25 years IPCC Data Distribution Centre at DKRZ and the Reference Data Archive for CMIP data

Martina Stockhause[1], Michael Lautenschlager[1]

[1]German Climate Computing Center (DKRZ), Hamburg, 20146, Germany

5   *Correspondence to*: Martina Stockhause (stockhause@dkrz.de)

Abstract. The Data Distribution Centre (DDC) of the Intergovernmental Panel on Climate Change (IPCC) celebrates its 25th anniversary in 2022. DKRZ is the ~~last~~ only remaining ~~founding member among the~~ DDC Partners from the original group jointly managing the DDC. In spite of changes in prioritization, it has been supporting the IPCC Assessments and long-term preserving the quality-assured, citable climate model data underpinning the Assessment Reports over these years. An active

10 and engaged collaborative community achieved advances in ~~The continuous d~~data standardization, data management best practices, and infrastructure developments. These evolving standards are reflected in the activities of the DDC. ~~Examples for such milestones are the NetCDF/CF data standard, the DataCite data DOIs enabling data citation, the data preservation and stewardship standards of the World Data System (WDS), or the IPCC FAIR Guidelines for the current Sixth IPCC Assessment Report (AR6)~~. The introduction of the IPCC FAIR Guidelines into the current Sixth IPCC Assessment Report (AR6) has

15 significantly changed the role of the DDC Partner DKRZ from an independent partner for long-term data preservation into an active partner involved in IPCC's Sixth Assessment cycle. As a result, the DDC has gained exposure and visibility, posing a challenge and an opportunity to operationalize IPCC's FAIR Guidelines and long-term preservation approaches. While the value of DDC services has been recognized, DDC sustainability remains unresolved and is currently being discussed within the IPCC as part of a general AR6 review process to formulate recommendations for AR7 data management. ~~Combined with a~~

20 ~~long-term DDC strategy, the DDC can be positioned in the developing FAIR digital object research landscape as well as in the political and societal realm of climate change.~~

1   History of TG-Data and the IPCC DDC

The current Data Distribution Centre (DDC, https://ipcc-data.org/, last access: 10 March 2022) of the Intergovernmental Panel on Climate Change (IPCC) is jointly managed under a Memorandum of Understanding (Xing, 2021) by four partners: the

25 German Climate Computing Center (DKRZ, Germany), the Center for International Earth Science Information Network (CIESIN, USA), the Spanish Research Council (CSIC, Spain), and MetadataWorks (UK). DKRZ is the only remaining founding partner, and it~~who~~ has been now operating the DDC for 25 years. ~~The DDC Partners signed a Memorandum of Understanding defining the DDC's operations (Xing, 2021).~~ The DDC is overseen by the Task Group on Data Support for Climate Change Assessments (TG-Data, https://ipcc.ch/data, last access: 10 March 2022), which is a non-permanent part of

30 IPCC's structure (Fig. 1). TG-Data member experts are complemented by ex-officio Members representing the DDC Partners and the Technical Support Units (TSU) of the three IPCC Working Groups (WG). The core role of the DDC is the support for IPCC authors and ~~data~~ users of~~in~~ data and scenarios underpinning IPCC outputs (see DDC Guidance, IPCC, 2018a). DDC makes special efforts to support~~, especially those~~ users in developing countries.

The DDC was formally established at the Thirteenth Session of the IPCC (IPCC-13) on 22 and 25-28 September 1997 in the

35 Maldives (IPCC, 1997). Deutsches Klimarechenzentrum (DKRZ) in Germany and the Climatic Research Unit (CRU) in the United Kingdom were selected to execute~~as a~~ shared DDC operation and the Finnish Meteorological Institute (FMI) was to contribute guidance and training (see figure 2). During the Second Assessment Report (SAR; IPCC, 1995) cycle, IPCC Working Group II (WGII) had requested for lowering the barriers to using data from future climate scenarios provided by the

Coupled Model Intercomparison Project (CMIP) of the World Climate Research Programme (WCRP). On the IPCC Workshop on Regional Climate Change Projections for Impact Assessment (London 24-26 September 1996) and subsequent meetings of the established IPCC Task Group on Climate Scenarios for Impact Assessment (TGCIA), requirements for data availability, data standardization and data quality together with the need for guidance materials were formulated. TGCIA recommended

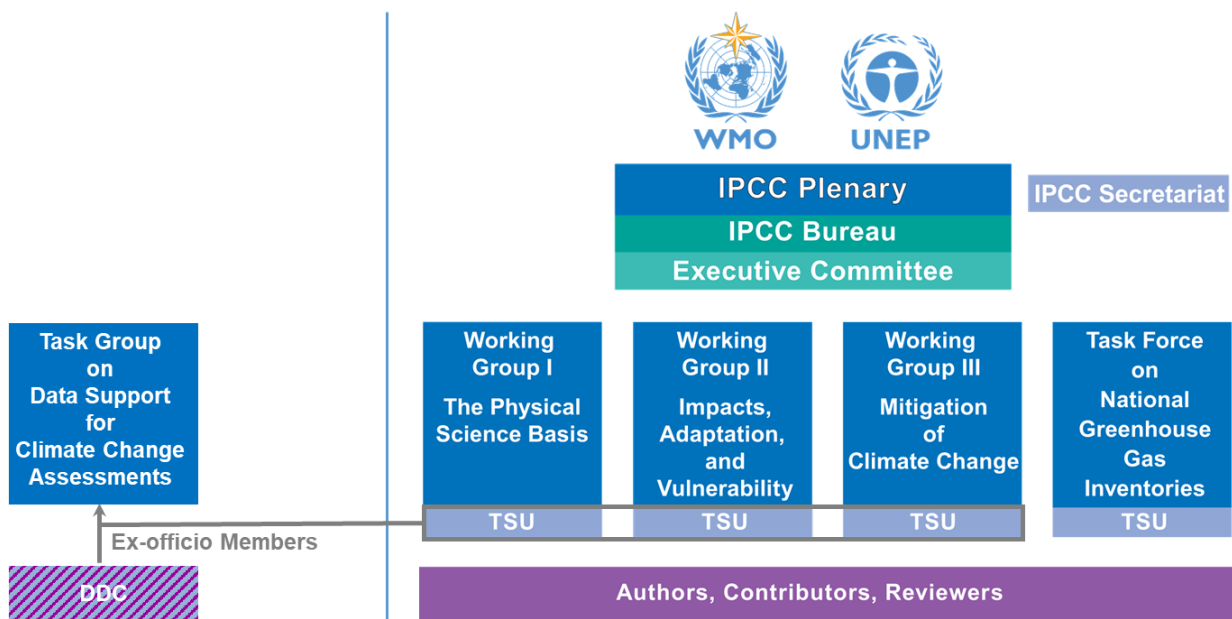



Figure 1: IPCC Structure (from IPCC webpage)

the establishment of a DDC to the IPCC Bureau in the same year. The IPCC Bureau requested governments on 22 July 1997 to nominate institutions to act as the DDC and to provide the necessary financial support to establish and maintain the DDC function.

The DDC Partners agreed that DKRZ should be responsible for the global climate model data related to CMIP, while the other DDC Partners took up the responsibility for other datasets in support of the IPCC WGs. Atmospheric near surface variables were collected, aggregated and disseminated by the DDC together with guidance material for the Third Assessment Report (TAR; IPCC, 2001). After this successful operation of the DDC, the IPCC Bureau received data requests from WGIII and WGI. WGIII's data request was similar to that of WGII and could be integrated into the existing DDC service while WGI's data needs were more complex, requesting most of the CMIP datasets. The Program for Climate Model Diagnosis & Intercomparison (PCMDI) accepted to establish the CMIP3 data archive in support of WGI. In cooperation with PCMDI, the DDC Partner DKRZ extracted the CMIP3 data subset for WGII and WGIII from the CMIP3 archive at LLNL for the Reference Data Archive of the Fourth Assessment Report (AR4; IPCC, 2007). In the Fifth Assessment Report (AR5; IPCC, 2013) cycle, increasing CMIP5 data volumes led to the development of a federated data archive and the ESGF (Earth System Grid Federation) data infrastructure. The DDC refocused on the long-term preservation of the CMIP5 data underpinning the AR5 by transferring the CMIP5 data at the WGI snapshot date into the DDC AR5 Reference Data Archive. For nearly two decades, the IPCC DDC has closely cooperated with the CMIP data infrastructure.

TThe Task Group TGCIA was renamed thee Task Group on Data and Scenario Support for Impact and Climate Analysis (TGICA) in 2003. TGICA developed a vision paper in 2015 and described the challenges TGICA was facing related to its limited capacity to deliver on its mandate. The IPCC Panel decided to revise TGICA's mandate and to hold an IPCC Expert Meeting on the Future of TGICA (IPCC, 2015, 2016). Vaughan (2016) summarizes TGICA's challenges and emphasizes that TGICA needs to be strengthened to be able to contribute in new ways to improve the access and use of climate data and scenarios for research and decision making through the DDC. A sharpened mandate, the clear identification of specific goals, and a realistic sense of the resources required to accomplish these goals are recommended. An IPCC Ad-hoc Task Force

TGICA took up the results from the IPCC Expert Meeting and formulated revised Terms of Reference for the reestablished Task Group on Data Support for Climate Change Assessments (TG-Data) and a revised Guidance for the DDC, which were approved on IPCC-47 ~~the 47th IPCC Plenary session in Paris, 13-16~~in March 2018 (IPCC, 2018ba). The focus of the DDC was narrowly focussed on data support tasks.

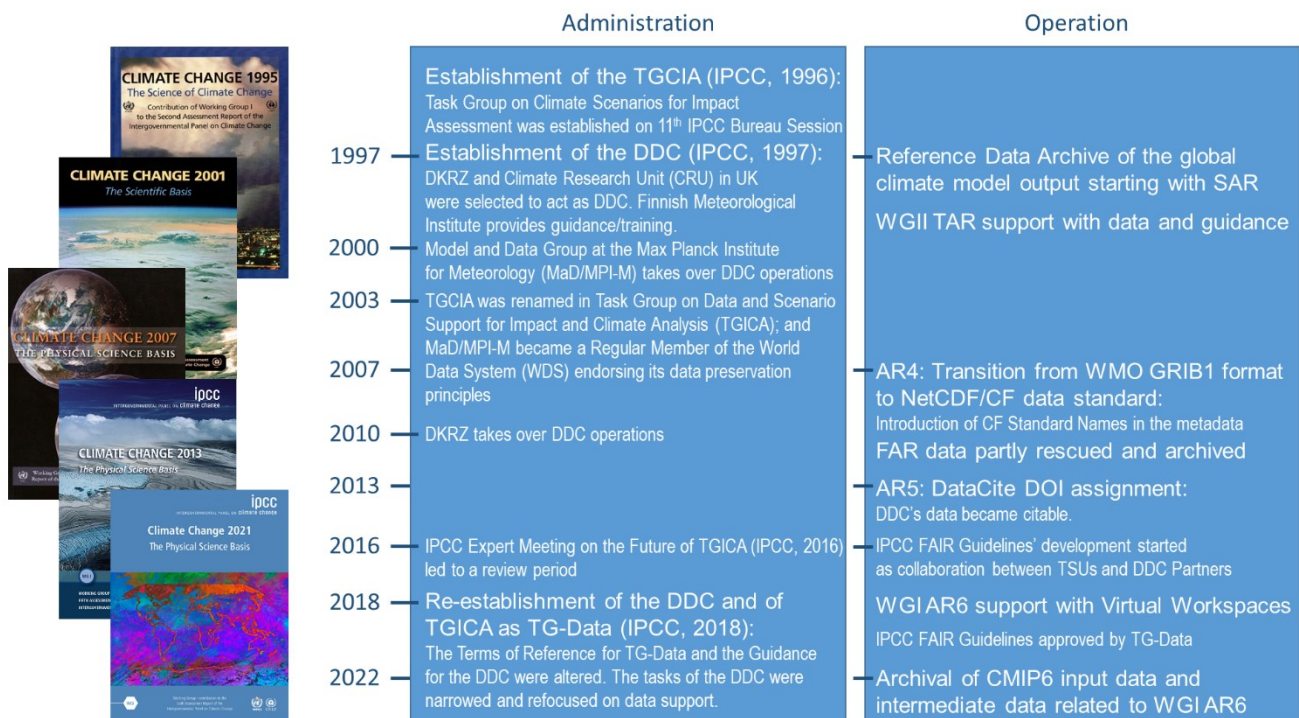



Figure 2: IPCC DDC history and main achievements of the DDC Partner DKRZ over the past 25 years (images from cover pages of the IPCC ARs starting with SAR 1995)

2   The Reference Data Archive at the DDC at DKRZ

DKRZ is the DDC Partner responsible for the long-term preservation of the global climate model data, ~~which is strongly related to Working Group I (WGI) of the IPCC and~~ provided by ~~the research project Coupled Model Intercomparison Project (CMIP) of the World Climate Research Programme (WCRP)~~CMIP. Starting with the Second Assessment Report (SAR; IPCC, 1995), core variables for the characterization of the state of the Earth System (table A1 in appendix A) from ~~the~~ model ~~applications for~~projection of the future climate ~~projections~~ were long-term archived at DKRZ, building the Reference Data Archive for the global climate model data underpinning IPCC's ARs~~-~~.

During data archival~~During the long-term archival~~, the data are stored on tape and the metadata are enriched and quality-assured to provide sufficient and high-quality information for various downstream users without specific knowledge of climate model applications. Added metadata include context information on project, experiments and models as well as discovery information on spatial-temporal coverage, parameters and contact information. In SAR, this information was gathered mostly from the data providers by the DDC. With the increasing level of organization and standardization of CMIP, this labor-intensive and non-standardized metadata gathering from data providers could be partially replaced by machine-access of CMIP resources, e.g. accessing the ESGF index.

The size of the Reference Data Archives for the different ARs increased from around 10 GBytes and 400 datasets for SAR and TAR to~~via~~ ca. 1 TByte and 1 500 datasets for AR4 and then to 1.7 PBytes and 910 000 datasets for AR5 (Fig. 32). The reasons are an increased number of archived variables per model run, an increased number of models participating in CMIP, and the inclusion of ~~change in archival strategy from monthly to~~daily and sub-daily data in addition to monthly data. A subset of data

95 underpinning ~~Data of~~ the FAR were rescued in their original formats in collaboration with NCAR (National Center for Atmospheric Research) from its data archive and added to the DDC in 2008 ~~in 2008 in the original formats, in which they were rescued. As it was not possible for all datasets to identify whether these data were available at the time of FAR preparation, only a subset of the available datasets was archived in the DDC.~~ Because of the low level of standardization, these datasets are difficult to (re-)use. Data underpinning the IPCC Special Report on Global Warming of 1.5°C (SR1.5; IPCC, 2018c) were transferred into the DDC Reference Archive in 2018. The archival of the CMIP6 data subset underpinning the AR6 is ongoing.

100 Download statistics show the long-term interest of users in the DDC Reference Data (Fig. 4~~3~~).

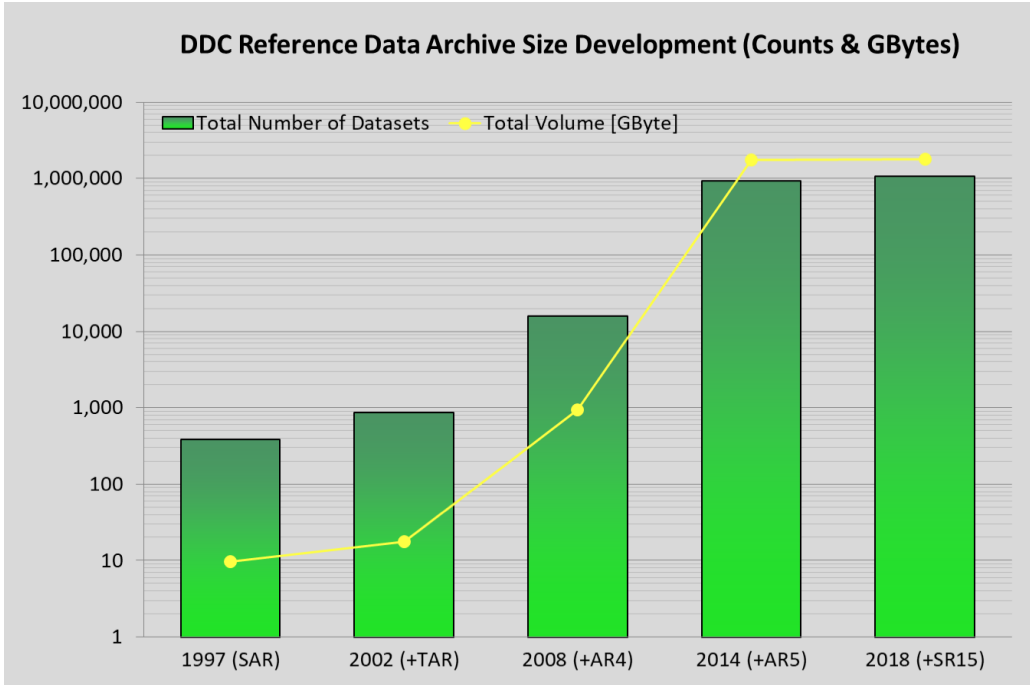
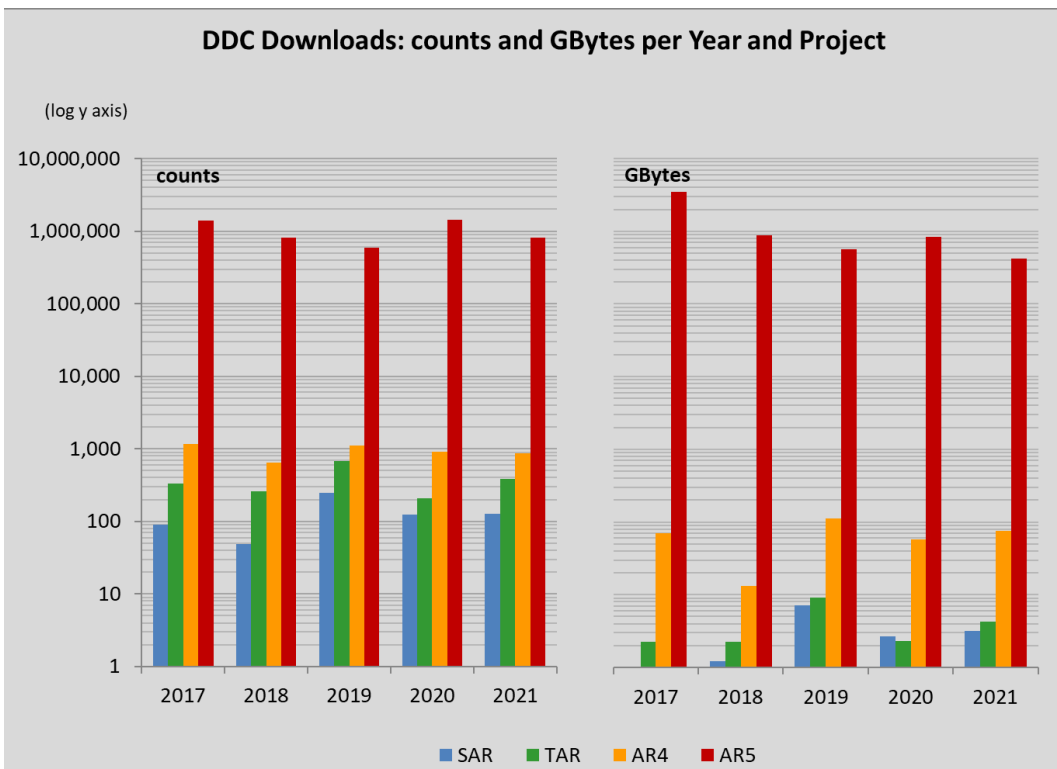



Figure 3~~2~~: Size development of the DDC Reference Data Archive for the global climate model data.

105 **Figure 4~~3~~: Downloads in number of datasets [counts] and data volume [GByte] from the DDC Reference Data Archive over the last 5 years per Assessment Report (~~AR6 incl. SR1.5 and~~ FAR is~~are~~ left out because of the incomplete Reference Archives~~ and AR6~~ including SR1.5 because of the ongoing data archival).**

As ~~a~~ DDC Partner, DKRZ ~~has~~ committed to ~~keep~~ ensuring its DDC data ~~holdings~~ remains accessible and reusable o~~ver~~n the long-term, which involves cyclic renewal of hardware, continuous maintenance of software, and metadata and data curation. A copy of the data on tape is stored off-site at the Max Planck Society's computing center in Garching, Germany. New generations of hardware (tape system), for example, require the copying of the DDC data holdings on new cartridges. Software updates for data discovery, access and exchange are required to comply ~~with~~to new standards and interfaces in order to enhance the user experience and to meet evolving user needs. An example for a metadata curation measure was the addition of Climate and Forecast (CF) standard names to the metadata of the DDC SAR and TAR Reference Data Archives. To overcome the data volume barrier for DDC data reuse of IPCC users located in developing countries with low internet bandwidths, the DDC introduced a service whereby users can order a set of preselected variables for seven regions on DVD and USB devices.~~An important metadata curation measure was the introduction of the Climate and Forecast (CF) Standard Names (https://cfconventions.org/Data/cf-standard-names/current/build/cf-standard-name-table.html, last access: 10 March 2022) and the mapping of proprietary model variable names to this new community standard in the DDC metadata. The data format of the datasets in the DDC reflects the change in the community's standard data format from the WMO GRIB1 (https://community.wmo.int/activity-areas/wis/grib-edition-1, last access: 10 March 2022) standard used for SAR and TAR towards NetCDF/CF (https://cfconventions.org/Data/cf-conventions/cf-conventions-1.9/cf-conventions.html, last access: 10 March 2022) used for AR5 with the provision of the datasets in both formats for AR4. In support of the IPCC users located in developing countries with low internet bandwidths, the DDC introduced the service of sending DVD and USB devices with preselected variables for seven regions per airmail. Data of the FAR were added to the DDC in 2008 in the original formats, in which they were rescued. As it was not possible for all datasets to identify whether these data were available at the time of FAR preparation, only a subset of the available datasets was archived in the DDC. Because of the low level of standardization, these datasets are difficult to~~

DKRZ adjusted to evolving best practices for~~new~~ data management ~~best practices evolving over time~~. The DKRZ long-term data archive including the IPCC DDC Reference Data Archive was approved in 2003 as WDC Climate (WDCC) by~~in~~ the ICSU World Data Center system. DKRZ became a Regular Member of the ISC World Data System (WDS, https://www.worlddatasystem.org/, last access: 10 March 2022) in 2008, the year of WDS's establishment. Therefore, the DDC Partner DKRZ complies with WDS's~~the~~ common research repository standards ~~adopted by the WDS, currently the CoreTrustSeal (https://www.coretrustseal.org/, last access: 10 March 2022)~~. With the founding of DataCite in 2009, registering data DOIs in order to make data citable became a community expectation, which was taken up by the DDC Partner DKRZ for ~~For~~ the AR5 Reference Data Archive published in 2013 and 2014~~, DDC Partner DKRZ registered DOIs via the 2009 founded registration agency DataCite, making the data citable~~. The long-term archival of AR5 provided further major changes in the workflow due to the extremely high data volume and several changes in the CMIP5 data infrastructure (https://pcmdi.llnl.gov/mips/cmip5/, last access: 10 March 2022; Taylor et al., 2012):

- The data were disseminated by the newly developed federated and decentral infrastructure of the Earth System Grid Federation (ESGF, https://esgf.llnl.gov, last access: 10 March 2022; Williams et al., 2016);
- detailed model and experiment documentations were gathered from the CMIP5 participants by the Earth System Documentation project (ES-DOC, https://es-doc.org, last access: 10 March 2022); and
- a three level quality control procedure (CMIP5 QC, https://cmip5qc.wdc-climate.de, last access: 10 March 2022) was applied to ensure basic data quality, the consistency of metadata, and metadata conformance with community standards like NetCDF/CF and project standards like the Data Reference Syntax (DRS). Passing the three quality control levels was the prerequisite for the ~~integration of the data in~~acceptance by the IPCC DDC for the IPCC AR5 Data Reference Archive (Stockhause et al., 201~~2~~4).

The size of the CMIP5 data archive required a high level of automation for metadata and data ingest as well as for the quality control checks. New interfaces to the infrastructure components ESGF, ES-DOC and DataCite had to be developed for

150   insertion of use and discovery metadata and data DOI registration. The long-term archived DDC AR5 data were made searchable and accessible through the ESGF, which has become the standard infrastructure for climate-related data. ETH Zurich collected a CMIP5 data subset in support of the IPCC AR5 authors in an alternate data structure. Due to ~~the~~ difficulties ~~y to~~ relating the individual datasets back to the CMIP5 reference datasets, the DDC AR5 Reference Data Archive was supplemented by an~~data archive of the ETH Zurich was added as additional~~ IPCC Working Group I AR5 snapshot ~~to the DDC~~

155   ~~AR5 Reference Data Archive~~. Discussions with the ETH Zurich provided valuable input for the IPCC FAIR Guidelines adopted~~concept~~ for AR6 and the long-term archival of the CMIP6 input data in the DDC~~underpinning the AR6~~.
       The DDC relies in its efforts and services on data provided by CMIP6 participants and on the standardization community efforts of several organizations and institutions. PCMDI led the AMIP and CMIP data standardization, other groups worked on the NetCDF/CF data standard (https://cfconventions.org, last access: 10 March 2022), the CoreTrustSeal research repository

160   standard (https://www.coretrustseal.org/, last access: 10 March 2022) or the DataCite DOI data publishing standard.
       ~~The IPCC DDC provides quality assured, citable IPCC-relevant reference climate model data for all IPCC assessment reports. Other DDC partners provide aggregated climate model data and guidance material to lower the boundary for downstream data users.~~

3   AR6 and the IPCC FAIR Guidelines

165   TGICA was under review of the IPCC from the start of the Sixth Assessment Cycle in January 2016 until the re-established TG-Data held its First Meeting in November 2019. This was little less~~. This was less~~ than a year prior to the original WGI literature and data cut-off date of~~n~~ 30 September 2020, which was postponed to 31 January 2021 due to the COVID-19 pandemic. The lack of the coordinating task group hampered ~~was a difficult situation for~~ the formulation and implementation of the FAIR Guidelines ~~into~~ for the Sixth Assessment Report (AR6).

170   The idea for adopting the FAIR Guidelines was born during the IPCC Expert Meeting on the Future of TGICA in January 2016. The aim was to enhance~~of the FAIR Guidelines is the enhancement of~~ the transparency of the IPCC AR6 and thereby~~us to~~ contribute to IPCC's integrity. The IPCC FAIR Guidelines implement the established data management principles of FAIR (Findable, Accessible, Interoperable, Reusable; Wilkinson et al., 2016) for data and TRUST (Transparency, Responsibility, User Focus, Sustainability, Technology; Lin et al., 2020) for repository operations into the Sixth Assessment cycle. The FAIR

175   data principles describe requirements for datasets to become an integral part of the research environment. The TRUST principles for repositories and its implementation in the CoreTrustSeal complement these essential data properties by best practices for repository operations in long-term data preservation and data stewardship. ~~The CTS certification is part of the Regular Membership application and approval procedure of the ISC World Data System.~~
       The ~~concept~~ development of~~for~~ the FAIR Guidelines started at the First IPCC AR6 Data Workshop in Hamburg, Germany,

180   19-20 September 2017 (Stockhause et al., 2017) and continued at the second virtual meeting on 20 February 2018. In collaboration with the WDS, which started at the Data Repository Day 2018 (WDS, 2018), the FAIR Guidelines concept was formulated in Stockhause et al. (2019). This concept was discussed with IPCC authors of WGI and WGII on the IPCC Expert Meeting on Assessing Climate Information for Regions in Trieste, 16-18 May 2018 (IPCC, 2018d~~b~~). The implementation of the FAIR Guidelines into tools supporting the authors was the topic of a WGI Training on Data and Software Development in

185   Oberpfaffenhofen, Germany, 6-7 June 2019 (IPCC, 2019). An early draft of the FAIR Guidelines were formally approved by TG-Data on its first meeting in Montreal, Canada, 6-8 November 2019, and the official version 1.0 was adopted by TG-Data in a virtual meeting in 2022.
       The IPCC FAIR Guidelines (Pirani et al., 2022) call for increased attention to~~can be described along~~ three aspects:
       •   Traceability of key statements and of figure and table creation: Information on input datasets like CMIP6 (Eyring et

190   al., 2016), final data displayed in figure~~graphie~~s, and ~~the~~ analysis scripts generating the figures are collected from the authors

by the WGI AR6 TSU. This information is ~~provided~~ recorded for every figure ~~in~~ as part ~~special section~~ of the Supplementary Materials ~~for each~~associated with each chapter. Moreover, bidirectional references between the digital AR6, final datasets and input datasets will enable users to navigate between these AR6 products~~interlink the different AR6 products~~ (Fig. 5~~5~~).

195 •   Credit for input data: Input datasets used by the authors are cited in the AR6 in compliance with Good Scientific Practices (DFG, 2019). In case of CMIP6 data, data citation is required by the Creative Commons licenses (CC, https://creativecommons.org/, last access: 10 March 2022), ~~CC BY SA and CC BY NC SA,~~ under which CMIP6 data were published. CMIP6 data are cited in a summarized form in Appendix II of the WGI AR6 (https://www.ipcc.ch/report/sixth-assessment-report-working-group-i/, last access: 10 March 2022; IPCC, 2021), the provenance metadata of the IPCC WGI Interactive Atlas (https://interactive-atlas.ipcc.ch/, last access: 10 March 2022), and for each figure in the Supplementary 200 Materials.

Figure 5~~5~~: Schematic vision of the bi-directional references between report, input data and final datasets in IPCC AR6 enabling users to navigate among these AR6 results (screenshots from IPCC, CEDA and DKRZ webpages).

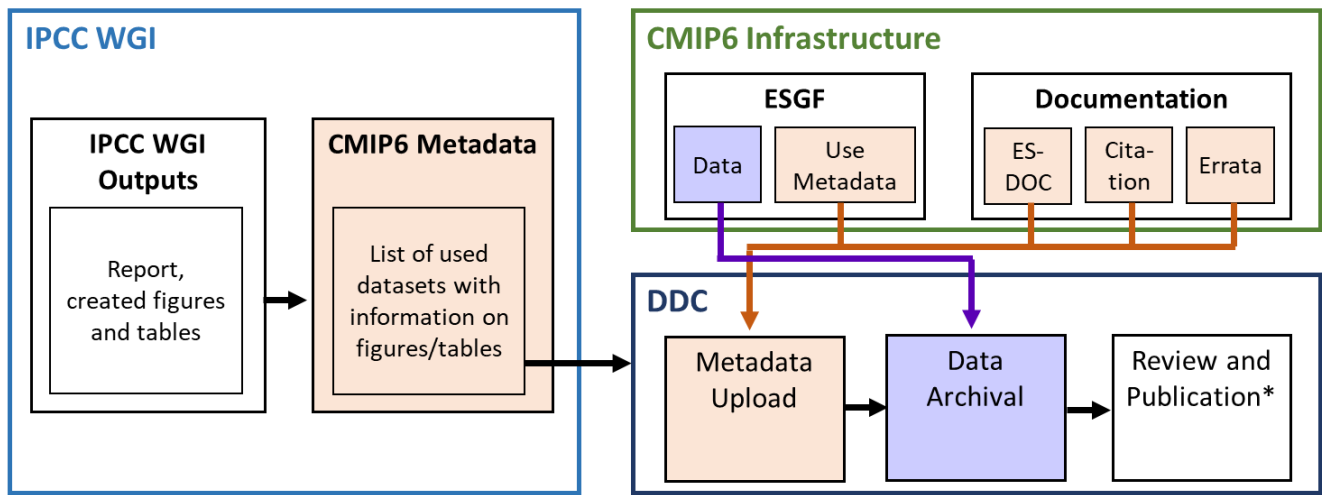
205 •   Long-term preservation of input data, scripts, and final data: The information, scripts and final datasets collected by the WGI TSU are transferred to the designated repository for long-term preservation. DOI registration makes the data and scripts citable and enables data users to give credit to chapter scientists for them. In case of CMIP6, the TSU compiled dataset lists for the DDC Partner DKRZ based on the data usage information collected from the authors~~dataset lists are compiled by the TSU a provided.~~ ~~Based on these lists~~For data long-term archival, the listed CMIP6 datasets are replicated, use metadata 210 are accessed from the ESGF, and further documentations from the Citation Service (Stockhause and Lautenschlager, 2017) and if available from ES-DOC (Pascoe et al., 2020)~~ are added~~. The long-term archival workflow is depicted in Figure 6~~6~~. The implementation of the IPCC FAIR Guidelines required a close cooperation between WGI TSU and the DDC Partners and relied on the CMIP6 infrastructure partners and information provided by the CMIP6 participants.~~.~~

215

Figure 6~~6~~: CMIP6 input data archival workflow to build the DDC AR6 Reference Data Archive

4   Changed role of the DDC Partner DKRZ in AR6

The implementation of the FAIR Guidelines expanded~~changed~~ the role of the DDC Partner DKRZ from a responsibility limited to a long-term data archive, operating mostly independent of WGI and the assessment cycle, ~~in~~to a~~n~~ more active partner with

220   an enhanced role within the Sixth~~s~~ Assessment cycle. ~~The c~~Closes~~t~~ cooperation was required with the WGI TSU to~~on the~~ formulat~~cion~~ and implement~~ation of~~ the FAIR Guidelines. Thus, DDC Managers participated in the IPCC Expert Meeting on Assessing Climate Information for Regions in May 2018 and jointly organized the WGI Training on Data and Software Development in June 2019 together with the WGI TSU. Advice based on the DDC's long experience in data management ~~we~~was~~re~~ provided for gathering the necessary information on data usage required ~~of~~from the authors, best practices in data

225   citation and the definition of machine-actionable interfaces. The DDC Manager at DKRZ joined the WGI AR6 authors as contributing author and reviewed the First and Second Order Drafts of the report to provide expert advice on data management aspects.

This active role of the DDC in AR6 increased the DDC's visibility and resulted in requests for further support of the IPCC author teams during the preparation of the AR6. The DDC Partner DKRZ and former DDC Partner CEDA provided Virtual

230   Workspaces (Stockhause, 2020; Fig. 8~~7~~) for the authors co-funded by the EU project Infrastructure for the European Network for the Earth System Modelling (IS-ENES, http://is.enes.org, last access: 10 March 2022). These collaboration platforms provided storage and compute resources for the chapter author groups together with access to requested core datasets and common software packages. Moreover, DKRZ supported the technical aspects of the ESMValtool (https://www.esmvaltool.org/, last access: 10 March 2022; Eyring et al., 2020) development ~~in technical aspects~~ and hosts the

235   webpage with CMIP evaluation results (https://cmip-esmvaltool.dkrz.de/, last access: 10 March 2022). On the national level, the DDC Manager at DKRZ joined the authors' subgroup of the German IPCC Coordination Office (https://www.de-ipcc.de/, last access: 10 March 2022) as German contributor to the IPCC AR6.~~.~~

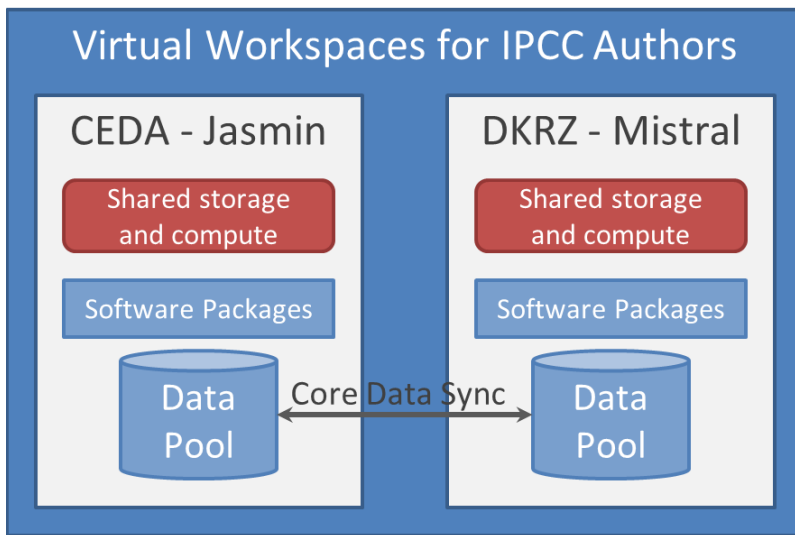



Figure 77: Virtual Workspaces provided by CEDA and DKRZ for IPCC AR6 authors (co-funded by IS-ENES).

DuringThe implementation of the FAIR Guidelines, introduced questions arose that had to be solved with the IPCC Bureau. One of these involved the original licenses modelling groups attached to their CMIP6 data, which were too restrictive for a general reuse of IPCC data products, e.g. final data or Atlas data. like license issues. The IPCC had to ask the CMIP6 participants through the Working Group on Coupled Modelling (WGCM) for an exemption of the CMIP6 data licenses., allowing the DDC Partners to publish the derived final datasets underpinning figures and intermediate and final datasets underpinning the Interactive Atlas under the CC BY license. As representatives of TG-Data, DDC Partner DKRZ and former DDC Partner CEDA were responsible for helping to ensure that IPCC technical requirements were met by the CMIP infrastructure being developed under the coordination ofrepresented the technical requirements of IPCC in the WGCM Infrastructure Panel (WIP, https://www.wcrp-climate.org/wgcm-cmip/wip, last access: 10 March 2022) and contributed data aspects to the IPCC Informal Group on Publications.

Independent of the FAIR Guidelines, the DDC Partners intensified their collaboration. The new UK DDC Partner MetadataWorks set up a joint DDC catalogue to improve the discovery of DDC data holdings. The DKRZ's DDC Manager contributed to the development of the DDC's profile of the Data Catalog Vocabulary standard (W3C DCAT, https://www.w3.org/TR/vocab-dcat-3/, last access: 10 March 2022) and provided the metadata of its Reference Data Archive in December 2021. A central DDC help desk was set up to coordinate the DDC user support. The revision of the DDC webpages is ongoing with the aim of retiringaiming to retire outdated pages and to refocusing the content on IPCC-related data, as called for under according to the renewed DDC Guidance.

5   Position of the DDC within the climate infrastructure and Rrole of CMIP6 for the AR6 cycleinfrastructure and other projects for the AR6 cycle

All of the IPCC Assessments have heavily drawn cycles relied on the latest climate change research on climate change projections provided by the WCRP project CMIP project. The core work of IPCCthe authors' work is the assessment of the latest peer-reviewed literature. CMIP data were used in the peer-reviewed literature and more directlyas well as for the creation of several IPCC report figures. With the introduction of the FAIR Guidelines, the dependency on CMIP-related literature and CMIP data were complemented by the dependency on CMIP6 infrastructure components (Petrie et al., 2021) and further DDC support activities (see section 4). For the CMIP6, data and infrastructure coordination, the WIP was formed by WGCM in 2014 to coordinate the development of the CMIP infrastructure across multiple institutions and agencies.set up, which endorsed concepts for infrastructure contributions and project data standards. The standardization of CMIP6 data is important for the reusability of the data. This includes compliance to the NetCDF/CF standard and specific file name conventions, a uniform

directory structure, and the collection and dissemination of the CMIP6 Controlled Vocabularies (CMIP6-CVs,
270 https://github.com/WCRP-CMIP/CMIP6_CVs, last access: 10 March 2022; Taylor et al., 2018)

The necessary infrastructure components include the data infrastructure Earth System Grid Federation (ESGF; Williams et al., 2016; Cinquini et al., 2014), which disseminates the data and provides use metadata and references to further information like data citation through its index. The CMIP6 Citation Service (http://cmip6cite.wdc-climate.de, last access: 10 March 2022; Stockhause and Lautenschlager, 2017) contributes data references to the long-term archival concept (Stockhause et al., 2015)
275 and thus links the data infrastructure to the long-term data preservation of the CMIP6 data subset in the DDC AR6 Reference Data Archive.

Apart from these necessary infrastructure components, ES-DOC provides detailed information on models, experiments and errata for further metadata enrichment in the DDC Reference Data Archive. ~~While t~~The Citation Service succeeded in having full data citation coverage for all datasets published in the ESGF on the literature and data cut-off date for WGI AR6 on 31
280 January 2021, whereas the coverages of ES-DOC model descriptions and errata information are ~~disappointingly~~ low. That means that the Citation Service and DDC mostly rely on the brief model descriptions in the CMIP6-CV ~~contributed~~ provided by the CMIP6 participants during the registration process~~, which are available in the CMIP6 CVs~~.

The agreed data standards and the high volume of the contributed data require thorough quality checks by the participants to ensure compliance and quality of the CMIP6 data. The conformance with the NetCDF/CF and the additional project metadata
285 rules is automatically checked during ESGF data publication. The DDC's services rely on these CMIP6 infrastructure and standardization achievements.

At the same time, the DDC contributes unique services for the international climate community:

1. The DDC is the only provider with a long-term commitment for data preservation and data services; it ensures that data will remain FAIR over time. Neither the ESGF data nodes nor the Copernicus Climate Data Store (CDS,
290 https://cds.climate.copernicus.eu/, last access: 10 March 2022) have made such a commitment. Their focus lies on serving recent climate data.

2. The DDC preserves the data underpinning key statements of the Assessment Reports and thus the data on which several political decisions are based. The CMIP data subset in the DDC containing the scientific information are essential for tracing back these decisions to the scientific basis.

295 3. As part of the IPCC Assessment process, the DDC reference climate data is quality-assured, enriched with metadata and made citable for its reusability by a wide range of future applications.

4. The DDC supports the IPCC authors and the IPCC TSUs during the Assessment cycles in multiple ways.

DDC's data services need to be integrated not only in the IPCC AR6 products but also in the landscape of climate data infrastructures. Examples are the data provision through well-established domain data catalogs like the ESGF and through
300 cross-domain infrastructures like the DataCite catalog, through European Open Science Cloud (EOSC, https://eosc.eu/, last access: 10 March 2022) services and Nationale Forschungsdateninfrastruktur (NFDI, https://eosc.eu/, last access: 10 March 2022) services. This technical integration requires the exchange of standardized metadata and the implementation of standard interfaces, which are developed by the international organizations including W3C, ISO, Research Data Alliance (RDA), WDS, CODATA, Open Geospatial Consortium (OGC) or the Coalition for Publishing Data in the Earth and Space Sciences
305 (COPDESS). The FAIR Digital Object Framework concept (De Smedt et al., 2019) provides guidance for the future interoperability of data and other digital objects.

However, the most important collaboration partners for the DDC Partner DKRZ are CMIP, the WIP, ESGF and further CMIP infrastructure partners. Through these collaborations, the DDC contributes its experiences to the CMIP future design.

~~For the provision of sustainable DDC services at DKRZ, it is essential to collaborate with a wide range of external partners:~~
310 ~~General Infrastructure Partners:~~
~~o National: Nationale Forschungsdateninfrastruktur (NFDI, https://www.nfdi.de/, last access: 10 March 2022)~~

- European: European Open Science Cloud (EOSC, https://eosc.eu/, last access: 10 March 2022)
  - International: W3C, ISO, Research Data Alliance (RDA), WDS, CODATA, DataCite
- Geophysical Infrastructure Partners:
  - Open Geospatial Consortium (OGC)
  - Coalition for Publishing Data in the Earth and Space Sciences (COPDESS, https://copdess.org, last access: 10 March 2022)
  - NFDI Consortium for Earth System Sciences (NFDI4Earth, https://www.nfdi4earth.de/, last access: 10 March 2022)
- Climate Infrastructure Partners:
  - European Network for Earth System Modelling Research Infrastructure (ENES-RI) Partners
  - ESGF
  - Copernicus Climate Data Store (CDS, https://cds.climate.copernicus.eu/, last access: 10 March 2022)
  - WIP and Climate Data Node Operation Team (CDNOT)

6   Conclusion and Perspectives

The IPCC DDC has provided quality-assured, citable IPCC-relevant reference climate data for all IPCC Assessment Reports and has supported the IPCC Assessments over the 25 years of its existence. The specific The role of the DDC has changed over the 25 years of its existence in order to adjust to evolving data management standards and evolving requirements from IPCC WGsI. Furthermore, the tasks responsibilities of the DDC Partner DKRZ have been adapted to developments in the CMIP6 infrastructure, which provides the data and documentation for the DDC's Reference Data Archive. DDC's data holdings provide valuable ancillary information for IPCC Assessment Reports. AR6 markedstands for a major change: On the one hand, tThe role of the DDC turned from maintaining an independent long-term data archive into an activeproviding general data servicess provider for the IPCC. At the same time, adoption ofOn the other hand, the IPCC FAIR Guidelines significantly enhanced the transparency of AR6 key findings. Their implementation posed a challenge to all partners: WGI TSU, IPCC authors and the DDC Partners. Data usage documentation in AR6 aund the long-term archival of related input and final datasets enable the traceability of results and the reuse of datasets. Long-term preservation of the data in the DDC ensures data availability and traceability on the long-term. Still, the DDC AR6 data archive remains but it is still incomplete esp. in the long-term preservation of input datasets. This was identified by the DDC Partners in 2020 as one of several areas for future improvements identified this as part of their gap analysis in 2020:

1. Exhaustive IPCC data archival,
2. Improved global data access (e.g. compute service for reduction of data transfer volume to support DDC users in developing countries and support for users from various domains),
3. Data Discovery,
4. Machine-accessible DDC data,
5. Regional to local data and data services,
6. Sustaining DDC Partners,
7. Collaboration with data infrastructure networks, e.g. RDA, WDS or CODATA,
8. Collaboration with cognate data providers, e.g. IPBES.

Some of these gaps are easy tohave been filled with the limited DDC resources,target like the collaboration aspects (gaps 7 and 8) or the and the data discovery issue (gap 3) was solved with the establishment of the joint DDC catalogue, but the remaining gaps require funding and are related to the missing long-term strategy for the DDC (gap 6). The current DDC Partner funding is provided by their IPCC member states for each Assessment cycle. In Germany, the DDC was funded as part of

research projects supporting the German contribution to CMIP and enabled the DDC to add the Reference Data Archive for the each AR cycle. Long-term data preservation and maintenance of data services of the DDC as for any other research repository and research data infrastructure including the CMIP infrastructure (ESGF, ES-DOC, Citation Service) still rely on voluntary contributions from institutions and individuals. There are several single points of failure in this practice. National funders of such international services like the DDC expect other nations to share in the costs. Thus, an internationalization of expenses for core international research data services and infrastructures is required.

TG-Data has targeted DDC's sustainability as part of its AR6 review process with the aim to formulate recommendations for the AR7 data management. Three options are under discussion:

1. IPCC members fund DDC Partners for an Assessment cycle,
2. IPCC members contribute to a DDC fund, or
3. IPCC members funding WG TSUs also fund the associated DDC Partner.

The DDC Partner funding for an Assessment cycle (1) is problematic, since data from the departing DDC Partner are to be transferred to the replacing DDC Partner and the experience of the DDC Partner in IPCC processes and procedures is cyclically lost. The data volume of the Reference Data Archives of the DDC Partner DKRZ is high (see Fig. 3) and the transfer is time-consuming and expensive without adding any value. Furthermore, the important collaboration with CMIP6 and the various infrastructure partners must be re-established by the new DDC Partner. Optimistically, this option will cause further significant but avoidable costs and, pessimistically, it is or will become impractical. Both option 2 and 3 can fund the addition of data for a new Assessment cycle to the DDC Reference Data Archive, but only option 2 can additionally support the long-term aspects of data preservation and the provision of customized data services. Option 3 could pose a problem for the IPCC as it further increases the already high costs associated with a TSU and might discourage IPCC members from nominating a co-chair for a WG. On the other hand, option 2 is subject of extensive discussions within the IPCC and therefore may not yet be available for AR7.

The value of data as scientific asset and the importance of open data has become undisputed. In "The Beijing Declaration on Research Data" (2019), CODATA emphasizes the importance of the broad reuse of data to address global challenges and recognizes the enormous challenge in the interoperability of data and responsible stewardship. The UNESCO states in its 'Recommendations on Open Science' (2021) that non-commercial infrastructures should facilitate ensuring the long-term preservation, stewardship and community control of research products including data. It recommends supporting these open infrastructures by direct funding and through an earmarked percentage of each funded grant. Regarding the IPCC DDC, the German Minister of Education and Research, Mrs. Stark-Watzinger, as representative of the German government explicitly mentioned Germany's involvement in the IPCC DDC and the importance of data for the IPCC process at the opening of the 55th Session of the IPCC and 12th Session of WGII on 10 February 2022 (IPCC, 2022).

Against the background of these recent developments in open science, the role of research infrastructures and the importance of long-term preservation of data as a valuable scientific asset, a solution for the sustainable funding of the DDC seems within reach. The first review of the FAIR Guidelines by TG-Data has started, in order to derive recommendations for the AR7 data management and to perpetuate the introduced FAIR data concept. Aspects under discussion include the review of the content of the FAIR Guidelines and their implementations, a better integration of the FAIR Guidelines into the IPCC assessment process and the formulation of a long-term strategy. For operationalizing the FAIR Guidelines' implementation, the integration of recommendations and standards from various infrastructure expert groups is required in order to unite best practices in data management with Good Scientific Practice. Thus, the DDC can be positioned in the developing FAIR digital object research landscape as described in the concept of a FAIR Digital Object Framework by De Smedt et al. (2020).

The IPCC FAIR Guidelines' implementation in the AR6 and in the IPCC WGI Interactive Atlas or even the input4MIPs project publication (Durack et al., 2018) contribute practice examples to the discussions on the integration of data into publications

Appendix A – Core Variables of the Reference Data Archive

air_pressure_at_sea_level
air_temperature
convective_precipitation_flux
dew_point_temperature
geopotential_height
global_average_thermosteric_sea_level_change
land_area_fraction
land_ice_area_fraction
large_scale_precipitation_flux
precipitation_flux
relative_humidity
sea_ice_amount
snowfall_amount
soil_moisture_content
specific_humidity
surface_altitude
surface_net_downward_shortwave_flux
surface_sensible_heat_flux
surface_downwelling_shortwave_flux_in_air
surface_snow_area_fraction
surface_snow_melt_flux
surface_snow_thickness
surface_temperature
water_evaporation_rate
wind_speed
x_wind
y_wind

Table A1: Core variables of the Reference Data Archive in CF Standard Name convention

400 **Code/Data availability**

No code nor data was created for this study.

Author contribution

MS wrote the manuscript draft and ML reviewed and contributed to the manuscript.

Competing interests

405 The authors declare that they have no conflict of interest.

Acknowledgements

We thank Tim Carter (former co-chair of IPCC TGICA, Finnish Environment Institute SYKE), who generously answered questions and shared unpublished materials about the early days of the IPCC DDC and TGCIA; Anna Pirani (head of IPCC WGI TSU) for the coordination of the formulation and FAIR Guidelines and their implementation in the AR6 WGI; and the DDC Partners, the WG TSUs esp. the WGI TSU colleagues, and the TG-Data members for their contributions to the IPCC FAIR Guidelines. The authors are grateful for the standardization efforts of various groups including PCMDI and the WIP members. Special thanks to the reviewers Karl Taylor and Paul Durack for their rich comments and valuable suggestions for the improvement of the manuscript.

References

Cinquini, Luca, Crichton, Daniel, Mattmann, Chris, Harney, John, Shipman, Galen, Wang, Feiyi, Ananthakrishnan, Rachana, Miller, Neill, Denvil, Sebastian, Morgan, Mark, Pobre, Zed, Bell, Gavin M., Doutriaux, Charles, Drach, Robert, Williams, Dean, Kershaw, Philip, Pascoe, Stephen, Gonzalez, Estanislao, Fiore, Sandro, and Schweitzer, Roland: The Earth System Grid Federation: An open infrastructure for access to distributed geospatial data, Future Generation Computer Systems, Volume 36, 2014, Pages 400-417, ISSN 0167-739X, https://doi.org/10.1016/j.future.2013.07.002, 2014.

CODATA, Committee on Data of the International Science Council, CODATA International Data Policy Committee, CODATA, and CODATA China High-level International Meeting on Open Research Data Policy and Practice, Hodson, S., Mons, B., Uhlir, P., and Zhang, L.: The Beijing Declaration on Research Data, Zenodo, https://doi.org/10.5281/zenodo.3552330, 2019.

De Smedt, Koenraad, Koureas, Dimitris, and Wittenburg, Peter: FAIR Digital Objects for Science: From Data Pieces to Actionable Knowledge Units. Publications 2020, 8(2), 21, MPDPI, https://doi.org/10.3390/publications8020021, 2021.

Deutsche Forschungsgemeinschaft: Guidelines for Safeguarding Good Research Practice, Code of Conduct, https://doi.org/10.5281/zenodo.3923602, 2019.

Durack, P. J.,Taylor, K. E.,Eyring, V.,Ames, S. K.,Hoang, T.,Nadeau, D.,Doutriaux, C.,Stockhause, M., and Gleckler, P. J.: Toward standardized data sets for climate model experimentation, Eos, 99, https://doi.org/10.1029/2018EO101751, 2018.

Eyring, V., Bock, L., Lauer, A., Righi, M., Schlund, M., Andela, B., Arnone, E., Bellprat, O., Brötz, B., Caron, L.-P., Carvalhais, N., Cionni, I., Cortesi, N., Crezee, B., Davin, E., Davini, P., Debeire, K., de Mora, L., Deser, C., Docquier, D., Earnshaw, P., Ehbrecht, C., Gier, B. K., Gonzalez-Reviriego, N., Goodman, P., Hagemann, S., Hardiman, S., Hassler, B., Hunter, A., Kadow, C., Kindermann, S., Koirala, S., Koldunov, N., Lejeune, Q., Lembo, V., Lovato, T., Lucarini, V., Massonnet, F., Müller, B., Pandde, A., Pérez-Zanón, N., Phillips, A., Predoi, V., Russell, J., Sellar, A., Serva, F., Stacke, T., Swaminathan, R., Torralba, V., Vegas-Regidor, J., von Hardenberg, J., Weigel, K., and Zimmermann, K.: Earth System Model Evaluation Tool (ESMValTool) v2.0 - an extended set of large-scale diagnostics for quasi-operational and comprehensive evaluation of Earth system models in CMIP, Geosci. Model Dev., 13, 3383-3438, https://doi.org/10.5194/gmd-13-3383-2020, 2020.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

IPCC: Climate Change 1995: The Scientific of Climate Change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change [Houghton, J.T., L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg, and K. Maskell (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1995.

IPCC: Report of the thirteenth session of the IPCC, Maldives, 22, 25-28 September 1997. https://www.ipcc.ch/site/assets/uploads/2018/05/thirteenth-session-report.pdf, last access: 10 March 2022, 1997.

IPCC: Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Houghton, J.T.,Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell, and C.A. Johnson (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 881pp, 2001.

IPCC: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp, 2007.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp., 2013

IPCC: 41st Session of the IPCC, 24-27 February 2015, Nairobi, Kenya, Decisions adopted by the Panel, https://www.ipcc.ch/site/assets/uploads/2018/05/p41_decisions.pdf , last access: 10 March 2022, 2015.

IPCC: Expert Meeting on the Future of the Task Group on Data and Scenario Support for Impact and Climate Analysis (TGICA), 26-27 January 2016, Geneva, https://www.ipcc.ch/site/assets/uploads/2018/05/EMR_TGICA_Future.pdf, last access: 10 March 2022, 2016.

IPCC: Guidance for the core functions of the IPCC Data Distribution Centre (DDC). https://www.ipcc.ch/site/assets/uploads/2018/12/Guidance_DDC.pdf, last access: 10 March 2022, 2018a.

IPCC: Report of the forty-seventh session of the IPCC, Paris, France, 13–16 March 2018, https://www.ipcc.ch/site/assets/uploads/2018/03/final_report_p47.pdf, last access: 10 March 2022, 2018b.

IPCC: Summary for Policymakers. In: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty [Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield (eds.)]. Cambridge University Press, Cambridge, UK and New York, NY, USA, pp. 3-24, https://doi.org/10.1017/9781009157940.001, 2018c.

IPCC: Expert Meeting on Assessing Climate Information for Regions, Trieste, Italy, 16-18 May 2018, http://www.ipcc.ch/apps/eventmanager/documents/52/120920180323-INF5EM-Regions.pdf, last access: 10 March 2022, 2018de.

IPCC: Working Group I Training on Data Access and Software Development, Oberpfaffenhofen, Germany, 6-7 June 2019, https://www.ipcc.ch/event/wgi-training-on-data-and-software-development/, last access: 10 March 2022, 2019.

IPCC: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press. In Press, https://doi.org/10.1017/9781009157896.

IPCC: 55th Session of the IPCC, Opening Ceremony, 14th February 2022, https://www.youtube.com/watch?v=00_0udBrw_w, last access: 10 March 2022, 2022.

Lin, Dawei, Crabtree, Jonathan, Dillo, Ingrid, Downs, Robert R., Edmunds, Rorie, Giaretta, David, De Giusti, Marisa, L'Hours, Hervé, Hugo, Wim, Jenkyns, Reyna, Khodiyar, Varsha, Martone, Maryann E., Mokrane, Mustapha, Navale, Vivek, Petters,

Jonathan, Sierman, Barbara, Sokolova, Dina V., Stockhause, Martina, and Westbrook, John: The TRUST Principles for digital repositories, Sci Data 7, 144. https://doi.org/10.1038/s41597-020-0486-7, 2020.

490 Pascoe, C., Lawrence, B. N., Guilyardi, E., Juckes, M., and Taylor, K. E.: Documenting numerical experiments in support of the Coupled Model Intercomparison Project Phase 6 (CMIP6), Geosci. Model Dev., 13, 2149–2167, https://doi.org/10.5194/gmd-13-2149-2020, 2020.

Petrie, R., Denvil, S., Ames, S., Levavasseur, G., Fiore, S., Allen, C., Antonio, F., Berger, K., Bretonnière, P.-A., Cinquini, L., Dart, E., Dwarakanath, P., Druken, K., Evans, B., Franchistéguy, L., Gardoll, S., Gerbier, E., Greenslade, M., Hassell, D.,

495 Iwi, A., Juckes, M., Kindermann, S., Lacinski, L., Mirto, M., Nasser, A. B., Nassisi, P., Nienhouse, E., Nikonov, S., Nuzzo, A., Richards, C., Ridzwan, S., Rixen, M., Serradell, K., Snow, K., Stephens, A., Stockhause, M., Vahlenkamp, H., and Wagner, R.: Coordinating an operational data distribution network for CMIP6 data, Geosci. Model Dev., 14, 629–644, https://doi.org/10.5194/gmd-14-629-2021, 2021.

Pirani, Anna, Alegria, Andres, Al Khourdajie, Alaa, Gunawan, Wuwan, Gutiérrez, José Manuel, Holsman, Kirstin, Huard,

500 David, Juckes, Martin Kawamiya, Michio, Klutse, Nana, Krey, Volker, Matthews, Robin, Milward, Adam, Pascoe, Charlotte, van der Shrier, Gerard, Spinuso, Alessandro, Stockhause, Martina, and Xing, Xiaoshi: The implementation of FAIR data principles in the IPCC AR6 assessment process, Zenodo, https://doi.org/10.5281/zenodo.6504469, 2022.

Stockhause, M.: IPCC Virtual Workspace at DKRZ, http://bit.ly/IPCC_DKRZ_Virtual_Workspace, last access: 10 March 2022, 2020.

505 Stockhause, M., Höck, H., Toussaint, F., and Lautenschlager, M.: Quality assessment concept of the World Data Center for Climate and its application to CMIP5 data, Geosci. Model Dev., 5, 1023-1032, 2012, https://10.5194/gmd-5-1023-2012, 2012.

Stockhause, M., Juckes, M., Chen, R., Moufouma Okia, W., Pirani, A., Waterfield, T., Xing, X. and Edmunds, R.: Data Distribution Centre Support for the IPCC Sixth Assessment, Data Science Journal, 18(1), p.20, http://doi.org/10.5334/dsj-2019-020, 2019.

510 Stockhause, Martina, Juckes, Martin, Pirani, Anna, Poloczanska, Elvira, and Waterfield, Tim: First IPCC AR6 Data Workshop Minutes, Zenodo, https://doi.org/10.5281/zenodo.1036460, 2017.

Stockhause, M. and Lautenschlager, M.: CMIP6 Data Citation of Evolving Data. Data Science Journal, 16, p.30, http://doi.org/10.5334/dsj-2017-030, 2017.

Stockhause, M., Toussaint, F., and Lautenschlager, M.: CMIP6 Data Citation and Long-Term Archival, Zenodo,

515 https://doi.org/10.5281/zenodo.35178, 2015.

Taylor, Karl E., Juckes, Martin, Balaji, V., Cinquini, Luca, Denvil, Sébastien, Durack, Paul J., Elkington, Mark, Guilyardi, Eric, Kharin, Slava, Lautenschlager, Michael, Lawrence, Bryan, Nadeau, Denis, and Stockhause, Martina: CMIP6 Global Attributes, DRS, Filenames, Directory Structure, and CV's, 10 September 2018 (v6.2.7), https://goo.gl/v1drZl, last access: 10 March 2022, 2018.

520 Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, Bulletin of the American Meteorological Society, 93(4), 485-498, https://doi.org/10.1175/BAMS-D-11-00094.1, 2012.

Vaughan, Catherine: An institutional analysis of the IPCC Task Group on Data and Scenario Support for Impacts and Climate Analysis (TGICA), https://www.researchgate.net/publication/301602432_An_institutional_analysis_of_the_IPCC_Task_Group_on_Data_and_S

525 cenario_Support_for_Impacts_and_Climate_Analysis_TGICA, last access: 10 March 2022, 2016.

World Data System: Data Repositories Day 2018, Gaborone, Botswana, 9 November 2018. https://www.worlddatasystem.org/community/wds-members-forum/data-repositories-day-2018, last access: 10 March 2022, 2018.

UNESCO: UNESCO Recommendations on Open Science, SC-PCB-SPP/2021/OS/UROS, https://en.unesco.org/science-

530 sustainable-future/open-science/recommendation, last access: 10 March 2022, 2021.

Wilkinson, M., Dumontier, M., Aalbersberg, I., Appleton, Gabrielle, Axton, Myles, Baak, Arie, Blomberg, Niklas, Boiten, Jan-Willem, da Silva Santos, Luiz Bonino, Bourne, Philip E., Bouwman, Jildau, Brookes, Anthony J., Clark, Tim, Crosas, Mercè, Dillo, Ingrid, Dumon, Olivier, Edmunds, Scott, Evelo, Chris T., Finkers, Richard, Gonzalez-Beltran, Alejandra, Gray, Alasdair J.G., Groth, Paul, Goble, Carole, Grethe, Jeffrey S., Heringa, Jaap, 't Hoen, Peter A.C, Hooft, Rob, Kuhn, Tobias,

535   Kok, Ruben., Kok, Joost, Lusher, Scott J., Martone, Maryann E., Mons, Albert, Packer, Abel L., Persson, Bengt, Rocca-Serra, Philippe, Roos, Marco, van Schaik, Rene, Sansone, Susanna-Assunta, Schultes, Erik, Sengstag, Thierry, Slater, Ted, Strawn, George, Swertz, Morris A., Thompson, Mark, van der Lei, Johan, van Mulligen, Erik, Velterop, Jan, Waagmeester, Andra, Wittenburg, Peter, Wolstencroft, Katherine, Zhao, Jun, and Mons, Barend: The FAIR Guiding Principles for scientific data management and stewardship, Sci Data 3, 160018. https://doi.org/10.1038/sdata.2016.18, 2016.

540   Williams, D. N., Balaji, V., Cinquini, L., Denvil, S., Duffy, D., Evans, B., Ferraro, R., Hansen, R., Lautenschlager, M., and Trenham, C.: A Global Repository for Planet-Sized Experiments and Observations, Bulletin of the American Meteorological Society, 97(5), 803-816, https://doi.org/10.1175/BAMS-D-15-00132.1, 2016.

Xing, Xiaoshi, Stockhause, Martina, Gutiérrez Llorente, José Manuel, and Irwin, Courtney: Memorandum of Understanding (MoU) for Operation of the IPCC Data Distribution Centre, Zenodo, https://doi.org/10.5281/zenodo.5914482, 2021.