

IPCC DATA AND CODE LICENSING GUIDELINES

Recommendations from the IPCC Task Group on Data Support for Climate Change Assessments

David Huard, Anna Pirani, Robert Chen, José Manuel Gutiérrez, Martin Jukes,
Volker Krey, Alessandro Spinuso, Martina Stockhause

Version 1.1, December 2022

This guidance document is an agreed product of the IPCC Task Group on Data Support for Climate Change Assessments (TG-Data). The [mandate](#) of the TG-Data was approved by the IPCC, but this does not imply the Panel's endorsement or approval of documents emanating from the TG-Data or any recommendations contained herein. Furthermore, this document has not been subject to the procedural IPCC review processes.

Document history, acknowledgements and citation details

This document, IPCC Data and Code Licensing Guidelines: Recommendations from the IPCC Task Group on Data Support for Climate Change Assessments, serves as “Supporting Material” of the Intergovernmental Panel on Climate Change (IPCC), according to the Appendix A (Procedures for the Preparation, Review, Acceptance, Adoption, Approval and Publication of IPCC Reports) to the Principles Governing IPCC Work. The document, which was prepared by the IPCC Task Group on Data Support for Climate Change Assessments (TG-Data), has not been subject to the procedural IPCC review process. Nonetheless, the document was reviewed by the IPCC Legal Officer. The authors gratefully acknowledge the comments and suggestions received. The information contained herein, including opinions and possible errors are nonetheless solely the responsibility of the authors.

Initial draft	M. Jukes	December 2020
Comments and suggestions	M. Stockhause, R. Chen, J.M. Gutierrez Llorente, D. Huard	January 2021
Review & discussion	TG-Data	February 2021
Review	J. Lew Schneider	February 2021
Requested changes	M. Jukes	April 2021
Review & discussion	TG-Data	April 2022
Requested changes	D. Huard, A. Pirani	June 2022
Review & discussion	TG-Data	July 2022
Preparation of final draft	A. Pirani, D. Huard, M. Stockhause	August 2022
Final edits	A. Pirani	September 2022
Finalisation of document, Code licensing	TG-Data meeting, A. Spinuso, V. Krey	October 2022
Rewording from policy to guidelines	D. Huard	December 2022

This document should be referenced as: Huard, D., Pirani, A., Chen, R., Gutiérrez, J. M., Jukes, M., Krey, V., Spinuso, A., and M. Stockhause, 2022: The IPCC Data and Intellectual Property: TG-Data Recommendations, Supporting Material, Intergovernmental Panel on Climate Change Task Group on Data Support for Climate Change Assessments (TG-Data), 8 pp, DOI: 10.5281/zenodo.7431834 (<https://zenodo.org/record/7431834>).

Table of Contents

1 Introduction.....	4
2 Motivation.....	5
3 The material to be licensed	6
4 TG-DATA Recommendations	7
References	7
Appendix 1: IPCC Data and Code Licensing Guidelines	8

1 Introduction

The IPCC assessment reports are based on the expert assessment of scientific and technical literature. Underlying datasets published in the literature are used extensively in the assessment process to corroborate, illustrate and communicate assessment conclusions. The tables and figures of assessment reports become important resources for activities associated with the climate change science-policy interface and the implementation of policy decisions taken on the basis of the reports. Over the Sixth Assessment Report (AR6) cycle, one of the main objectives of the Task Group on Data Support for Climate Change Assessments (TG-Data)¹, alongside the IPCC Data Distribution Centre [1], has been to make data underlying IPCC data products publicly available, in accordance with the FAIR Principles [2] [3].

IPCC reports are published under a copyright license that prohibits commercial use and the creation of derivative products, unless discussed first and then given permission by the IPCC Secretariat (see <https://www.ipcc.ch/copyright/>, Aug 19, 2022). This license is applied to protect IPCC reports from distortion since these are accepted by member governments, or approved in the case of the Summary for Policymakers, and adopted in the case of the Synthesis Report. If the same license was applied to data products, it would severely limit their usefulness and value. A different IPCC data license is required to allow the creation of derivatives for the pursuit of research and the re-use of IPCC data-based products for national assessments, adaptation and mitigation policies. Guidance is also needed to navigate license arrangements with data providers, publishers and users in relation to the FAIR data access objectives of IPCC data products.

This document summarises in [Section 2](#) the motivation to establish IPCC Data Guidelines, and describes the material to be licensed in [Section 3](#). TG-Data recommendations to ensure the availability and use of IPCC data products and help clarify the rights and obligations of both users and contributors are presented in [Section 4](#). The IPCC Data Policy recommended by TG-Data is presented in [Appendix 1](#).

¹ Per its terms of reference, one of the purposes of TG-Data is to “Facilitate in cooperation with the Data Distribution Center (DDC) the availability and use of climate change related data resulting from the activities of the IPCC in accordance with the mandate of the IPCC.” The IPCC and many of the organisations responsible for data products which underpin IPCC assessments are moving towards clearer formulations of license and intellectual property right (IPR) policies.

2 Motivation

Licensing of IPCC material, with clear and consistent meaning in all legal jurisdictions, is essential to facilitate its appropriate use to address pressing climate change challenges, while protecting the rights of data providers. Open access and transparency of scientific data have been widely recognized as the “bedrock” of modern scientific progress (Science International, 2015 [4]). A restrictive license of the form used for the publication of assessment reports would severely limit access and use of IPCC data products.

The IPCC restricts the rights of users over content that is accessed from its web domain (ipcc.ch) and published reports. This is standard practice for printed publications in order to prevent distortion and manipulation of content by others. This ensures that content that carries the IPCC seal of approval is fully protected. The application of the same restrictions to IPCC data products would conflict with the goal of implementing FAIR data principles in the IPCC process.

There are many complexities associated with data licensing, especially when applied to an activity on the scale of the IPCC assessment reports, which combine international collaborative efforts involving thousands of scientists and the rigorous review and approval process involving all member states of the IPCC. The license requirements for the IPCC need to be clearly communicated to users around the world that operate in a variety of languages and legal jurisdictions. Another issue is handling restrictions imposed by multiple copyright holders, which may be incompatible to one another and curtail the publication of derivative products.

In developing IPCC Data Guidelines, TG-Data recognises that many organisations seeking to build on the findings of IPCC assessments rely on guidance from commercial services, and those services need access to the data underlying IPCC assessments. It also takes into account the fact that a “ShareAlike” (also known as “copyleft”) licensing constraint can act as a barrier to commercial users. Some scientific journals refuse to publish data licensed under ShareAlike licenses and such licenses could restrict further use of these data in interdisciplinary research and applications that build on the IPCC assessment reports. The proposed Data Policy seeks to avoid such constraints, enable re-use and promote iterative improvements, a widely used approach in the publishing of software and research data.

3 The Material to be Licensed

The assessment reports include tables, figures and maps illustrating key concepts, analyses and conclusions. These typically synthesise data from many different sources. For instance, a figure or a map may combine climate projections averaged over multiple climate model simulations, with observations from satellite instruments. TG-Data distinguishes three categories of data: input data, intermediate assessment data, and final assessment data.

Input Data

Input data denotes the source data that underpins information in the assessment reports. It is typically authored by credible, authoritative, trusted sources, who decide under which license it is published. In some instances, key datasets published under restrictive conditions can be made available to IPCC authors through special exemptions and waivers.

Intermediate Assessment Data

Intermediate assessment data is the outcome of data processing and analysis performed as part of the assessment as an intermediate step in the generation of final assessment data. Data is only defined as intermediate if it has gone through non-trivial processing to be considered an original product, distinct from the input data. These intermediate datasets may be of high value in terms of ensuring transparency around the assessment process, as an outcome of an expert-based assessment of the underlying evidence for the report, and as a resource for users trying to replicate IPCC results or to implement policies based on the IPCC assessment findings. Restrictions may be imposed by input data licenses on this derivative product's license.

Final Assessment Data

Final assessment data refer to data which is directly presented in data tables or graphically displayed (e.g. as a line graph or a spatial map) in the report. As with intermediate assessed data, input data licenses may impose restrictions on this derivative product's license.

Source Code

Source code refers to scripts, online code repositories, and software libraries written to create intermediate and final assessed data, as well as the figures included in the reports. Distribution of this code by the IPCC facilitates the re-use and further development by users of IPCC data products.

4 TG-DATA Recommendations

TG-Data recommends the following so that stakeholders benefit as much as possible from the fruits of the IPCC's efforts.

- 1 Input data copyright holders are encouraged to adopt well-known licenses enabling broad usage, including commercial use, and avoid "ShareAlike" licenses. ShareAlike clauses grant freedom to use and modify the data, but derivative products must adopt the same license as the original. This can severely limit the ability to combine multiple datasets whose licenses conflict, and some scientific journals do not publish data under a ShareAlike license.
- 2 Intermediate and final assessment data should be licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, where this does not infringe the interests of relevant license holders. The Creative Commons family of licenses are designed to provide legal interoperability across virtually all jurisdictions. The CC BY 4.0 license requires users of the data to properly credit its authors, but does not restrict how the data is used or distributed.
- 3 When input datasets are published under restrictive licenses, waivers or exemptions can be sought for the IPCC assessment reports. These waivers should be negotiated with copyright holders by Working Group co-chairs, with guidance from TG-Data representatives. These waivers would ensure that derivative products can be licensed by the IPCC under CC BY 4.0, and that the version used by the assessment report is curated in a long-term archive, either by IPCC DDC [4] or another trusted data repository.
- 4 To ensure maximal reusability of source code, similarly to data, code should be published under permissive (non-copyleft) open source licenses that do not restrict commercial use. Some popular permissive software licenses include ISC, Apache 2.0, MIT, BSD2. Particularly, Apache 2.0 explicitly allows users to claim patents on derivative projects, requiring major changes to be described in any released derivation. Concerning software developed for the preparation of the report, applying Apache 2.0 does not authorise the reuse of branding and trademarks associated with the IPCC, as well as with organisations owning the software modules used by the original code.

References

- [1] "Guidance for the core functions of the IPCC Data Distribution Centre (DDC)" (https://www.ipcc.ch/site/assets/uploads/2018/12/Guidance_DDC.pdf): "The purpose of the IPCC Data Distribution Centre (DDC) is to archive and provide transparency, traceability, and stability of data and scenarios that are relevant in the context of the IPCC."
- [2] Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>
- [3] Anna Pirani, Andrés Alegria, Alaa Al Khourdajie, Wawan Gunawan, José Manuel Gutiérrez, Kirstin Holsman, David Huard, Martin Juckes, Michio Kawamiya, Nana Klutse, Volker Krey, Robin Matthews, Adam Milward, Charlotte Pascoe, Gerard van der Shrier, Alessandro Spinuso, Martina Stockhause, & Xiaoshi Xing. (2022). The implementation of FAIR data principles in the IPCC AR6 assessment process. Zenodo. <https://doi.org/10.5281/zenodo.6504469>
- [4] Science International (2015): Open Data in a Big Data World. Paris: International Council for Science (ICSU), International Social Science Council (ISSC), The World Academy of Sciences (TWAS), InterAcademy Partnership (IAP). . https://council.science/wp-content/uploads/2017/04/open-data-in-big-data-world_long.pdf

Appendix 1: IPCC Data and Code Licensing Guidelines

About these guidelines

Data needs to be accessible to both IPCC authors and external users. Undue restrictions should not be placed on the users, while, at the same time, the legitimate interests of the owners of data should be protected. In the interests of preserving clarity and transparency while protecting the rights of data providers, the Creative Commons Attribution license (CC BY 4.0) is the recommended data license for the IPCC. The CC BY 4.0 license would be used for derived data products generated in IPCC assessments, as and where possible. TG-Data and DDC representatives will provide guidance to seek exemptions from any restrictions imposed by input data licenses. Open-source licenses are recommended for code that underlies IPCC data products.

The IPCC data and code licensing guidelines

- 1 Input data, i.e. data curated by the DDC as originally published by the data providers, shall be licensed under the same license terms and conditions imposed by the data providers.
- 2 Data produced as part of the IPCC assessment, be it intermediate or final assessment data, shall be published, wherever possible, under the CC BY 4.0 license. If exemptions cannot be obtained from the copyright owners, the applicable licenses of input data will apply.
- 3 Source code that underlies IPCC data products is encouraged to be published using permissive (non-copyleft) open source licenses.