional profiles (CVs)	Docx/ Pdf	200+	Y
Video recordings of qualitative interviews	Mp4	100	N
Interview transcriptions	docx	100	Y
Interviewer notes and summaries	docx	100	Y
Qualitative Data coding	xml	3	Y
Teaching Curriculum	docx / MP4 links / pdf	3	Y

3a. In the table above, the last column indicates the data that you intend to share and/or publish openly.

Data that are instrumental for your publications and that you consider to be of long-term value are to be marked "Yes". Ensure that you keep these data. Outline steps you will take to manage, store, and publish your data. These steps may have budget implications which should be accounted for in question 9.

If your answer is "No", provide a justification for not sharing or publishing the data. The IDRC *Open Data Statement of Principles* provides several reasons for when data restrictions are warranted. For example, are there ethical and legal issues that might prevent data from being made open? Have you explored procedures for mitigating potential ethical and legal concerns (also see guidance provided in the Stage 1 DMP on sensitive and security-related data)?

Guidance resources:

In addition to the resources listed below, if you plan to deposit your data in a repository, consult their guidance. Some universities and research institutions also provide data management support.

Harvey, R. (2008), *Appraisal and Selection*. DCC Briefing Papers: Introduction to Curation. Edinburgh: Digital Curation Centre. Handle: 1842/3325. Available online: <http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/appraisal-and-selection>

Natural Environment Research Council (UK), *Data Value Checklist*. Available online: <https://nerc.ukri.org/research/sites/data/policy/data-value-checklist/>

Whyte, A. & Wilson, A. (2010), *How to Appraise and Select Research Data for Curation*. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available online: <http://www.dcc.ac.uk/resources/how-guides>

CVs and video recordings of the qualitative interviews of past fellows and non-fellows will be shared after taking voluntary informed consent from the participants. CVs may hold personally sensitive information like- phone numbers, addresses, etc. of individual fellows. Similarly, Video recordings of the interviews and the transcriptions may also contain sensitive information, for instance, where an individual faced gendered challenges within the domestic sphere/ within the community/ in higher education and beyond in the research market and professional development. In cases, where an individual does not wish to disclose his/her identity, anonymised transcripts will be made available with conditional access – on direct request in order to mitigate the issues of data accessibility. Video recordings will not be shared openly to avoid the risks of scams. Audio recordings of the interviews can be shared after taking voluntary consent from the participants. The interviewer notes and summaries of interviews will be made available and accessible. This data will also be catalogued and reflected in the reports and paper publications that emerge from this data. The qualitative data coding tree shall be a part of the relevant publication emerging from this analysis or shall be submitted at associated data with peer-reviewed publications outputs from the research. The qualitative data coding will also be made openly accessible. The training module emerging from the data shall be open access only towards the end of the project period/after country stakeholder dissemination workshops as this document will go through constant modifications over 2 years based on feedback and assessment of fellows' research in multiple stages and stakeholder dialogues. The training modules will be published both on the project website and on the websites of all three universities involved in this project.

All data produced from the research will be collected, catalogued, and stored on a secure cloud space (Google/Dropbox) that will be purchased and allow various levels of conditional access to the project consortium researchers, and assistants. This will especially allow for remote data management for this multi-country project. In addition, data will also be stored in an online data repository (as indexed in Re3data) and the link of the published data with DOI will be made available on the project website to direct the users to the stored location. Open-access publications will be made either in the form of open-access journals, or reports/working papers published on the online data repository like Mendeley data/ OSF Preprints/ Zenodo.

3b. Describe the value or importance of data in the project.

If a Stage 1 DMP was submitted, refer to Stage 1 question #3 and update your response here. Describe the value or importance of the data. For example: are the data unique or built on existing published data; would the data be costly to reproduce; might the data contribute to future research efforts; and are the data likely to be cited or referenced in a project publication?

The data is unique as the sample used and data produced emerges specifically from the past 15 years collaboration of this particular consortium of partners in capacity building projects, which tried to promote gender-responsive and inclusive water science and climate education. The open access data therefore will be helpful in understanding how the capacity building has helped the fellows overcome gendered challenges faced in higher education and beyond in the research market and professional development. University records of past fellows will be used as starting points to identify samples of fellows and follow up on contact details and CVs of past fellows through direct contact and open-access social media profiles.

Production of the data will require significant human resources in the form of research assistants, interviewers, research design coordinators/disciplinary experts, as well as the cost of accessing online video calling platforms such as Zoom, video, and document data storage repositories, transcription and translation software/expertise, and qualitative data analysis software.

qualitative data collected through online interviews after taking voluntary informed consent from the participants shall be published through online data repositories with a formal DOI and further cited in project publications. The project website will also be used to link catalogued and summarised data for citation in publications. This database can be used for future research efforts as well as the interviews will cover many institutional, market, and social variables which can be valuable data for other related research on science education, gendered constraints to professional development, and institutional issues in education.

3.c Where do you intend to publish your data?

Consult Re3data (or other sources) and identify up to three data repositories for archiving your data. Briefly justify your selection. Repositories will provide guidance to researchers on how to make your data available. Note any implications for managing your data. Indicate if you intend to use identifiers (e.g., DOI and ORCID) to facilitate data sharing.

If undecided, discuss your options and next steps.

Guidance resources:

Re3data: Registry of Research Data Depositories. Available online: <https://www.re3data.org/>
Re3data provides detailed information on data repositories globally, which can be browsed by subject, country or content type.

Ball, A. & Duke, M. (2015), *How to Cite Datasets and Link to Publications*. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available online: <http://www.dcc.ac.uk/resources/how-guides>.

Digital Object Identifier (DOI). Available online: <https://www.doi.org/>. The DOI system provides a framework for persistent identification of data and other intellectual content. The use of DOIs is strongly encouraged.

ORCID. Available online: <https://orcid.org/>. ORCID provides a permanent digital identifier that distinguishes researchers and supports automated linking between them and their professional activities (including datasets), ensuring that their work is recognizable and citable.

The research consortium will require some capacity building on data repositories since the consortium partner institutions, primarily comprising South Asian universities, have limited practice of formal secure data archiving on institutional repositories. However, the research team has identified three data repositories to publish the project based on the below-mentioned criteria:

- Free
- DOI
- Ability to update files
- Directory structure

1. Open Science Framework (OSF)

OSF is a good repository for scientific data. OSF is free. One gets a DOI for the repository. There is a version control system. It supports directory structure in repositories. One can update files even after the publication and the history of the repository are tracked. The default file size limit is 5 GB. But it is possible to extend the limit with add-ons. The OSF interface is more advanced than in other repositories.

2. Mendeley Data

Mendeley Data is an open research data repository, where researchers can store and share their data. Datasets can be shared privately between individuals, as well as publicly with the world. The benefits include:

- the findings can be verified and reproduced- the data can be reused in new ways
- the discovery of relevant research is facilitated
- funders get more value from their funding investment."

Datasets uploaded to Mendeley Data go into a moderation process where they are reviewed. This ensures the content constitutes research data, is scientific, and does not contain a previously published research article. Researchers can upload and store their work free of cost on Mendeley Data.

3. Zenodo

It is a free data repository. There is a version control system. The DOI is provided. One can meter page views and downloads. The file size limit is 50GB per dataset but one can have an unlimited number of datasets.

However, one cannot create folders with files. One can upload each folder as a separate dataset or compress each folder into an archive and upload it.

At the initial phase, Google Drive, Dropbox, physical servers and institutional LAN, and website domains are currently used for data archiving which will be used to archive and share the data emerging from this project including the secure cloud space to be purchased and dedicated to this project.

For data that must remain private, primarily some specific audio and transcripts, as well as some notes, we will look to utilise the consortium cloud domain with restricted access. Undertakings for data sharing and privacy policies will be signed by all researchers involved in the project. Handing over procedures of data will also be managed strictly.

4. What metadata and supporting documentation will accompany your research data (e.g., code book, methods for data analysis)? Identify the metadata standards or requirements in use by the data repositories you identified above.

Documentation and metadata allow your data to be understood and discovered by others. Metadata describes who, how, and when the data were created, as well as identifiers that facilitate data discovery and understanding. There are various metadata standards which can help you describe your data in a consistent way, and these standards differ by discipline and by repository. Librarians, data repositories, or your colleagues may be able to advise on relevant standards.

Guidance resources:

These resources provide a general introduction to metadata standards. Data repositories are likely to recommend a metadata standard for you. We advise that you follow their guidance.

FAIRsharing.org. *The Basics of Standards*. Available online:

<https://fairsharing.org/educational/>

Higgins, Sarah (2007), *What are Metadata Standards?* Available

online:<http://www.dcc.ac.uk/resources/briefing-papers/standards-watch-papers/what-are-metadata-standards>

Research Data Alliance maintains an open directory of metadata standards for research data at <http://rd-alliance.github.io/metadata-directory/>

UQAM, *Préservation et organisation*. Available online.

<http://guides.bibliotheques.uqam.ca/themes/172-Gestion-des-donnees-de-recherche?tab=1044>

Metadata for the project will primarily catalogue updated CVs of past fellows, audiovisual material (interview recordings), interview transcripts and summaries. Simple XML documentation will be carried out comprising the following categories of information: Name of the respondent, Origin University, Year of Capacity Building under prior initiatives, Date of Interview, Interviewer Name and University, Audiovisual recording location (link for internal cloud storage location), CV document location (link for internal cloud storage location), Transcript document location (link for internal cloud storage location), and interview summary and notes. All other data will be made available online through publications, report annexures, and journal publication-linked data submissions with DOI markers. It is not planned to use a formal standard or related tools to document and describe the data. However, it is important to mention that most repositories collect additional standard metadata as part of the data submission process. The research team will provide the metadata when they will be depositing data (e.g. by filling in a form) and some will be automatically generated by the repository. The final format will depend on which data repository researchers finally choose for data sharing.

5. Has your project undergone an ethical review to ensure appropriate data management and permission for data sharing?

Select “Yes” if your project was subject to an ethics review. Describe the outcome and implications for your data management and sharing plans.

If you are preparing documentation for an ethics review, describe your proposed plan relating to consent, confidentiality, anonymization, and other ethical considerations, where appropriate. When submitting your documentation for ethics approval, include your informed consent form, if applicable, and draw attention to language that would permit you to share the data.

Select “No” if your institution does not require an ethics review and briefly explain why. Discuss procedures related to data management and data sharing/openness. Make explicit mention if you are working on sensitive data or data with human subjects. Consider ways to share the data or metadata while adhering to ethical standards. Familiarize yourself with principles of ethical research conduct as well as privacy laws or regulations in the countries where data will be collected, as these may differ from that of your research institute.

Guidance resources:

Australian National Data Service (2018), *Publishing and Sharing Sensitive Data*. Available online: https://www.ands.org.au/data/assets/pdf_file/0010/489187/Sensitive-Data-Guide-2018.pdf

Meyer, M.N. (2018), *Practical Tips for Ethical Data Sharing*. Advances in Methods and Practices in Psychological Science. Vol. 1 Issue 1, March 2018.

<https://journals.sagepub.com/doi/10.1177/2515245917747656>

UK Data Service, *Legal and Ethical Issues*. Available online:

<https://www.ukdataservice.ac.uk/manage-data/legal-ethical>

	Yes	<input checked="" type="checkbox"/>
	No	<input type="checkbox"/>

The research methodology which includes survey design, sample size, methods for qualitative data collection, data storage and data publication, will undergo ethics review process at the lead institute Post Graduate Institute of Agriculture, University of Peradeniya, Sri Lanka (https://agri.pdn.ac.lk/Pages/about/downloads/research_ethical_committee/Operational_Procedure.pdf). IDRC's Open Data Statement of Principles will be shared with the ethics board. This application will be prepared and applied for within the first 4 months of the project. Human subjects will be used for the project, therefore access to interview recordings and transcripts will be made open access only after taking informed consent from the participants. If the participants show concern about data attribution, then anonymised transcripts will be shared for data publication. All interviews will be conducted with formal written/recorded informed consent from respondents with complete transparency about the research objectives; ethics review process and approval, and modes of data utilisation. A consortium letter detailing the ethics review outcomes and documented process of ethical management of the project will be submitted to each institution in the consortium for formal consent.

6. Describe how copyright and intellectual property rights (IPR) issues affect your data management and sharing plans.

If you are using data from other researchers/organizations, or you anticipate your data will have commercial applications, discuss potential limitations to data sharing.

Guidance resources:

Australian National Data Service (2018), *Publishing and Sharing Sensitive Data*. Available online: https://www.ands.org.au/data/assets/pdf_file/0010/489187/Sensitive-Data-Guide-2018.pdf

Carroll, Michael. (2015), *Sharing Research Data and Intellectual Property Law: A Primer*. PLoS Biology. 13 (8). e1002235. DOI: [10.1371/journal.pbio.1002235](https://doi.org/10.1371/journal.pbio.1002235).

The data collected for this research will require collaborative data collection initiative from each project country and will be consolidated for regional level analysis at the regional research lead partner organisation/team. Data sharing and ownership within the consortium will be of import for the smooth functioning, non-conflictual publishing and research dissemination, and effective partnership in the consortium. At the inception stage of the project norms of data ownership within the consortium will be agreed upon along with a schedule of expected publications and sharing of authorship. Those who will be contributing to developing data collection instruments and collecting data will be recognized as authors of published data sets, however, the participants will not be considered authors.

7. How will the data be licensed?

If you plan to share your data, including a data licence will provide clarity on how others may use your data. Identify the licence you intend to use and provide your rationale.

Guidance resources:

Ball, A. (2014), *How to License Research Data*. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available online: <http://www.dcc.ac.uk/resources/how-guides/license-research-data>

CESSDA, *Licensing Your Data*. Available online: <https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/6.-Archive-Publish/Publishing-with-CESSDA-archives/Licensing-your-data>

Choosing a licence for your data or software:

- Creative Commons licences: <https://creativecommons.org/licenses/>
- Tool for selecting a licence: <https://ufal.github.io/public-license-selector/>

Where data objects can be shared openly we will apply the cc0 waiver for data outputs. For open access reports CC BY license will be used. For peer reviewed publications embargo and licenses of the respective journals will apply. To make it easier for others to re-use tools, data or other content that will be produced in this project 'By-Attribution, Non-Commercial' (CC BY-NC) license, i.e. a very common Creative Commons licence will be used. It means that anyone will be able to use the information in any way they like, as long as they attribute it to the creator of the data and don't use it for commercial purposes. Those data outputs that cannot be shared publicly will only be made available under restricted usage terms on specific requests and detailed usage proposals.

8. How will the data be stored and backed up during the research? How will you manage access and security?

Explain how you will manage the data and security procedures. What roles and responsibilities have been assigned within the research team or with external organizations? Will you be using secure online storage or transferring the data? If so, will you encrypt the data? If using online services, where is your data hosted and does this location raise any legal issues?

Bear in mind the need to back up your data. Fully managed file services with automated backup might be available to you through your institution. Such services could be used in combination with portable storage or cloud computing to meet particular needs.

Guidance resources:

Research Data MANTRA. Available online: <https://mantra.edina.ac.uk/>. Consult the *Storage and Security* training module.

CESSDA, *Data Management Expert Guide*. Available online: <https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide>. Consult modules 4. Store and 5. *Protect*.

A dedicated cloud space domain will be purchased and utilised for the audiovisual and documentary data that will emerge from this project. A dedicated web manager will provide project website maintenance as well as cloud and data maintenance services. Coordination costs have been considered for the overall coordination support of the consortium activities which will also include the responsibilities of cataloguing, storing, and enabling access to data collected from 3 project countries. Postdoctoral researchers/ research associates at each institution, under the supervision of co-PIs will be incharge of cataloguing country level datasets and regular uploading of data on to the common secure cloud storage. Once the project ends, the core team members of the project (i.e. PI, Co-PIs, regional research Lead and Regional research coordinator) will be holding the data for future uses. Data which will be collected after taking informed consent from the participants will be published on an open access repository.

9. If you intend to restrict access to the data (e.g., establish an embargo period before publishing your data or apply a restrictive data sharing licence), justify your rationale.

Specify and explain any delays to data publishing or access restrictions, such as embargo periods or restricted access, and ensure these are properly justified.

Only sensitive information such as audiovisual recordings of interviews and transcripts of interviews where participants will not give consent for open access publication, those will be restricted for access as they would tend to contain sensitive personal or institutional information. While no embargo will be established, access to this data (anonymised) would

only be made on personal request to the Lead PI along with a formal justification and detailed proposal for the use of this data. Only upon a careful assessment of the request anonymised data/partial data will be provided after discussion with the core team members who will be having access to the data.

10. If you are preparing or updating this DMP after your project budget has been approved, identify any significant budget changes to managing and sharing your data.

If you have not prepared a budget yet, consult advice in the Stage 1 DMP related to budgeting, and outline activities and related costs.

Data Management has been incorporated into the budget including human resource with overlapping roles and responsibilities of data management along with their core project research responsibilities.