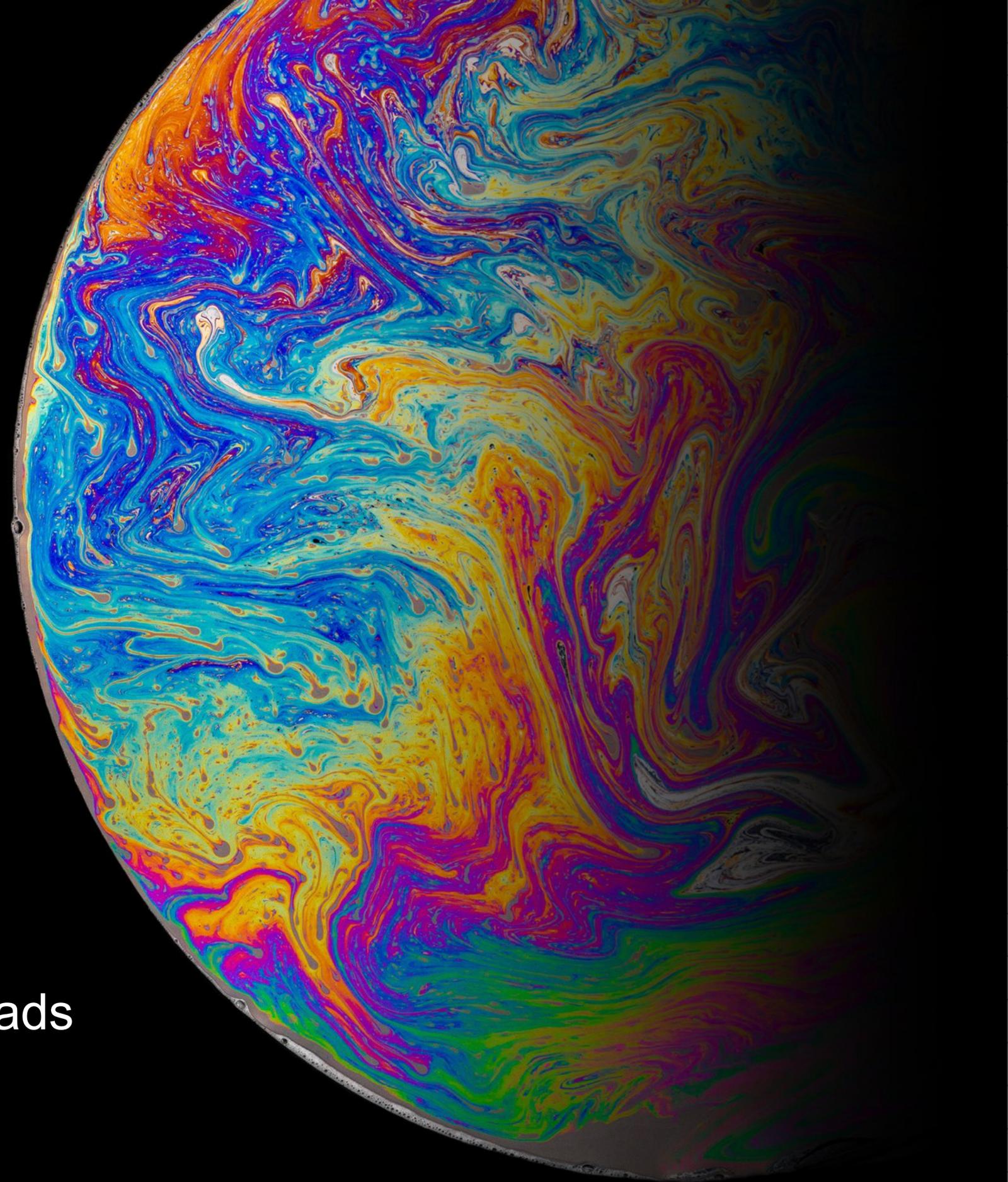


Updates from the Model Benchmarking Task Team

Birgit Hassler and Forrest M. Hoffman, Co-Leads



What are we doing on what time scale? (1)

- Meeting monthly with Task Team members
- Developing two papers for the community
 - Retrospective on CMIP6 model evaluation — What worked and what did not? What was novel? What tools were used? → ideally this should be submission-ready by spring 2024
 - Big ideas — Charting the path for new advances for CMIP7 → probably submission-ready by the end of 2024
- Planning for a face-to-face meeting in spring (one week, main focus on finishing draft of retrospective paper, but maybe also to work on a new framework?)

What are we doing on what time scale? (2)

- Airtable/Cards and videos for benchmarking packages, served on the CMIP website → work in progress right now, hopefully finished by spring 2024
- Presenting to the science community at WCRP OSC
- Convening benchmarking session at AGU, EGU, AOGS (some also have ECRs in the convener team)

Resources, Risks/Issues, Challenges?

- Resources include a diverse TT membership, many with experience developing/using benchmarking tools
- Desire to build out a comprehensive evaluation framework that can run in an automated fashion as output data are published (*e.g.*, at ESGF)
 - The community wants quick-look assessments / diagnostics
 - Unsure if we have the human resources to build them
 - Metadata (and data!) QA/QC remains a challenge / barrier

Engagement Opportunities for Fresh Eyes

- We met with representatives for Fresh Eyes Subgroups
 - Data Analysis
 - Model Evaluation
- They are knowledgeable, enthusiastic and optimistic!
- We hope to engage them in
 - assessing tools / packages,
 - developing/testing evaluation metrics, and
 - helping to convene and giving presentations at conferences & workshops

Needs from CMIP Panel and/or WIP?

- Consider encouraging better QA/QC for data / metadata
 - Improved tools for automated file checking
 - Stricter enforcement vs. ease of publication
 - Encouraging modeling centers to follow up on errata
 - Requires unified effort from TTs on Data Access, Data Citation, and Data Request
- Continued support for meetings & conference activities
- Advocacy for a framework to unify products from existing packages, possibly with funding agencies