EarthCube Research Coordination Network Grant - What About Model Data?

Draft Virtual Workshop Agenda: Workshop #1, Boulder, CO, May 5-7, 2020, **Times in MDT.**

Primary Goal: Develop initial set of model descriptors and descriptor rankings.
Needs: model descriptors, ranking categories
Breaks: Breaks will be provided as the need arises during each session.

Day 1: Tuesday, 5 May, 9:00 - 12:00 MDT
Join Hangouts Meet

https://meet.google.com/iae-tqqq-mbm

Join by phone

+1 651-571-1493 PIN: 516 517 377#

8:30-9:00: Virtual coffee to test remote connection effectiveness and work out tech bugs.
9:00-9:15: Welcome, Workshop Goals and Expectations, Agenda Review, Logistics, How to ask questions, Tech Information.
9:15-11:00: Current Requirements and Strategies in Model Storage and Reproducibility

- The goal of this session is to describe the state-of-the-science in current requirements and practices.

Presenter Name	Organization/role	Title
Sheri Mickelson ~9:15	Software Engineer, NCAR/CISL	CESM's New Data Workflow for CMIP6
Bryan Lawrence ~9:30	UK NCAS: Director of the Models and Data Division	When should a simulation be FAIR?
James Stagge ~10:00	Ohio State University: Assistant Prof, Civil Eng/Hydrologist.	Minor adjustments, major benefits: Assessing reproducibility in hydrology research
Brian Gross ~10:30	NOAA/NCEP	Moving to Community Modeling In the National Weather Service
Pat Hogan ~10:45	Oceanographic Sciences Branch (OSB) Chief, NCEI	Best Practices for Model/Data Preservation at NOAA/NCEI
Mike Friedman ~11:00	AMS publications: Senior Manager of	Model Data and Publishers: The View

	Publishing Operations	from AMS

- Discuss importance of defining terms: e.g., reproducibility
- Acknowledging adjacent issues: e.g., sharing, where to store, who pays for it

11:00-12:00 Open Discussion

Day 2: Wednesday, 6 May, 9:00 - 12:00 MDT

Join Hangouts Meet

https://meet.google.com/dei-pqsw-opu

Join by phone

+1 786-540-5428 PIN: 545 787 707#

9:00- 9:30 Review themes from day one and provide guidance for breakout groups

9:30-11:15 Breakout #1: Model Descriptor Brainstorming

- Breakout Sessions have 8-10 people, 5 groups
    - Separate Google invite for each breakout session.
- Develop model descriptors and definitions.

11:15-12:00 Reports from breakouts and open discussion

12:00-12:45 Leaders only: Work on combining Day 2 lists into one document

Day 3: Thursday, 7 May, 9:00 - 12:00 MDT

Join Hangouts Meet

https://meet.google.com/aqz-qbqf-hch

Join by phone

+1 904-580-9631 PIN: 918 972 791#

9:00-9:30: Discussion: Review/discuss combined model descriptor list and provide guidance for breakout groups

9:30-11:15: Breakout #2: Simulation Descriptors and Classes

- Descriptor list: combine, refine, add
    - “Themes” are added for organizational purposes, will not be part of public rubric
- Classes: start filling in class definitions; focus on edge cases (Classes 1 and 3)
    - Classes are only meant to describe the range of a given descriptor
    - Most simulations will be a mix of different classes (i.e., Class 1 for some descriptors, Class 3 for others, etc)

How will this be used?

- Use case testing (future workshops) will hopefully reveal simulations that group into similar descriptor/class combinations
- Best practices for how much data to save (data tier), for how long, and where to save it (private/shared) will be developed from these groupings

11:15-12:00: Reports from breakouts, combine descriptor rankings, and workshop wrap-up

Breakout session organization:

- Separate Google Hangout Breakout for each breakout session -5 breakouts. All participants will receive calendar invites to the Google Hangouts Meet breakout groups that they are assigned to.
- Each Breakout will be coordinated by a leader. Leaders guide discussion, leaders and participants to fill out Google docs and sheets (Google Drive for each workshop). Google media will be accessible for all participants to edit (anyone with the link can edit).