FORCE2017-14
FORCE2017

Data citation workflow changes in climate science: CMIP6 and IPCC AR6

Martina Stockhause
German Climate Computing Center (DKRZ), orcid.org/0000-0001-6636-4972

Climate science data is of large, broad and long-term interest. It is re-used in very different scientific and political contexts. Transparency of the findings by the documentation of data and software sources is of high importance. An example workflow for climate sciences (Stockhause and Lautenschlager, 2017, doi:10.5334/dsj-2017-030) will be presented focusing on the research project CMIP6 (Coupled Intercomparison Project Phase 6) and its connections to surrounding projects and the IPCC AR6 (Intergovernmental Panel on Climate Change 6th Assessment Report). The workflow is characterized by high-volume data, an early usage of the evolving project data, and an increasing complexity in terms of the number of included research experiments, the number of project participants and the number and diversity of data user communities.

Service requests from the project participants for data citations of evolving CMIP6 data collections and tracking of data usage in literature have been targeted by the CMIP6 Citation Service. However, CMIP5 has shown that the readiness to cite data is lower among climate scientists than the desire to be cited and get credit for data. The WDCC as publisher of data citations for input4MIPs, CMIP6 and IPCC DDC has the chance to overcome this discrepancy by raising awareness among CMIP6 participants that they are also data consumers of input4MIPs data providers. Challenges in the workflow lie - apart from its complexity - in the citation of evolving data, the citation of Earth System Models consisting of several independently developed components, and in the integration of (grey to peer-reviewed) literature publication workflows.