



EARTH CUBE
TRANSFORMING GEOSCIENCES RESEARCH



"What About Model Data?" Best Practices for Preservation and Replicability Rubric Use Case Template

-Updated 03/30/2023

[Rubric, Usage Instructions and Use Case Examples are all archived under https://doi.org/10.5065/q936-q118](https://doi.org/10.5065/q936-q118)

User Guidance

Document Purpose: Template for a rubric use to complete after filling out the rubric and getting a composite score and model output recommendation.

Use Case Template - *Yellow Highlighted text to be completed by rubric user*

- Use Case Description
 - High-level overview of the use case
 -
 - Science goals and basic workflow
 -
- For all projects: what to preserve (see sections below for details)
 - simulation code
 - initialization data
 - simulation setup (e.g., parameterization selection)
 - pre-processing code
 - post-processing code
- What use-case specific additional materials should be preserved and shared? [the lists below are possibilities]
 - Data
 - Inputs to model
 - Description
 -
 - *Total data volume preserved in a repository by the PI*
 - *Total data volume preserved in a repository maintained by an outside data provider (e.g. NCEI)*
 - *Total data volume not preserved in a repository? (might be retained on PI's local working storage)*
 - Raw model output
 - Description
 -

- *Total data volume preserved in a repository by the PI*
 - *Total data volume not preserved in a repository? (might be retained on PI's local working storage)*
 - Processed model output
 - Description
 -
 - *Total data volume preserved in a repository by the PI*
 - *Total data volume not preserved in a repository? (might be retained on PI's local working storage)*
 - Software
 - Model configuration
 -
 - Model code
 -
 - Preprocessing and Postprocessing code
 -
 - Other
 - Documentation [everything should have this, but maybe there are special kinds of documentation produced for particular use cases]
 -
 - Visualizations or images [products intended to be used visually, distinguished from processed output that exists as numerical data]
 -
- Why should these things be preserved and shared?
 - General
 -
 - Reasons why the things listed above are important
 - Note expected/intended audience and what they expect/need
 - Are there specific people who will be using the data downstream?
 -
 - Possible/aspirational users?
 -
 - Note any temporal considerations, such as particular products that become more/less useful over time
 -
 - Could refer to individual rubric descriptors in this section - which descriptors are most important/useful to guide the preservation recommendations for each case?
 -
- Broader Impacts: Note that stakeholders in the following section are any set of users that will utilize the output in decision making for their community, e.g., resource management, hazard mitigation, climate adaptation, etc. In particular, collaboration with stakeholders from minority and indigenous populations should be addressed directly in the following answers.
 - How will output from this project be used by stakeholders?

-
- How were stakeholders involved in the data curation decision-making?
-
- How will stakeholders be compensated for their participation in the data curation decision-making process?
-
- Do you have any concerns about misuse of your data or software? If so, what concerns do you have, and what are the reasons for those concerns?
 -