





## **Disclaimer**

The right of publication in print, electronic and any other form and in any language is reserved by the World Climate Research Programme (WCRP). Short extracts from WCRP publications may be reproduced without authorization provided that the complete source is clearly indicated. Editorial correspondence and requests to publish, reproduce or translate this publication, either as a whole or in part, should be addressed to: [wcrp@wmo.int](mailto:wcrp@wmo.int).

The designations employed and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariats of WCRP Sponsor Organizations – the World Meteorological Organization (WMO), the Intergovernmental Oceanographic Commission (IOC) of UNESCO and the International Council for Science (ICSU) – concerning the legal status of any country or territory, or its authorities, or concerning the delimitation of the frontiers of any country or territory.

Recommendations of working groups and panels shall have no status within WCRP and the Sponsor Organizations until they have been approved by the Joint Scientific Committee (JSC) of WCRP. The recommendations must be concurred with by the Chair of the JSC before being submitted to the designated constituent body or bodies.

### **Executive Summary**

*WCRP encourages its projects and community to foster free and open access to science data sets and databases. Data should be published freely, openly and with meta-data that make the data discoverable, intelligible, assessable and usable. Data publishing should make use of established repositories, infrastructure and international standards. WCRP data policy emphasizes the importance of data citation, long-time preservation and access to published research findings based on these data. To foster and ensure interoperability of datasets, the international community is encouraged to converge on and promote a small set of copyright licenses, which constitute the concrete manifestations of a given data policy. In the long term, open data publishing and curation should aim for full technological and cultural integration into the research and knowledge production cycle.*

## **Contents**

<b>1. INTRODUCTION</b>	<b>1</b>
<b>2. WCRP DATA POLICY: PROMOTING AND ENCOURAGING OPEN ACCESS TO SCIENTIFIC DATA</b>	<b>1</b>
<b>3. OVERVIEW: DATA POLICY DOCUMENTS RELEVANT TO THE WCRP COMMUNITY</b>	<b>2</b>
<b>4. UNIFYING AND HARMONIZING OPEN-ACCESS DATA POLICIES AND DATA SET COPYRIGHT TERMS</b>	<b>3</b>
<b>5. FUTURE DEVELOPMENTS</b>	<b>4</b>
<b>6. REFERENCES</b>	<b>5</b>



## 1. Introduction

The World Climate Research Programme (WCRP) advances its mission – facilitating the analysis of Earth’s climate – through a wide array of coordinative instruments. The majority of these are inextricably linked to aspects of scientific data. Research coordination activities by WCRP, realized together with its extensive range of partners and collaborations, concern the full spectrum of the data value chain, in particular: data collection or production, data storage, data stewardship, data quality control, data rescue, data dissemination, data analysis, data fusion, data ingestion, data synthesis, reanalysis and data-model comparison; as well as infrastructure for data upload, storage, processing and distribution. WCRP activities continuously rely on the commitment to free and open exchange of data that the international climate science community has established to date. In support of its mission and of climate research, as well as for the benefit of society, WCRP promotes an open-access approach to science data policy.

## 2. WCRP Data Policy: promoting and encouraging open access to scientific data

WCRP activities are co-funded and co-supported by a broad community of agencies, universities, researchers, national as well as international bodies and co-operations. Ownership of data and data publishing rights hence vary widely across WCRP-sponsored, WCRP-endorsed and WCRP-related projects and will often lie with entities outside WCRP. In recognition of these constellations, the WCRP data policy is three-fold:

- 1) Where formal ownership of research data lies with WCRP, its Core Projects, Grand Challenges, Working Groups, Advisory Councils, their sub-projects or with any other WCRP bodies, these data should be made freely and openly accessible, through adequate mechanisms and to the maximum extent possible. In some cases, WCRP projects may hold copyright on databases they assembled even when the copyright for all individual data sets in the collection lies with external entities [28]. In these cases, it is likewise encouraged to make available the database itself under an appropriate open-access license.
- 2) For WCRP-related projects for which formal ownership of research data lies not with these projects themselves, but with another collaborating institution, WCRP encourages that these data be made freely and openly accessible by the owner, to the maximum extent possible;
- 3) By promulgating the WCRP Data Policy to its community, as well as through WCRP’s dialog at international level, with funding agencies, learned societies, research institutions, agencies, ministries, expert panels and other groups involved in science data policy planning, WCRP promotes the development, installment and harmonization of, as well as advocacy for, interoperable open-data policies in climate research and Earth system science, to the maximum extent possible, for the benefit of science and society.

In all of the above, free and open access to data is understood to imply (see, e.g., [4,6,11,13,15, 16,17,18,21,22,26]) that the data is:

- a) available free of cost if the data is obtained via internet download, and never incurring any charge for the data itself beyond cost of dissemination for any other means of data provision; available with minimum delay for validation or other necessary processing (possibly allowing for a timely release of preliminary data in cases where extensive post-processing is involved); available under minimum restrictions on use (including on searching, mining, and linking), re-use, re-publication and re-distribution, and in a non-discriminatory fashion;
- b) easily technically accessible, made available according to any relevant applicable international standards and not requiring user registration except where unavoidable;
- c) published via citable, persistent and traceable means, for example via dataset DOIs and, where appropriate and possible, described in the open-access data literature and with appropriate means to account for evolving data (e.g., [22,23,27]). Where possible, peer-reviewed scientific studies that used, produced or evaluated such a data set or are otherwise strongly related to it should be published in open-access mode as well (see e.g. [12,13,24]);
- d) complete, quality-controlled, with mechanisms for error reporting and correction in place, and published without unnecessary alterations from the version issued by the primary publisher. Data distributed in real or near-real time may, as appropriate, be re-issued in a time-delayed mode after additional quality control and documentation;
- e) published and equipped with meta-data in ways that make the data discoverable, intelligible, assessable and usable (“intelligently open”, see e.g. [4,21]). This also includes reporting or registering the data and meta-data with appropriate registries and meta-databases to enhance discoverability. Further, data should be published under uniform data formats and quality standards, with uniform meta-data, and with a particular view on facilitating their use in data assimilation, model intercomparison, data-model comparison and reanalysis;
- f) published and stored in ways that ensure long-term preservation, integrity and access; and that minimize technological redundancy by utilizing established standards and infrastructure, including national and international repositories (for example, the Earth System Grid Federation ESGF utilized by e.g. CMIP [25], the Carbon Dioxide Information Analysis Center CDIAC, the PANGAEA library, or the general-purpose science data archive Dryad).

WCRP furthermore calls upon data users to ensure attribution, citation and recognition of the origin and producers of (climate) research data (e.g., [22,23,27]).

WCRP data policy is intended to be compatible with those of its co-sponsors (WMO [1], IOC [2,3] and ICSU [4,5]) and collaboration partners (e.g., GCOS [9] and GEOSS [10]).

### **3. Overview: Data policy documents relevant to the WCRP community**

A wide range of data policy documents have been promoted by the international community within the areas of both climate research in particular as well as general science and data policy.

Data policy documents put forth by the three WCRP co-sponsors are:

- WMO resolutions 40, 25 and 60 as well as the Geneva Declaration adopted by the 12<sup>th</sup>, 13<sup>th</sup>, 17<sup>th</sup> and 13<sup>th</sup> World Meteorological Congress, respectively [1];

- the IOC Oceanographic Data Exchange Policy [2] as well as the 2013-2016 IOC Strategic Plan for Oceanographic Data and Information Management [3];
- the 2015 Science International accord on “Open Data in a Big Data World” by the International Council for Science (ICSU), the InterAcademy Partnership (IAP), The World Academy of Sciences (TWAS) and the International Social Science Council (ISSC) [4] as well as the ICSU World Data System (WDS) Data Sharing Principles [5].

Data policy documents adopted by WCRP projects are:

- the CLIVAR Data Policy [6];
- the CMIP6 Data Licensing terms [7,8] and Data Citation Regulations [23];

–

Data policy documents by partners and collaborators of WCRP are, among others:

- the GCOS implementation plan [9];
- the GEOSS Data Sharing Principles [10];

–

Data policy documents put forth for international science and research in a broader scope are, e.g.:

- the OECD Principles and Guidelines for Access to Research Data from Public Funding [11];
- the EU Horizon 2020 Open Research Data Pilot under Article 29.3 of the Horizon 2020 Model Grant Agreement [12], the EU RECODE project consortium report [13] and the EU INSPIRE Directive [14];
- the report by the United Nations Secretary-General’s Independent Expert Advisory Group on the Data Revolution for Sustainable Development [15].

## 4. Unifying and harmonizing open-access data policies and data set copyright terms

Science-coordinating bodies as well as research institutions and data initiatives across the world have, in absence of a universal international standard for data sharing in (climate) research, individually advanced data policies in support of open access to data (see previous section). While each speaking to different stakeholders, they together provide a broad basis of support for open data principles. At the same time, diversity in data policies can hinder interoperability when fusing, combining or re-publishing data under slightly differing policies – and where different data policies are fully compatible, opportunities may exist to unify policy across issuing entities [4,13,16,17,18].

Previous surveys and meta-analyses have provided intercomparisons of the existing spectrum of data policies in the (climate) science data community (e.g., [16,18]). Almost all policies have in common a call for open access to data to the maximum extent possible while also acknowledging the need for a limited range of exceptions to accommodate privacy, security, safety, ethical or commercial concerns. This common principle is sometimes described as “open data by default”, “qualified openness” or “open data with limits”. Policies then differ in the way allowance for exceptions to universal open access is incorporated. The most common approaches introduce such a trade-off by one or more of the following distinctions between:

- a set of essential data variables to be shared without restrictions vs. non-essential data variables on which restrictions, e.g. for commercial use, may be placed (e.g., [1])
- open-access data available freely for non-commercial use vs. data subject to restrictions for commercial use (e.g., [1,7,11])
- open access for research purposes vs. limited access for all other purposes (e.g., [1,10])
- privileged period of exclusive use for the data producers (data exploitation phase, or “embargo” phase) for a limited time vs. immediate open access [13,16]
- open access to all data by default vs. restrictions on datasets on a case-by-case basis or for data collected by sub-networks operating under different policies (e.g., [10,11,19]).

While such data policies determine commitments within an organizational entity, it is the data set copyright licenses under which data sets are finally published that ultimately determine the usability of and (legal) interoperability between open-access data sets. Entities increasingly include with their data sharing policies also copyright licenses under which data are to be published to be in compliance (see, e.g., [7,16,17,18,20]). WCRP encourages the global community to converge on a limited number of data set copyright licenses in order to harmonize the range of data policies in effect, to enable comparison of the actual consequences which a certain data policy implies and to facilitate the (legal) interoperability and compatibility of published datasets.

Suitable candidates for such licenses might be versions 4.0 (or the otherwise latest version) of the Creative Commons licenses CC-BY (“attribution required”) or CC0 (“public domain”) for both single data sets and entire databases (also see [28]).

## 5. Future developments

**Maintenance of data center infrastructure:** One common provision in data sharing policies states that data are to be provided “free of charge or allowing for a maximum charge of the cost of reproduction, without charge for the data and products”. Such policy can be useful when data are shared via shipping of hardware media (especially in the case of limited bandwidth or high volumes of data preventing internet transfer), through closed networks for operational data, by satellite or otherwise technically intensive means. For many scenarios in climate research, data is today shared by centralized or distributed internet download. In such cases, any “cost of distribution” (or also “marginal cost of distribution”) cannot easily be separated from the cost of build-up, software programming, storage hardware and operating costs of the high-volume, high-throughput data centers serving the data. WCRP argues that policy for data set provision which is primarily effected through the internet should follow recent developments that promote the fully cost-free access to data (e.g., [10,12,16]). Funding agencies as well as national and international research programs must, through appropriate means, take into account the necessity of continuing, long-term support for data storage, stewardship, and digitalization, as well as for data rescue of endangered archives.

**Integrating data management and stewardship into the research process:** As noted in [29], the international community has over recent years moved from discussing whether researchers should publish their data to discussing how the community can best organize to publish research data in a consistent, reliable manner. Increasingly, the latter entails identifying

pathways for integration of data publication management into the research cycle. Tasks here include, for example, data version control [27], allowing bidirectional links between data and publications as well as supporting the publication of software needed to derive published results from research data [29]. More abstract considerations examine the possibility of mutual interfaces, process models, protocols and infrastructures for all aspects of data and the research value chain [29,30,31]. WCRP supports the international community in its ongoing challenge to provide open data already today while over time converging to seamlessly integrated designs that technically and habitually incorporate the open publication of data, software and findings in the research and knowledge creation process.

**Open access to data and the findings derived from them:** Open access to data constitutes only one half of the resources required for traceability, reproducibility and comparability of scientific findings. Significant information is held in the continually produced works of scientific literature that describe the generation, collection, use, quality control, comparison, assessment of, and knowledge extraction from research data. Researchers in developing countries in particular are limited by access restrictions on scientific findings that complement raw research data. WCRP hence advocates for a two-fold approach to making (climate) science data support the traceability, reproducibility and comparability of scientific findings. In addition to promoting open data access, WCRP encourages the global (climate) science community and the entities determining the policies under which it operates to strive for an expansion of policy that equally promotes open access to peer-reviewed scientific publications (see, e.g., [12,13,24]).

## 6. References

- [1] WMO resolutions 40, 25 and 60 as well as the Geneva Declaration adopted by the 12th, 13th, 17th and 13th World Meteorological Congress, respectively. Available from <http://public.wmo.int/en/our-mandate/what-we-do/data-exchange-and-technology-transfer> .
- [2] IOC Oceanographic Data Exchange Policy, adopted by the 22<sup>nd</sup> session of the IOC Assembly ass Resolution IOC-XXII-6. Available from [http://www.iode.org/index.php?option=com\\_content&view=article&id=51&Itemid=95](http://www.iode.org/index.php?option=com_content&view=article&id=51&Itemid=95) .
- [3] IOC Strategic Plan for Oceanographic Data and Information Management (2013-2016). IOC Manuals and Guides, 66. Available from <http://unesdoc.unesco.org/images/0022/002230/223030M.pdf> .
- [4] 2015 Science International accord on “Open Data in a Big Data World”, by the International Council for Science (ICSU), the InterAcademy Partnership (IAP), The World Academy of Sciences (TWAS) and the International Social Science Council (ISSC). Available from <http://www.icsu.org/science-international/accord>
- [5] ICSU WDS Data Sharing Principles, available from [https://www.icsu-wds.org/files/WDS\\_Data\\_Sharing\\_Principles\\_2015.pdf](https://www.icsu-wds.org/files/WDS_Data_Sharing_Principles_2015.pdf) .
- [6] CLIVAR Data Policy. Available from <http://www.clivar.org/resources/data/data-policy> .
- [7] Lawrence, B.N. and the WCRP WGCM Infrastructure Panel: CMIP6 Licensing and Access Control. V1.0, available from [https://www.earthsystemcog.org/projects/wip/position\\_papers](https://www.earthsystemcog.org/projects/wip/position_papers) .
- [8] Eyring, V., Bony, S., Meehl, G., Senior, C., Stevens, B., Stouffer, R., and Taylor, K.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937-1958, doi:10.5194/gmd-9-1937-2016, 2016.

- [9] 2010 Update of the Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC, by WMO, IOC, UNEP and ICSU. Available from <http://www.wmo.int/pages/prog/gcos/Publications/gcos-138.pdf> .
- [10] GEOSS Data Sharing Principles. Available from <http://www.earthobservations.org/dswg.php> . Also see [https://www.earthobservations.org/geoss\\_dsp.shtml](https://www.earthobservations.org/geoss_dsp.shtml) .
- [11] OECD Principles and Guidelines for Access to Research Data from Public Funding. Available from <https://www.oecd.org/sti/sci-tech/38500813.pdf> .
- [12] European Commission, Directorate-General for Research & Innovation: Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020 (Section 4). Available from [http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-pilot-guide\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf) .  
See in particular Article 29 of the Horizon 2020 Model Grant Agreement, available from [http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/amga/h2020-amga\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf)
- [13] V. Tsoukala et al: EU RECODE, Policy RECommendations for Open access to research Data in Europe, Deliverable D5.1: Policy guidelines for open access and data dissemination and preservation, Feb. 2015, available from <http://recodeproject.eu/wp-content/uploads/2015/02/RECODE-D5.1-POLICY-RECOMMENDATIONS- FINAL.pdf> .
- [14] Directive 2007/2/EC of the European Parliament and Council of 14 March 2007, establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), and its Implementing Rules. Available from <http://inspire.ec.europa.eu/index.cfm/pageid/3> .
- [15] A World that Counts. Mobilising the Data Revolution for Sustainable Development. Report prepared at the request of the United Nations Secretary-General, by the Independent Expert Advisory Group on a Data Revolution for Sustainable Development. November 2014. Available from <http://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf>
- [16] Current Best Practice for Research Data Management Policies. A Memo for the Danish e-Infrastructure Cooperation and the Danish Digital Library Simon Hodson and Laura Molloy, ICSU-CODATA, May 2014. Available from <http://dx.doi.org/10.5281/zenodo.27872>
- [17] Draft International Principles on the Legal Interoperability of Research Data, and Draft Implementation Guidelines for the Principles on the Legal Interoperability of Research Data. RDA-CODATA Working Group on Legal Interoperability of Research Data. Available from <https://rd-alliance.org/groups/rdacodata-legal-interoperability-ig/wiki/rdacodata-legal-interoperability-ig-wiki-index.html> .
- [18] Interpretation of the “full and open” access to and use of (geographic) data: existing approaches. Living paper of the GEO Data Sharing Working Group October 2013. Available from [https://www.earthobservations.org/documents/dsp/201310\\_full\\_and\\_open\\_principle\\_interpretation\\_living\\_paper.pdf](https://www.earthobservations.org/documents/dsp/201310_full_and_open_principle_interpretation_living_paper.pdf) .
- [19] Implementation Guidelines for the GEOSS Data Sharing Principles. Available from [https://www.earthobservations.org/documents/geo\\_vi/07\\_Implementation%20Guidelines%20for%20the%20GEOSS%20Data%20Sharing%20Principles%20Rev2.pdf](https://www.earthobservations.org/documents/geo_vi/07_Implementation%20Guidelines%20for%20the%20GEOSS%20Data%20Sharing%20Principles%20Rev2.pdf) .
- [20] White Paper: Mechanisms to Share Data as Part of GEOSS Data-CORE. GEO Data Sharing Working Group. Available from [https://www.earthobservations.org/documents/dswg/Annex%20VI%20-%20Mechanisms%20to%20share%20data%20as%20part%20of%20GEOSS%20Data\\_CORE.pdf](https://www.earthobservations.org/documents/dswg/Annex%20VI%20-%20Mechanisms%20to%20share%20data%20as%20part%20of%20GEOSS%20Data_CORE.pdf) .

- [21] Science as an open enterprise. The Royal Society Science Policy Centre report 02/12. Available from <https://royalsociety.org/topics-policy/projects/science-public-enterprise/report/> .
- [22] Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014. Available from <https://www.force11.org/group/joint-declaration-data-citation-principles-final> .
- [23] Martina Stockhause, Frank Toussaint, Michael Lautenschlager for the WCRP WGCM Infrastructure Panel: CMIP6 Data Citation and Long-Term Archival. V5, available from [https://www.earthsystemcog.org/projects/wip/position\\_papers](https://www.earthsystemcog.org/projects/wip/position_papers) .
- [24] United States Executive Office of the President, Office of Science and Technology Policy, 2013 Memorandum on “Increasing Access to the Results of Federally Funded Scientific Research”. Available from [https://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp\\_public\\_access\\_memo\\_2013.pdf](https://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf) .
- [25] Report of the 34th Session of WCRP Joint Scientific Committee, 2013. Available from [http://www.wcrp-climate.org/images/documents/reports\\_flyers/WCRP\\_JSC34\\_report.pdf](http://www.wcrp-climate.org/images/documents/reports_flyers/WCRP_JSC34_report.pdf) .
- [26] Belmont Forum Data Policy and Principles. Adopted at the Belmont Forum’s 10th Plenary Meeting, Oslo, 2015. Available from <https://belmontforum.org/announcements/2016/belmont-forum-adopts-open-data-principles-environmental-change-research> .
- [27] Data Citation of Evolving Data, Recommendations of the RDA Working Group on Data Citation (WGDC). A. Rauber, A. Asmi, D. van Uytvanck, S. Pröll, 2015. Available from [https://rd-alliance.org/system/files/RDA-DC-Recommendations\\_151020.pdf](https://rd-alliance.org/system/files/RDA-DC-Recommendations_151020.pdf) .
- [28] [https://en.wikipedia.org/wiki/Sui\\_generis\\_database\\_right](https://en.wikipedia.org/wiki/Sui_generis_database_right) , also see <https://wiki.creativecommons.org/wiki/Data> .
- [29] Austin, C., et al. for the RDA-WDS Publishing Data Workflows group: Key components of data publishing: using current best practices to develop a reference model for data publishing. International Journal on Digital Libraries International Journal on Digital Libraries. doi:10.1007/s00799-016-0178-2
- [30] Peng, Ge, et al. "Scientific Stewardship in the Open Data and Big Data Era: Roles and Responsibilities of Stewards and Other Major Product Stakeholders." D-Lib Magazine 22.5 (2016): 1. Available from <http://www.dlib.org/dlib/may16/peng/05peng.html> .
- [31] Castelli, Donatella, Paolo Manghi, and Costantino Thanos. "A vision towards scientific communication infrastructures." International Journal on Digital Libraries 13.3-4 (2013): 155-169.

**The  
World Climate  
Research Programme  
(WCRP)**

*facilitates analysis and  
prediction of Earth system change  
for use in a range of practical  
applications of direct relevance,  
benefit and value to society.*

