Assumption: Rubric to be used to assist a researcher in determining what simulation outputs should be deposited in a FAIR aligned community repository to communicate knowledge.

Simulation workflow outputs are assumed to be produced by a combination of the simulation run and simulation post processing workflow components.

Rubric Usage Instructions and Use Case Examples can be accessed at: https://modeldatarcn.github.io/rubrics-worksheets/Rubric-Instructions-and-Use-Cases.pdf

	Simulation / Experiment Descriptors		Simulation / Experiment Descriptor Classes					
			Preserve Less Output		Preserve More Output			
Big Picture Question	**Descriptor**	**Descriptor definition**	**Class 1**	**Class 2**	**Class 3**	**User Entered Score. (Integers only) Score Range: 1 -Class 1 2 -Class 2 3 -Class 3**	**Suggested Weight (If score > 1)**	**Weighted Score (Weighted Score = Entered Score x Recommended Weighting when score > 1)**
Section Theme: Community Commitment								
If you score high (e.g. 13-18) in this section this may be a "Data Production" use case. "Data Production" use cases are projects with output intended for large numbers of downstream users.								
Is it anticipated that your simulation workflow outputs will have broad community impact and downstream reuse?	Used in a "Highly Influential Scientific Assessment"	As defined, for example, by OMB "Revised Information Quality Bulletin for Peer Review" (2004 Apr 15): a scientific assessment whose "dissemination could have a clear and substantial impact on important public policies (including regulatory actions) or private sector decisions with a potential effect of more than $500 million in any one year or that the dissemination involves precedent setting, novel and complex approaches, or significant interagency interest."	No, not used in any HISA.	Subset of output may enable fact checking, e.g. all output are not needed, but selected or derived products (e.g. ensemble mean and spread) will provide adequate scientific representation.	Used in a HISA. Need to keep output for future fact checking.		**2**	
	Part of larger community set - Continuum of coordinated experiments vs solo/smaller events	Is this simulation output part of a larger set, that is of value as a whole? (e.g., intercomparisons)	No, not part of a larger set	Subset of data may be more appropriate for some kinds of ensemble experiments.	Yes, output is part of a larger set of related experiments.		**2**	
	Community Benchmark Dataset	Is this simulation output potentially a community benchmark for comparison?	No, not a benchmark or community reference dataset.		Yes, output is a community reference dataset (e.g. global reanalysis).		**2**	
					Section Total Raw Score. (Min=3, Max=9)		**Section Total Weighted Score. (Min=3, Max=18)**	
Section Theme: Repository Data Accessibility								
This section is important for those intending to produce high volume, "Data Production" datasets to be accessed by many downstream users (e.g. weighted score between 13-18 from above "Community Commitment" section above).								
Does the FAIR aligned community repository that you plan on archiving your data in provide adequate data access capabilities for the volume of data that you plan on depositing?	Repository Supported Data Transfer	Do bandwidth limitations impede data transfer options from the community data repository expected to archive the simulation output?	Data is volume is too large to effectively transfer and no data volume reduction capabilities are provided by the repository.		Data volume is small enough, or data volume reduction services are provided by the repository to support data effective data transfer.		**2**	
	Repository Supported Data Analysis	Is there a capability to access/use data analysis compute resources colocated with the community data repository, where the simulation output will be archived?	No publicly accessible data analysis compute capabilities are co-located with the data repository expected to host the simulation output.		Publicly accessible data analysis compute capabilities are co-located with the data repository expected to host the simulation output.		**2**	
					Section Total Raw Score. (Min=2, Max=6)		**Section Total Weighted Score. (Min=2, Max=12)**	

Assumption: Rubric to be used to assist a researcher in determining what simulation outputs should be deposited in a FAIR aligned community repository to communicate knowledge.

Simulation workflow outputs are assumed to be produced by a combination of the simulation run and simulation post processing workflow components.

Rubric Usage Instructions and Use Case Examples can be accessed at: https://modeldatarcn.github.io/rubrics-worksheets/Rubric-Instructions-and-Use-Cases.pdf

	Simulation / Experiment Descriptors			Simulation / Experiment Descriptor Classes					
				Preserve Less Output		Preserve More Output			
Big Picture Question	**Descriptor**	**Descriptor definition**		**Class 1**	**Class 2**	**Class 3**	**User Entered Score. (Integers only) Score Range: 1 -Class 1 2 -Class 2 3 -Class 3**	**Suggested Weight (If score > 1)**	**Weighted Score (Weighted Score = Entered Score x Recommended Weighting when score > 1)**
Section Theme: Simulation Workflow Accessibility									
Would it be straightforward for others in your academic discipline to rerun your simulation model run workflow steps?	Model Source Code Availability	How accessible is this particular version of the model/code? Are there IP barriers, embargo periods for new model development?		Community validated version of a highly accessible model was used.	Model source code is shareable, but specific changes were implemented that make it unique. Code is lightly documented.	Model source code is difficult to acquire		**1**	
	Model Source Code Documentation/Ease of use	Is the source code well documented and easy to use?		Source code is well documented and easy to install and run.		There is very little code and supporting documentation. Source code is difficult to understand and manage.		**1**	
	Model Compute Platform/System Dependencies	How specialized of a platform is needed to execute the model (specific hardware, compilers, software libraries needed)?		Does not require special hardware, niche software libraries, and licensed compilers to execute. This could include a containerized version of a model.		Requires resources that are more difficult to get access to. E.g. specialized HPC, niche software libraries, and licensed compilers.		**1**	
	Simulation Input Accessibility	How much effort is it to get and manage all the inputs used by the simulation?		Simulation inputs/boundary conditions are easy to acquire & manage.		If simulation inputs/boundary conditions are difficult to acquire & manage, retaining output lowers burden for others who might want to re-run model or use outputs.		**1**	
						Section Total Raw Score. (Min=4, Max=12)		**Section Total Weighted Score. (Min=4, Max=12)**	
Section Theme: Simulation Post Processing Workflow Accessibility									
Would it be straightforward for others in your academic discipline to rerun your	Post Processing Source Code Availability	How accessible is this particular version of the post processing code? Are there IP barriers, embargo periods for new model development?		Community validated version of a highly accessible post processing workflow was used.	Post processing source code is shareable, but specific changes were implemented that make it unique. Code is lightly documented.	Post processing source code is difficult to acquire		**1**	
	Post Processing Source Code Documentation/Ease of use	Is the post processing source code well documented and easy to use?		Source code is well documented and easy to install and run.		There is very little code and supporting documentation. Source code is difficult to understand and manage.		**1**	

Assumption: Rubric to be used to assist a researcher in determining what simulation outputs should be deposited in a FAIR aligned community repository to communicate knowledge.

Simulation workflow outputs are assumed to be produced by a combination of the simulation run and simulation post processing workflow components.

Rubric Usage Instructions and Use Case Examples can be accessed at: https://modeldatarcn.github.io/rubrics-worksheets/Rubric-Instructions-and-Use-Cases.pdf

	Simulation / Experiment Descriptors		Simulation / Experiment Descriptor Classes					
			Preserve Less Output		Preserve More Output			
Big Picture Question	**Descriptor**	**Descriptor definition**	**Class 1**	**Class 2**	**Class 3**	**User Entered Score. (Integers only) Score Range: 1 -Class 1 2 -Class 2 3 -Class 3**	**Suggested Weight (If score > 1)**	**Weighted Score (Weighted Score = Entered Score x Recommended Weighting when score > 1)**
discipline to rerun your simulation post processing workflow steps?	Post Processing Compute Platform/System Dependencies	How specialized of a platform is needed to execute the post processing code (specific hardware, compilers, software libraries needed)?	Does not require special hardware, niche software libraries, and licensed compilers to execute. This could include a containerized version of a post processing workflow.		Requires resources that are more difficult to get access to. E.g. specialized HPC, niche software libraries, and licensed compilers.		**1**	
					Section Total Raw Score. (Min=3, Max=9)		**Section Total Weighted Score. (Min=3, Max=9)**	
Section Theme: Research Workflow Output Accessibility								
Would it be straightforward for others across academic disciplines to use your simulation workflow outputs?	Simulation/Post Processing Output Usability	How easy is it to use the outputs outside the original context? Does it adhere to community standards/conventions (e.g. CF NetCDF)? Are the metadata sufficient for someone else to understand the output.	Simulation outputs provided in proprietary format. Obscure or undefined standards make usability and long term curation difficult.		Simulation outputs structured, formatted, and aligned with community conventions. Data can be easily read by common software and understood in the future.		**2**	
					Section Total Raw Score. (Min=1, Max=3)		**Section Total Weighted Score. (Min=1, Max=6)**	
Section Theme: Research Feature Reproducibility								
Would it be feasible for others in your academic discipline to reproduce a physical feature generated through your simulation?	Simulation Feature Reproducibility	The ability to reproduce specific (atmospheric) features (of given scale) within an acceptable statistical range of error.	No issues with specific feature reproducibility	Would be difficult to reproduce some feature details, but general findings are robust	Would be difficult to reproduce due to nonlinearity of phenomena being studied		**3**	
					Section Total Raw Score. (Min=1, Max=3)		**Section Total Weighted Score. (Min=1, Max=9)**	
Section Theme: Cost of Running Simulation Workflow								
What is the cost to produce your simulation workflow outputs?	Computational Cost of Running the Simulation Workflow	The economic cost (combination of run time and computer access costs) of completing simulation workflow	Small computational cost and no special platform needs	Moderate computational cost, but access to needed platforms straightforward	High computational cost. Need a large compute capability and/or can only be produced with specialized platforms		2	
	Human Resource Cost of Producing the Simulation Workflow	Person-hours required to reproduce a simulation dataset	Trivial effort required to replicate simulation for most end users.		Significant time & expertise required to replicate simulation. Likely will require contact with & guidance from original data producer(s).		2	

Assumption: Rubric to be used to assist a researcher in determining what simulation outputs should be deposited in a FAIR aligned community repository to communicate knowledge.

Simulation workflow outputs are assumed to be produced by a combination of the simulation run and simulation post processing workflow components.

Rubric Usage Instructions and Use Case Examples can be accessed at: https://modeldatarcn.github.io/rubrics-worksheets/Rubric-Instructions-and-Use-Cases.pdf

	Simulation / Experiment Descriptors		Simulation / Experiment Descriptor Classes					
			Preserve Less Output		Preserve More Output			
Big Picture Question	**Descriptor**	**Descriptor definition**	**Class 1**	**Class 2**	**Class 3**	**User Entered Score. (Integers only) Score Range: 1 -Class 1 2 -Class 2 3 -Class 3**	**Suggested Weight (If score > 1)**	**Weighted Score (Weighted Score = Entered Score x Recommended Weighting when score > 1)**
					Section Total Raw Score. (Min=2, Max=6)		**Section Total Weighted Score. (Min=2, Max=12)**	
Section Theme: Repository Data Management Services Cost								
What is the cost for you to archive the output in a FAIR aligned community repository to preserve and provide access to your simulation workflow outputs for a minimum period of time?	Repository Supported Data Curation Cost	The economic cost of curating simulation output in a community repository, for a minimum time period.	Community repository data curation expenses are prohibitive due to large volume of the expected model outputs.		Would be inexpensive to curate the complete simulation workflow output for a minimum number of years in a community repository.		4	
					Section Total Raw Score. (Min=1, Max=3)		**Section Total Weighted Score. (Min=1, Max=12)**	
					Rubric Total Raw Score. (Min=17, Max=51)		**Rubric Total Weighted Score. (Min=17, Max=90)**	

Rubric Use Case Examples can be accessed at: https://modeldatarcn.github.io/rubrics-worksheets/Rubric-Instructions-and-Use-Cases.pdf

Rubric Total Weighted Score < 48	**48 <= Rubric Total Weighted Score <= 72**	**72 < Rubric Total Weighted Score**
Preserve few simulation workflow outputs	Preserve selected simulation workflow outputs	Preserve the majority of simulation workflow outputs
Preserve and provide access to simulation workflow configuration and code components	Preserve and provide access to simulation workflow configuration and code components	Preserve and provide access to simulation workflow configuration and code components
See Use Case 1, Knowledge Production	*See Use Case 2, Knoweldge Production*	*See Use Case 3, Data Production*