CMIP6 data in the IPCC AR6: The IPCC FAIR Guidelines

Martina Stockhause (DKRZ/IPCC DDC; stockhause@dkrz.de)

Anna Pirani, Martin Juckes, José Gutierrez, Robin Matthews, Charlotte Pascoe, Robert Chen, Xiaoshi Xing, David Huard and many others

WGCM-24, 2021-12-09

CMIP6 data in the IPCC AR6: Concepts

CMIP6 infrastructure endorsed by the WIP includes:

- ESGF providing CMIP6 dataset
- Data Citation Service providing DataCite DOIs for CMIP6 data: All CMIP6 data were citable on the AR6 WGI data and literature cut-off date 2021-01-31); and
- Long-Term Preservation of CMIP6 data in the IPCC Data Distribution Centre (DDC) at DKRZ

CMIP Phase 6 (CMIP6)

References:
Stockhause, M. et al., 2015. CMIP6 Data Citation and Long-Term Archival (WIP white paper). https://doi.org/10.5281/zenodo.35178

IPCC FAIR Guidelines: Concept

Enhancing the transparency of IPCC outputs

IPCC FAIR Guidelines
complement the established transparent IPCC assessment procedure by FAIR data and CoreTrustSeal repository standards (FAIR and TRUST principles).

Digital objects included:
- Input data: used by IPCC authors, e.g. CMIP6
- Scripts: applied on input data to create figures and tables
- Final Data: created by IPCC authors for figures and tables

Aspects: Traceability – Credit – Preservation

AR6 Climate Change 2021: The Physical Science Basis

The Working Group I contribution to the Sixth Assessment Report is expected to be finalized in 2021.

Findable Accessible Interoperable Reusable

TRUST
Transparency Responsibility User focus Sustainability Technology

References:
Stockhause, M. et al., 2019. Data Distribution Centre Support for the IPCC Sixth Assessment. http://doi.org/10.5334/dsj-2019-020

CMIP6 data in the IPCC AR6: IPCC FAIR Guidelines

Enhancing the transparency of IPCC's outputs:
Traceability – Credit – Preservation
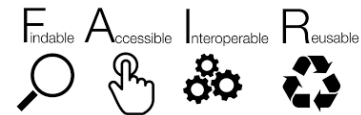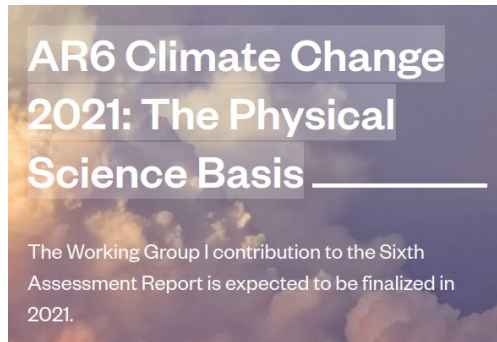
Providing traceable accounts for key findings
has been guiding IPCC's assessments.

Traceability of key statements of the report by
documentation of data usage and figure/table creation

- Supplementary Materials document data usage
  and figure creation enabling figure reproducibility
- References to data from digital AR6
- References to AR6 chapters are added to the
  DDC Data Archive for CMIP6 input data enabling
  machine-access (e.g. via Scholix).

CMIP6 data in the IPCC AR6: IPCC FAIR Guidelines

Enhancing the transparency of IPCC's outputs:
Traceability – Credit – Preservation

Credit for CMIP6, other input data, created scripts
and final data:

- CMIP6 data is cited in AR6 WGI Annex II and the Interactive Atlas
- Created final datasets and scripts (figures, tables) are citable by DOIs

Data Citation requirements:
- CMIP6 data licenses (BY in CC BY-[NC-]SA) require to give appropriate credit to the creators.
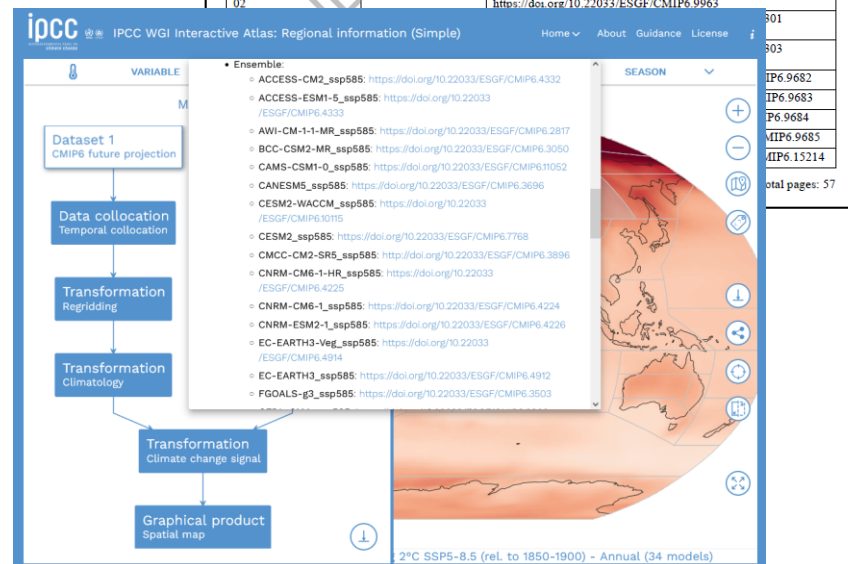- Citing original data sources, software, materials are part of Good Scientific Practice.

References:
DFG (2019). Guidelines for Safeguarding Good Research Practice. https://doi.org/10.5281/zenodo.3923602

CMIP6 data in the IPCC AR6: IPCC FAIR Guidelines

Enhancing the transparency of IPCC's outputs:
Traceability – Credit – Preservation

Preservation of curated core CMIP6 input data, scripts, and final data:

- CMIP6 data underpinning the AR6 is curated by enriching metadata with Controlled Vocabulary, Citation Service (authors, references) and ES-DOC information and preserved on the long-term in the IPCC DDC
- Scripts used for figure creation are preserved in GitHub/Zenodo
- Final data is preserved on the long-term at CEDA [Metadataworks]
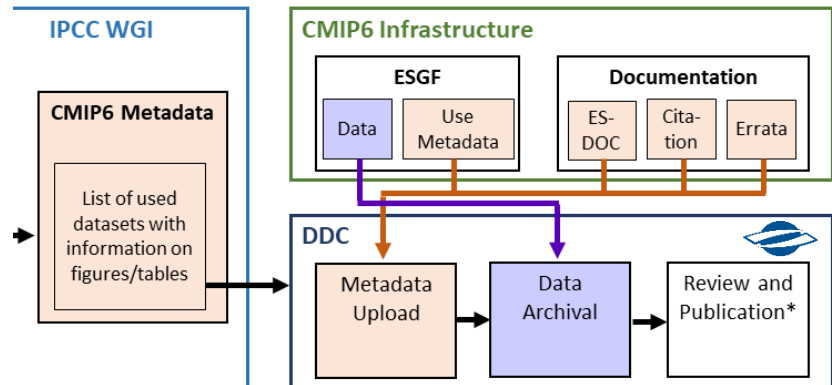
* DataCite publication and publication on IPCC webpages
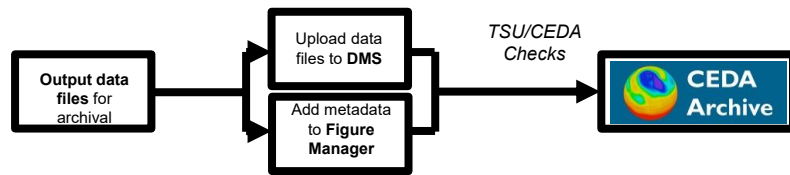
Future Perspectives

IPCC FAIR Guidelines introduced best practices in data management and thereby enhanced IPCC AR6's transparency.

IPCC FAIR Guidelines:

TG-Data derives recommendations for AR7 based on AR6 experiences, for

- Review of FAIR Guidelines' content and implementation, and hand-over to AR7
- Integration of the FAIR Guidelines into the IPCC assessment procedure (e.g. FOD/SOD and data cut-off date) and revisit protocols (e.g. Errata Protocol)
- Long-term strategy / sustainability / funding

Requests from IPCC TG-Data:

- Use of CC BY licenses - no exemption for IPCC required
- Earlier data availability
  - → schedule coordination
  - → prioritization of IPCC-relevant experiments or even definition of an IPCC CMIP7 subset
- Importance of data and metadata quality
- Make further input data citable, e.g. CORDEX; funding issue for the Citation Service

Future Perspectives

IPCC FAIR Guidelines serve as best practice example for climate information and climate service providers

Best practice examples:

- Credit for the data gets implemented in services like IPCC WGI Interactive Atlas (IA)
- Traceability/provenance gets built into software packages/tools like ESMValtool, Metaclip (IPCC IA)
- Long-term preservation of FAIR data: enriched documentation, long-term curation and adjustments to provide continuous access to the data underpinning policy decisions (CoreTrustSeal)

→ Promotion of IPCC's FAIR Guidelines as best practice examples…