



Data Management for PalMod-II – A FAIR-Based Strategy for Data Handling in Large Climate Modeling Projects

SWATI GEHLOT

KARSTEN PETERS-VON GEHLEN

ANDREA LAMMERT

HANNES THIEMANN

*Author affiliations can be found in the back matter of this article

COLLECTION:
DATA MANAGEMENT
PLANNING ACROSS
DISCIPLINES AND
INFRASTRUCTURES

PRACTICE PAPER

]u[ubiquity press

ABSTRACT

PalMod-II was a multi-institutional research project in Germany focusing on enabling and performing global numerical climate simulations with state-of-the-art coupled Earth System Models spanning a full glacial cycle from 130 000 years in the past to the present and beyond. The main project goal was the dataset resulting from these simulations and making it available for reuse by the climate science community in-line with the FAIR data principles. In this paper, we present the research data management (RDM) approach developed and employed in PalMod-II to progress towards that project goal. The RDM approach was implemented by RDM professionals specifically funded by PalMod-II, which made it possible to provide RDM services tailored specifically to the project needs. The compilation and maintenance of a project-wide data management plan (DMP) has proven essential for keeping the project on track and serving as a central focal point of any data-related aspects. These include the specification of data responsible scientists, allocation of storage and computational resources on a high-performance computing system, documentation of simulation output requirements, definition of data standardisation, and publication workflows in-line with the FAIR data principles. Since the RDM approach executed in PalMod-II was first-of-its-kind for all project partners, exhaustive communication at par with the scientists was required to create trust and a collaborative atmosphere within the project. Finally, the RDM approach implemented in PalMod-II facilitated the publication of a flagship dataset for global reuse, and will also be implemented in the follow-up project: PalMod-III.

CORRESPONDING AUTHOR:

Swati Gehlot

Deutsches Klimarechenzentrum
GmbH, DE

gehlot@dkrz.de

KEYWORDS:

research data management;
data management plan; FAIR;
Climate model data; data
standardization; ESGF

TO CITE THIS ARTICLE:

Gehlot, S, Peters-von Gehlen, K, Lammert, A and Thiemann, H. 2023. Data Management for PalMod-II – A FAIR-Based Strategy for Data Handling in Large Climate Modeling Projects. *Data Science Journal*, 22: 34, pp. 1–8. DOI: <https://doi.org/10.5334/dsj-2023-034>

1. INTRODUCTION

Organisation of data-related activities in a consortial project with many partners requires dedicated, project-based research data management (RDM). This especially applies to data-intensive projects, which are aimed at achieving maximum reusability of project outcomes. In this paper, we present our experiences and the domain-specific aspects of RDM planning for such a large consortial project in simulation-based climate science – the second phase of the PalMod project, PalMod-II.¹

The PalMod-II project was a German initiative aimed at performing novel, fully-coupled, global-climate model simulations, and enabling the intercomparison and evaluation with paleo-climate proxy data. PalMod-II data products include the output of three state-of-the-art coupled climate models of varying complexity and spatial resolutions simulating the climate of the past 130,000 years. This output typically comes in the form of continuous, geospatially-gridded numerical data. It is very large in volume (hundreds of Terabytes per simulation) and can be standardised in a mostly straight-forward manner, due to existing and widely-accepted community (meta)data standards. In addition to the long time series of model data, a comprehensive compilation of paleo-proxy data was produced to facilitate model-model and model-proxy intercomparison and evaluation. Paleo-proxy data come in the form of tree rings, pollen records, ice cores, marine organism carbon dating, and ocean or lake sediment cores (e.g., Ramisch et al. (2020), Reschke et al. (2021), and references therein). Unlike the model simulation output, these data are small in volume but standardisation for efficient reuse, e.g., for evaluating the PalMod-II model simulations, is challenging due to a lack of community-agreed (meta)data standards.

In contrast to PalMod-I, where dedicated RDM-activities were not a part of the project, PalMod-II is supplied with dedicated resources to handle the requirements of internal project and data management, as well as the definition of a project-wide RDM strategy for data handling and dissemination to the global climate science research community. Starting from the definition of the data request received as input from the working groups, the project specific DMP was compiled for PalMod-II. The DMP also includes definitions of workflows of standardization, publication, and long-term archiving of the PalMod-II coupled climate model data. The applied DMP approach is particularly useful as a reference to the rather new profession category of climate data managers/stewards, along with other large multidisciplinary natural science consortiums.

The data generated in flagship projects such as PalMod-II needs to be made available for efficient reuse to the community, which requires that the RDM activities laid out in the DMP are conformant to the FAIR principles (Wilkinson et al., 2016). As domain-specific indicators/metrics for FAIRness are still not fully agreed upon, they can therefore not be evaluated following standard approaches (de Miranda Azevedo and Dumontier (2020), Peters-von Gehlen et al. (2022)). Domain-specific DMPs focusing on FAIR project data represent valuable resources not only for the project scientists, but for the respective domain because they have the potential to drive the adoption of FAIR-aligned community practices.

In the following, we introduce the necessities of RDM-related activities in large consortial projects (Section 2), detail the planning of RDM in PalMod-II (Section 3), and present a discussion and our conclusions (Section 4).

2. ORGANISATION OF DATA-RELATED ACTIVITIES IN LARGE PROJECTS

The scope and diversity of Earth System Science (ESS) research directly leads to the primary data management challenge: dealing with high volumes of heterogeneous, non-systematic datasets. Within large interdisciplinary projects where various scientists work simultaneously on individual work packages and handle (produce, compile, exchange, store) the project data, a project-wide systematic RDM approach is needed. The primary goal of the project-wide, data management approach is to support different and often independent sub-project level workflows and transform them from multiple fragmented, task-oriented workflows to an integrated, project-oriented plan.

¹ <https://www.palmod.de/home>.

This plan is documented in the project-wide data management plan (DMP) which should serve as fulfilment of funder requirements and as a reference manual for the project scientists, providing a guiding reference for managing the project data. Depending on the progress of the project, data management approaches and practices may have to be adopted during project runtime. These are then reflected in the continuously updated DMP, which is preferably managed using software packages allowing for version control, e.g. RDMO (Anders et al., 2022a). RDM updates heavily depend on effective intra-project communication practices.

Apart from being a requirement imposed by research funding bodies, the compilation and maintenance of a DMP serves to document and organise the data flows and data dependencies in data intensive, collaborative projects. Especially if the project goals critically depend on the interaction between multiple working groups distributed over several institutions. The PalMod-II project therefore represents an excellent use case to illustrate the benefits of project-wide RDM in simulation-based climate science, for which the established DMP is the focal point guiding all activities like data creation, sharing, preservation, and reuse for the benefit of all project partners, as well as the general research community.

Owing to the data intensive nature of state-of-the-art (simulation-based) research projects in ESS (project data volumes can easily be on the order of PBs), RDM planning must also consider the most efficient use of available IT-resources and provide input for annual computational resource applications at IT service providers, e.g. DKRZ² (a topical infrastructure provider in ESS). Again, intra-project communication is key. The climate model simulations within PalMod-II were performed at different HPC (high-performance computation) facilities across Germany and, therefore, the data transfer workflow definition (across these HPC platforms) was also considered as an important component of the DMP.

Ultimately, the project outcomes should be made available for reuse by the scientific community according to the FAIR data guiding principles (Wilkinson et al., 2016). Depending on the project goals, the data made available can be anything from just the bare minimum of primary data to comprehensive data sets describing various aspects of the Earth System, e.g. model simulation output of global coverage spanning centennial or even millennial timescales. Planning the preparation of the datasets for publication and long-term archival in-line with the FAIR principles is, therefore, essential.

Planning and executing the above-mentioned aspects of RDM in large projects requires specialist expertise and can typically not be achieved by project researchers. Therefore, it is important to assign RDM experts to the task. This can either happen by means of institutional support or, even better, domain-expert RDM support staff hired specifically for the project.

The following aspects have to be considered in planning RDM in-line with the FAIR data guiding principles for large, data-intensive projects to provide support towards the project goals:

1. Data generation (definition of data formats and metadata standards for the generated/collected/reanalysis data).
2. Data sharing and exchanging within the project working groups (inter-dependencies, accountability, model/code/algorithm development tasks across various sub-groups within a project).
3. Data analysis (including evaluation or inter-comparison).
4. Defining data publication standards (usage, licences, confidentiality, legal issues, copyright).
5. Data dissemination (via repositories, servers, clouds/databases).
6. Archival and re-use after the project lifetime (long term).

In the present use case, we present the RDM planning and execution within the project PalMod-II along the above aspects in section 3.

3. RDM PLANNING FOR PALMOD-II

For organising PalMod-II RDM, a dedicated project working group, the “crosscutting data management group” (CC-DM), was part of the projects’ organisational structure.

² <https://www.dkrz.de/en>.

The need for such a dedicated working group was realised during the PalMod-II application phase. In PalMod-I (2016–2019), RDM activities did not play a significant role which was clearly identified as a disadvantage hindering the project progress. Furthermore, as the envisaged end goal of PalMod (PalMod is planned as a three-phase project, with the third phase starting in 2023) is the dissemination of flagship, paleo-climate datasets for reuse by the global community, RDM approaches and practices in-line with the FAIR principles are of particular relevance. Specifically, PalMod-II RDM ensured that produced datasets conform to discipline-specific (meta)data standards (I, R), are published and made available to the global community via the Earth System Grid Federation (ESGF) infrastructure under CC-BY 4.0 licence (F, A, R), have globally unique PIDs assigned at the dataset-level (F) and shall be archived for long-term reuse in the World Data Center for Climate (WDCC [Re3data:10.17616/R3989R](https://re3data.org/10.17616/R3989R)), a FAIR-aligned CoreTrustSeal-certified discipline-specific repository (see Petersvon Gehlen et al. (2022) for a detailed analysis of the FAIRness of WDCC-archived data).

The necessary impact of CC-DM was ensured by installing a PalMod-II-funded core data management team at DKRZ. Representatives from all other PalMod-II work packages supported the RDM process through communication and promotion of RDM practices in their everyday work environments. At the beginning of the project, we used email surveys to collect information on the research workflows, generated data and data interdependencies between subprojects (work packages). Communication at different levels was organised starting from the principal investigator to university students involved within the project. After a first set of collected surveys, we organised a project-wide RDM workshop (February 2020) in order to collect detailed information on individual process chains, workflows, and inter-dependencies required to advance towards the project goal. The workshop discussion report (and the resulting RDM strategy) was published on the project website (Ref: [PalMod official project website](#)) for any working groups' reference. Apart from identifying PalMod-II's data flows, the collected information also served as input for the initial version of the project-wide DMP, which was due as deliverable after the first half year of PalMod-II's project runtime. We note here that this DMP was not required by the project funder, but was included to facilitate RDM practices in the project.

Further, a direct application of the DMP can also be found in the IT resource application at the DKRZ. Only thanks to a carefully elaborated DMP could the scientific steering committee of the DKRZ be convinced to grant computational resources to this project, which enabled significant parts of the scientific project work to be performed at all.

We used the web based RDMO (Research Data Management Organizer) tool ([Anders et al., 2022a](#)) for creating and maintaining the PalMod-II DMP. RDMO offers the unique possibility to adapt the question catalogues and DMP templates to the specific project needs. For PalMod-II, we designed a question catalogue and DMP template suited for projects in simulation-based climate science to generate the first version of the DMP. Further adaptation of the defined DMP template was carried out over the last three years based on the progress of the project.

The overall PalMod-II RDM concept addressing all relevant aspects required to progress towards the project goal of supplying flagship paleo-climate datasets to the scientific community in-line with the FAIR principles, is shown in [Figure 1](#). The individual aspects shown in [Figure 1](#) are detailed in Section 3.1 – 3.6.

3.1 DATA GENERATION

Data generation in PalMod-II comprised the simulated coupled climate model data as well as the integrated paleo-proxy reconstructions, and the restart/forcing data required to perform the Earth System Model (ESM) simulations (e.g. land-sea-ice mask, greenhouse gas concentrations, aerosol concentrations, vegetation distribution etc.). To make most efficient use of available resources and reduce overheads, the ESM simulations are set up to output only fields actually required for downstream reuse. Definition of these output fields was achieved in iterative communication rounds among PalMod-II scientists and the CC-DM team. The timelines for data generation, lists of ESM output parameters, expected data volumes, documentation of project-internal interdependencies, and naming of responsible scientists are included in the DMP for reference purposes and project guidance.

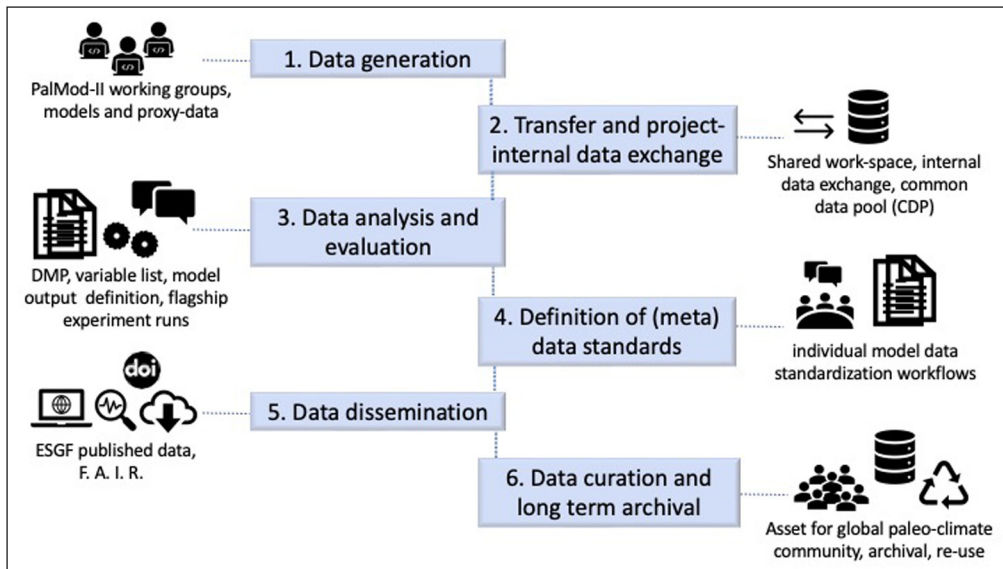


Figure 1 Research Data Management (RDM) Workflow within PalMod-II.

3.2 DATA SHARING AND EXCHANGE

During the active phase of the research, sharing and exchange of data among the project partners is vital to ensure targeted progress of the project. Along with defining clear interdependencies across the sub-projects, the RDM practices documented in the DMP also ensured that the data sharing process is easy and unambiguous. Using storage resources provided by DKRZ, a common data pool (CDP) was established specifically for PalMod-II in order to facilitate project-internal sharing of ad hoc data (intermediate, in-process model development/evaluation tasks across various groups). The CDP also provided an opportunity for project-internal data re-use, for example, the paleo-proxy reconstructions to be used in model development and evaluation. This was supported by a data transfer, access and retrieval recipe documented in the DMP, which also includes conditions for project-internal data reuse. The data hosted in the CDP proved especially essential to the CC-DM efforts to plan and test data standardisation workflows required to prepare the PalMod-II data for publication, in-line with the FAIR principles (see below).

3.3 DATA EVALUATION/ANALYSIS

Various evaluation and intercomparison processes were defined and performed in PalMod-II (model inter-comparison between the three PalMod-II models and model-data comparison with paleo-proxy data/observations). As a project deliverable, a defined list of model data (simulated model output variables) is a part of the DMP. Planning and executing the early definition of the list ensured efficient analysis and evaluation for intercomparable simulations within the working groups.

3.4 DATA STANDARDISATION

Publication of the PalMod-II ESM output via the Earth System Grid Federation (ESGF, Cinquini et al. (2014)) infrastructure, and subsequent long-term archival in the WDCC requires that the (meta)data conform to discipline-specific standards. Specifically, publication via ESGF requires alignment with CMIP conventions, which includes compliance with the Climate and Forecast (CF) conventions (Eaton et al., 2022). This is to ensure efficient findability, accessibility, interoperability, and reusability (FAIR by design) of published ESM output.

Planning the preparation of PalMod-II ESM-output to fulfil these requirements was a major task for the CC-DM team. Project-internal communication resulted in the determination of publication licenses, data usage restrictions, confidentiality agreements, and legal aspects. These aspects, together with conditions to maintain the data with embargo status until a scientific publication is accepted, are documented in the DMP. Because the three ESMs used in PalMod-II all produce differently structured raw output, planning and defining data standardisation workflows was an essential task for the CC-DM team in PalMod-II. The workflow designs arose out of the ongoing work in PalMod-II to reflect the needs of the scientists and the RDM requirements.

Ultimately, the implemented standardisation workflow is based on CMOR (Climate Model Output Rewriter, Doutriaux and Taylor (2010)), used within CMIP (Coupled Model Intercomparison Project (Meehl and Hibbard, 2007, Eyring et. al., 2016)), which is an established tool for defining and applying (meta)data standards to climate model data. The workflows had to be slightly adapted for different models and these are documented in the continuously updated DMP.

3.5 DATA DISSEMINATION

The PalMod-II model data available at the time of writing were standardised according to the workflows defined in Section 3.4 and subsequently published via ESGF,³ including the assignment of persistent identifiers, making it FAIR.

3.6 CURATION AND ARCHIVAL

Long-term archived data must be accompanied by appropriate metadata that describe their provenance so that other users are able to find them and understand them after the project lifetime (F and R aspects of FAIR). It is essential that the persons/institutes responsible for the data curation, after the project ends, are clearly mentioned in the DMP. Long-term archiving of PalMod-II data is done via the WDCC (cf. Section 3). Since the WDCC is hosted by the DKRZ, planning the process of long-term archiving for PalMod-II data was very efficient and did not produce large overhead.

Planning for the long-term archival of PalMod-II data in the WDCC is essential for compliance with the FAIR principles, as it focuses on the long-term reusability of data, originating from simulation-based climate science (Peters-von Gehlen et al. (2022)). WDCC-archived data have DOIs assigned and come with rich metadata such as information concerning citation (authors, links to related publications), provenance, contact, and data quality. Since the data files are the same as those published via the ESGF, all benefits associated with the highly-standardised data also apply to the WDCC-archived data.

Applying the RDM concept, as detailed above, made sure that all the data related issues were defined, continuously updated if needed, and made available to the project stakeholders. Appropriate data support was provided to the data creators as well as the data users. Unambiguous and clear data usage policies, data access, licensing, and the archival strategy were defined within the DMP, ensuring that the appropriate credit is given to the data creators (scientists and climate modellers in case of PalMod-II).

End products of PalMod-II, which consist of unique long term scientific paleo-climate data, are kept for re-use by the paleo-climate research community as well as other research disciplines (e.g., land-use, socio-economic studies etc.).

4. DISCUSSION AND CONCLUSION

PalMod-II was a project initiating a novel chapter in paleo-climate research by simulating the last glacial cycle using coupled climate models and paleo-proxy reconstruction data to enable more credible climate projections for the next millennia. A variety of research groups with diverse scientific focus worked within PalMod-II to establish and evaluate the climate model simulations. Being a project where the final end product is large-scale climate research data for re-use by the global scientific community, a dedicated RDM-concept, including the compilation and maintenance of a DMP as a living document, was applied within the cross-cutting working group for PalMod-II data management.

The RDM-concept in PalMod-II was centred around two main pillars: (i) compilation and maintenance of the DMP during the project and (ii) organising and facilitating the entire process associated with making the PalMod-II-produced datasets available to the global community in-line with the FAIR data principles. The latter, especially, involved continuous and intense communication with the project scientists to achieve maximum alignment with the needs of the project scientists and prospective data-re-users. In other words, the success of an RDM approach largely depends on the acceptance of RDM principles and their consistent application/compliance by all the stakeholders.

3 Earth System Grid Federation <https://esgf-data.dkrz.de/search/palmod/>.

The comprehensiveness of the executed RDM-approach exercised for PalMod-II was first-of-its-kind for the participating institutions. Thus, we list a couple of lessons-learned for future related projects and associated RDM requirements:

- 1) Without properly set-up communication in a large project consisting of several partnering institutions, a unified RDM approach (and also the project management itself) becomes very difficult to achieve.
- 2) The DMP clearly demonstrated the intra-project dependencies between the individual work packages including timelines, a feature well acknowledged by the project scientists.
- 3) The DMP significantly enhanced the efficiency of compiling IT resource requirements needed for the annual application for HPC resources at DKRZ.
- 4) As a living document, the DMP evolved as the central collection point of project decisions and actions taken towards the project (RDM) goal of making project data available in-line with the FAIR principles (see update history of the DMP submitted as supplementary material showing the timeline of DMP updates during the project duration), thus enhancing the transparency of project progress.
- 5) Domain-specific expertise of data management staff (data stewards) responsible for the RDM in a specific project is essential for achieving the project goals, because communication has to be at a semantic level comprehensive for domain scientists, and knowledge of typical workflows in the scientific field (Anders et al., 2022b) is a plus.

The PalMod-II RDM concept enables common RDM according to the FAIR data principles across all the working groups of large consortial projects using common workflows for the exchange of data and information along the process chain. Therefore, the approach detailed here is amenable for reuse by other, similar projects. In particular, the domain-specific project-DMP can be re-used or adapted for other spin-off projects of similar nature, making it a sharable and reusable asset for the scientific community.

More specifically, PalMod-III is scheduled to commence in the first half of 2023, and the RDM approach of PalMod-II will be continued.

FUNDING INFORMATION

The authors gratefully acknowledge the support from the German Federal Ministry of Education and Research (BMBF grant number 01LP1926, Project PalMod-II CC-TP1).

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR AFFILIATIONS

Swati Gehlot

Deutsches Klimarechenzentrum GmbH, DE

Karsten Peters-von Gehlen

Deutsches Klimarechenzentrum GmbH, DE

Andrea Lammert

Deutsches Klimarechenzentrum GmbH, DE

Hannes Thiemann

Deutsches Klimarechenzentrum GmbH, DE

REFERENCES

- Anders, I, Enke, H, Hausen, D, Henzen, C, Jagusch, G, Lanza, G, Michaelis, O, Peters-von Gehlen, K, Rathmann, T, Rohrwild, J, Schönau, S, Wedlich-Zachodin, K and Windeck, J. 2022a. "The Research Data Management Organizer (RDMO) – the Software and the Community behind Germany's Leading DMP Tool". *Data Science Journal*, in revision.
- Anders, I, Peters-von Gehlen, K and Thiemann, H. 2022b. Canonical Workflows in Simulation-based Climate Sciences. *Data Intelligence*, 4(2): 212–225. DOI: https://doi.org/10.1162/dint_a_00127

- Cinquini, L, Crichton, D, Mattmann, C, Harney, J, Shipman, G, Wang, F, Ananthakrishnan, R, Miller, N, Denvil, S, Morgan, M, Pobre, Z, Bell, GM, Doutriaux, C, Drach, R, Williams, D, Kershaw, P, Pascoe, S, Gonzalez, E, Fiore, S and Schweitzer, R. 2014. The Earth System Grid Federation: An open infrastructure for access to distributed geospatial data. *Future Gener. Comp. Sy.*, 36: 400–417. DOI: <https://doi.org/10.1016/j.future.2013.07.002>
- de Miranda Azevedo, R and Dumontier, M. 2020. Considerations for the Conduction and Interpretation of FAIRness Evaluations. *Data Intelligence*, 2(1–2): 285–292. DOI: https://doi.org/10.1162/dint_a_00051
- Doutriaux and Taylor. 2010. Climate Model Output Rewriter (CMOR), Version 2.0 (CMOR2), Livermore, CA: Program for Climate Model Diagnosis & Intercomparison (PCMDI). (2010).
- Eaton, B, Gregory, J, Drach, B, Taylor, K, Hankin, S, Blower, J, Caron, J, Signell, R, Bentley, P, Rappa, G, Höck, H, Pamment, A, Juckes, M, Raspaud, M, Horne, R, Whiteaker, T, Blodgett, D, Zender, C, Lee, D, Hassell, D, Snow, AD, Kölling, T, Allured, D, Jelenak, A, Soerensen, AM, Gaultier, L and Herlédan, S. 2022. NetCDF Climate and Forecast (CF) metadata conventions. [Available at: <http://cfconventions.org/Data/cf-conventions/cf-conventions-1.10/cf-conventions.pdf>].
- Eyring, V, Bony, S, Meehl, GA, Senior, CA, Stevens, B, Stouffer, RJ and Taylor, KE 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, 9: 1937–1958. DOI: <https://doi.org/10.5194/gmd-9-1937-2016>
- Meehl, GA and Hibbard, K. 2007. A Strategy for Climate Change Stabilization Experiments with AOGCMs and ESMs. Report from the Aspen Global Change Institute 2006 Session—Earth System Models: The Next Generation, (July 30–August 5, 2006) and joint WGCM/AIMES Steering Committee Meeting 27 September (2006, 2007).
- PalMod official project website. [Available at: www.palmod.de].
- Peters-von Gehlen, K, Höck, H, Fast, A, Heydebreck, D, Lammert, A and Thiemann, H. 2022. Recommendations for Discipline-Specific FAIRness Evaluation Derived from Applying an Ensemble of Evaluation Tools. *Data Science Journal*, 21(1): 7. DOI: <https://doi.org/10.5334/dsj-2022-007>
- Ramisch, A, Brauser, A, Dorn, M, Blanchet, C, Brademann, B, Köppl, M, Mingram, J, Neugebauer, I, Nowaczyk, N, Ott, F, Pinkerneil, S, Plessen, B, Schwab, MJ, Tjallingii, R and Brauer, A. 2020. VARDA (VARved sediments DATabase) – providing and connecting proxy data from annually laminated lake sediments. *Earth Syst. Sci. Data*, 12: 2311–2332. DOI: <https://doi.org/10.5194/essd-12-2311-2020>
- Re3data.org: World Data Center for Climate; editing status 2022-08-25; re3data.org – Registry of Research Data Repositories last accessed: 2023. DOI: <http://doi.org/10.17616/R3989R>
- Reschke, M, Kröner, I and Laepple, T. 2021. Testing the consistency of Holocene and Last Glacial Maximum spatial correlations in temperature proxy records. *J. Quaternary Sci.*, 36: 20–28. DOI: <https://doi.org/10.1002/jqs.3245>
- Wilkinson, M, Dumontier, M, Aalbersberg, I, et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*, 3: 160018. DOI: <https://doi.org/10.1038/sdata.2016.18>

TO CITE THIS ARTICLE:

Gehlot, S, Peters-von Gehlen, K, Lammert, A and Thiemann, H. 2023. Data Management for PalMod-II – A FAIR-Based Strategy for Data Handling in Large Climate Modeling Projects. *Data Science Journal*, 22: 34, pp. 1–8. DOI: <https://doi.org/10.5334/dsj-2023-034>

Submitted: 04 January 2023

Accepted: 25 May 2023

Published: 11 September 2023

COPYRIGHT:

© 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Data Science Journal is a peer-reviewed open access journal published by Ubiquity Press.