EGU2020-13189
https://doi.org/10.5194/egusphere-egu2020-13189
EGU General Assembly 2020


# The initial Data Management Plan for PalMod II - FAIR simulation and paleo data from the Last Interglacial to the Anthropocene

**Oliver Bothe**[1] and Karsten Peters[2]
[1]Helmholtz-Zentrum Geesthacht, Institute of Coastal Research, Geesthacht, Germany
[2]Deutsches Klimarechenzentrum GmbH (DKRZ), Abteilung Datenmanagement, Hamburg, Germany

The project PalMod II is the second phase of Germany's national paleoclimate modelling initiative (www.palmod.de) whose aim is to model the transient climate evolution from the last interglacial to the anthropocene with state of the art earth system models. The second phase more precisely wants to perform simulations for the last glacial inception, the marine isotope stage 3, and the last deglaciation. It further plans to compile paleo-observational proxy data over the full glacial cycle from about 130,000 years before present until today. Models of differing complexity (fully-coupled earth system models and models of intermediate complexity) will be used to assess the scientific questions posed in PalMod II. Model output will be combined with the compiled paleo-proxy data for validation purposes. The sheer data amount in excess of several petabytes and different data handling practices of the participating communities require dedicated management of the data workflow both in- and outside of the immediate PalMod community.

The PalMod II data management takes place in close collaboration between data management specialists and the scientists. The objectives include the standardisation of each simulation and proxy dataset, the facilitation of data sharing and data reuse between work packages, the access channels for external collaborators, and the long-term preservation of the data. The data management follows the concept of the "Active Data Management Plan", which foresees a continuous development of the data management plan (DMP), starting with an initial basic version. The DMP covers the entire life cycle of the research data generated in the project, from generation and analysis to data publication and archiving. This includes aspects such as data formats, metadata standards and data usage licenses. Ownership and responsibilities for simulation and paleo data sets as well as the input data during and after the end of the project will also be considered.

This contribution will present the initial DMP for PalMod II. It will describe the amount of data produced in the project, highlight how the above mentioned aspects will be dealt with, and present how the project aims to ensure the Findability, Accessibility, Interoperability, and Reusability, i.e. the FAIR data principles, of simulation output, post-processed model data, and paleo-proxy data from PalMod II.

* This contribution presents results of the full PalMod II initiative, the authors present them on behalf of the initiative.