

IPCC Sixth Assessment approaches towards FAIR data and an enhanced data reuse

Martina Stockhause¹, Alaa Al Khourdajie², Andres Alegria³, Robert Chen⁴, David Huard⁵, Martin Juckes⁶, Charlotte Pascoe⁶, Anna Pirani⁷, Robin Matthews⁸, Elvira Poloczanska⁹, Sebastian Vicuna¹⁰, Xiaoshi Xing⁴, and Özge Yelekçi⁸

¹German Climate Computing Center / Deutsches Klimarechenzentrum GmbH (DKRZ)

²Imperial College London

³IPCC WGII Technical Support Unit (TSU)

⁴Columbia University

⁵Ouranos

⁶STFC

⁷Technical Support Unit of IPCC WGI

⁸IPCC WGI Technical Support Unit (TSU)

⁹Intergovernmental Panel on Climate Change, WGII Technical Support Unit

¹⁰Centro de Cambio Global UC

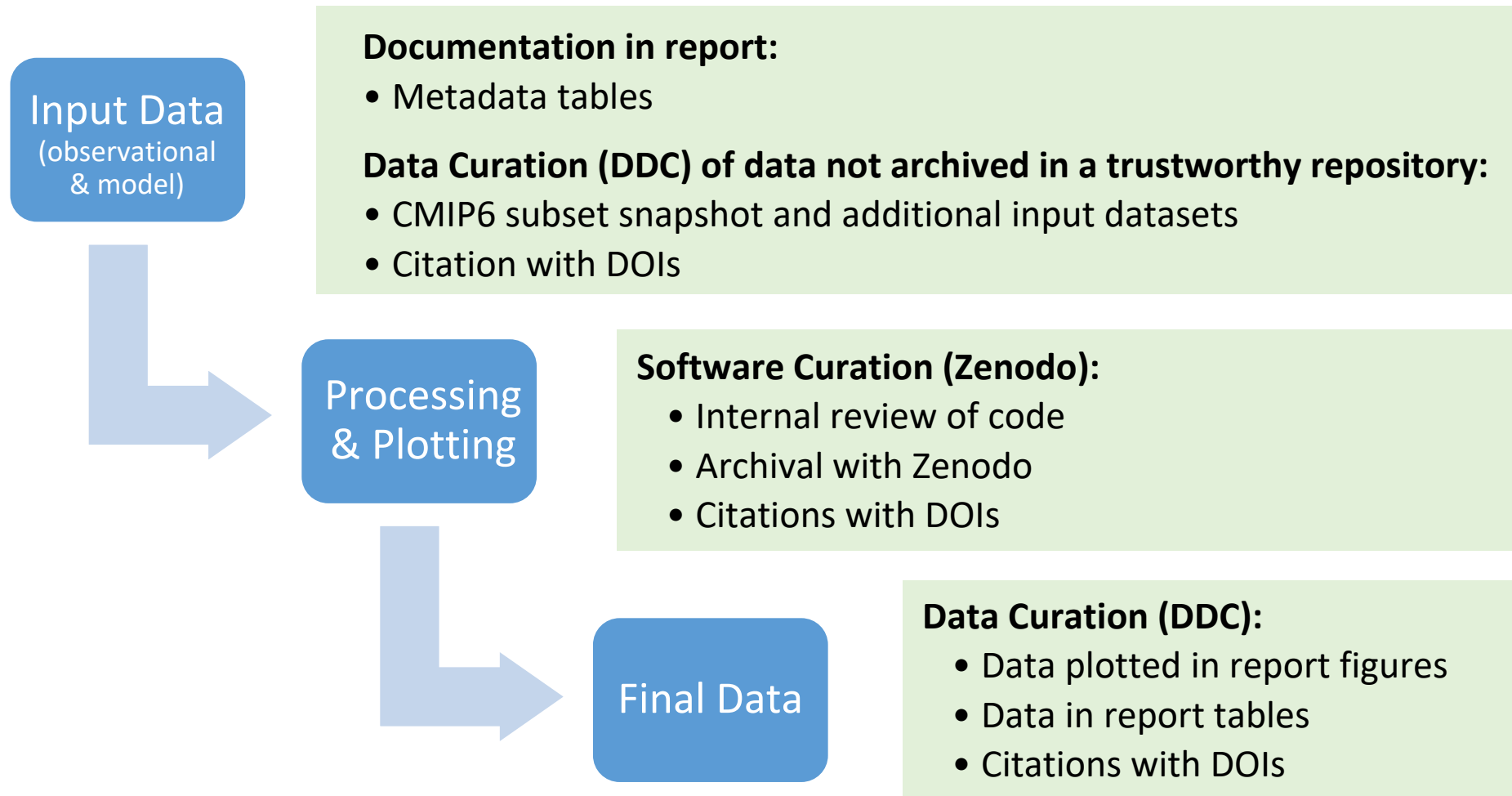
November 24, 2022

Abstract

The Intergovernmental Panel on Climate Change (IPCC) currently prepares its Sixth Assessment Report (AR6). Its authors assess peer-reviewed scientific literature and recent climate datasets to inform policy-makers about the current state of the science regarding climate change and its impacts, as well as adaptation and mitigation options. For AR6, efforts are underway to make its main results FAIR and preserve them in the TRUSTworthy repositories of the IPCC Data Distribution Centre (DDC), jointly managed by CEDA, DKRZ, and CIESIN. The AR6 FAIR initiative was kickstarted by the IPCC DDC and Working Group I (WGI) [Stockhause et al., 2019], then adopted by IPCC TG-Data (Task Group on Data Support for Climate Change Assessments) shortly after its creation. All three WGs have adopted the FAIR data guidelines. IPCC assessments are large and diverse in terms of scientists involved as well as included scientific objects. Challenges for digital data curation are related to the scale and diversity of papers, reports, datasets, the variety of software, and the different familiarity of the scientists with these technical aspects. The following priority areas for improved data stewardship were selected based on the aims to enhance the traceability of AR6 key findings and their reusability: preserve figure datasets in the DDC; - preserve analysis software; . preserve main input datasets in the DDC; . assemble datasets and provenance information on the figure creation from IPCC authors; and . interlink datasets to the IPCC report. Datasets are transferred to the DDC at the end of AR6. The DDC partners are responsible to preserve the data for future reuse by different stakeholders and under a variety of current and future scientific and policy-related questions. As the role for the DDC expands within the IPCC, new partners are sought. The TRUST principles provide a framework for the communication of DDC tasks to different stakeholders, e.g. to countries interested to host a DDC. The presentation will give an overview over the IPCC AR6 approaches towards FAIR data maintained in TRUSTworthy repositories, their challenges, their approach to meet these challenges and open questions, e.g. the integration of digital data into the IPCC Error Protocol, targeted within TG-Data.

IPCC Sixth Assessment approaches towards FAIR data and an enhanced data reuse

The approach towards FAIR data and software with long-term stewardship in the DDC provides the first step of a long-term effort.



IPCC Sixth Assessment approaches towards FAIR data and an enhanced data reuse

The motivation for these approaches is the enhancement of the traceability, transparency, and accessibility of IPCC products.

Open Issues / Gap Analysis

1. Exhaustive IPCC data archival / Recruitment of new partners
2. Integration into IPCC procedures and products
3. DDC service improvements around data discovery and access

IPCC Sixth Assessment approaches towards FAIR data and an enhanced data reuse

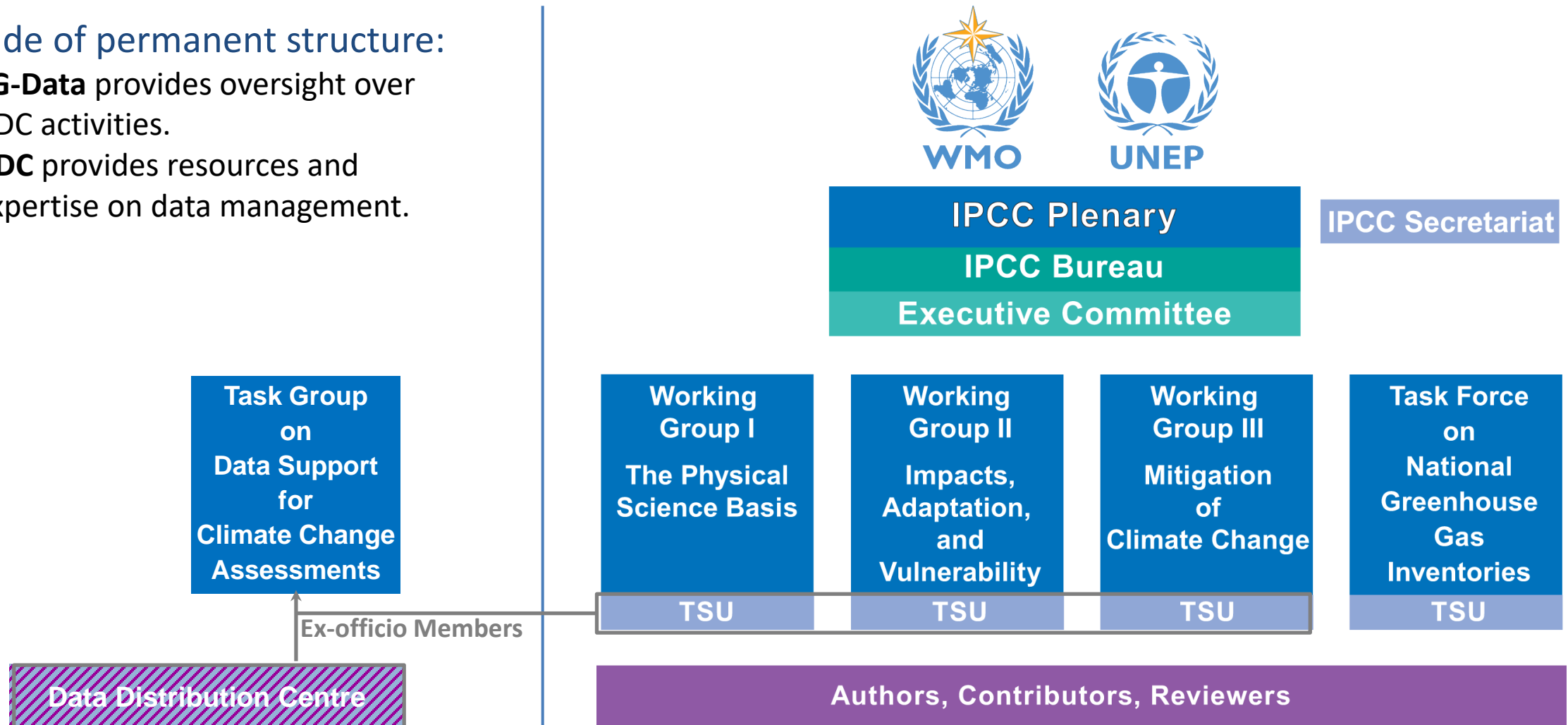
AGU Fall Meeting 2020
1-17 December 2020

Martina Stockhause, Alaa Al Khourdajie, Andres Alegria, Robert S Chen, David B Huard, Martin N Jukes, Charlotte L Pascoe, Anna Pirani, Robin Matthews, Elvira Poloczanska, Sebastian Vicuna, Xiaoshi Xing, and Özge Yelekçi

Introduction: Intergovernmental Panel on Climate Change (IPCC)

Outside of permanent structure:

- **TG-Data** provides oversight over DDC activities.
- **DDC** provides resources and expertise on data management.



DDC Partners

- Centre for Environmental Data Analysis (CEDA, UK)
- Deutsches Klimarechenzentrum (DKRZ, Germany)
- Center for International Earth Science Information Network (CIESIN, US)

DDC Partner Candidate

- Santander Meteorology Group (UC-CSIC, Spain)

DDC's Role

Main task (subject of the MoU):

- Long-term stewardship of data underlying the IPCC Assessment Reports complying with international standards: CoreTrustSeal – WDS Regular Membership

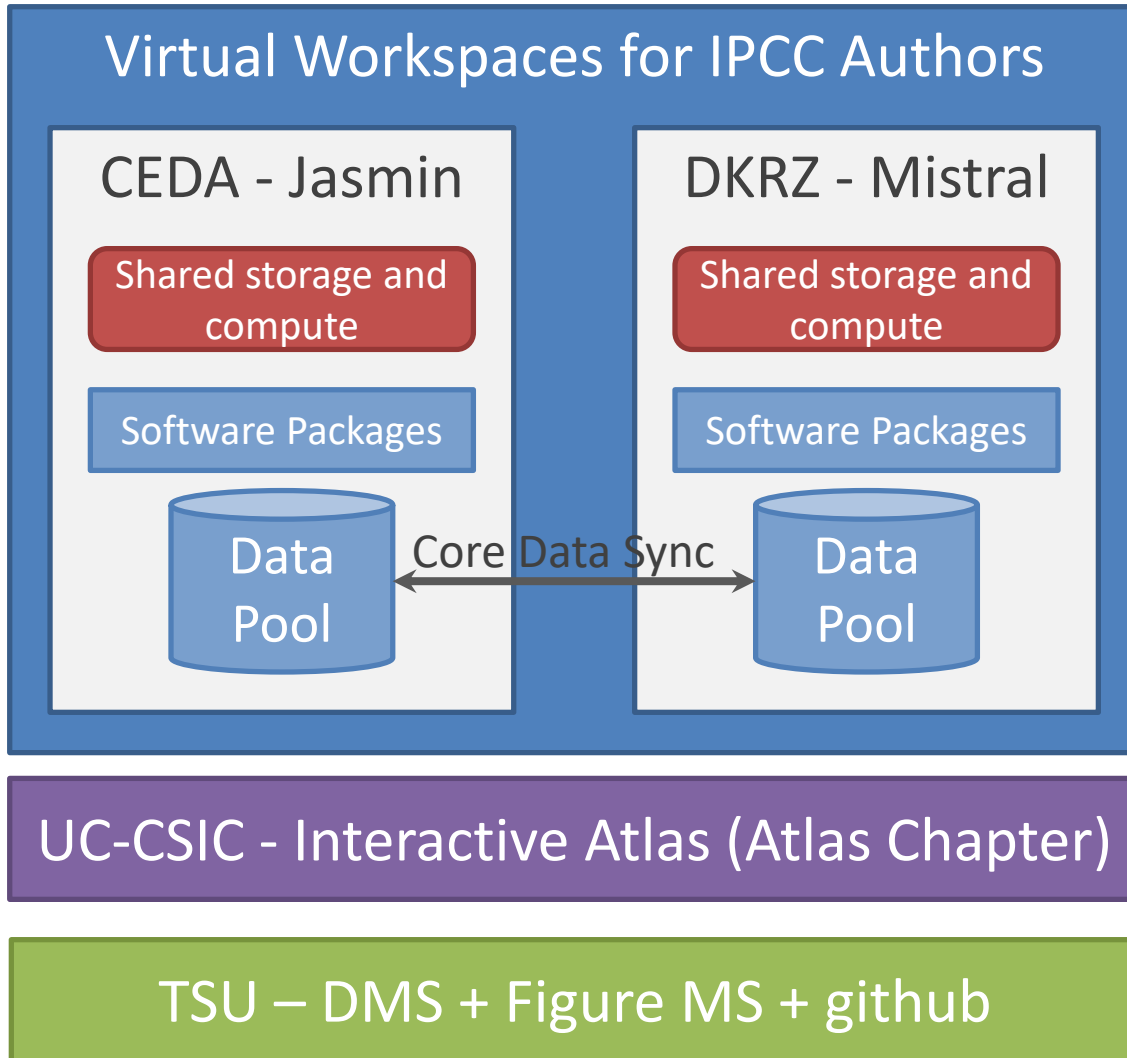
New tasks in AR6 (as best effort):

- provide support for the Working Groups and IPCC authors during the Assessment
- provide expert advice for data-related questions
- contribute to the integration of data into IPCC products, procedures, and protocols

II. FAIR Data Guidelines

FAIR Guidelines: Preparing the WGI AR6

Preparing WGI AR6



- Virtual Workspaces support IPCC Authors in analyzing **input datasets** and creating **final datasets** for figures for the AR6.
- **Atlas** chapter of AR6 WGI inherits an Interactive Atlas with figure creation functionality on the long-term.
- **TSU** coordinates IPCC authors in AR6 creation and collect dataset information.

31 January 2021: Literature acceptance cut off
(data snapshot date)

FAIR Guidelines

Transparency and Accessibility

- Implementation of IPCC Errata Protocol for assessed digital information
- Long-term stewardship of assessed digital information

Benefits/Purpose

- Contributes to IPCC's integrity
- Enables long-term reuse
- Enables external checking

Traceability

- Documentation of input data used
- Citing input data
- Interlinking report, final data and input data

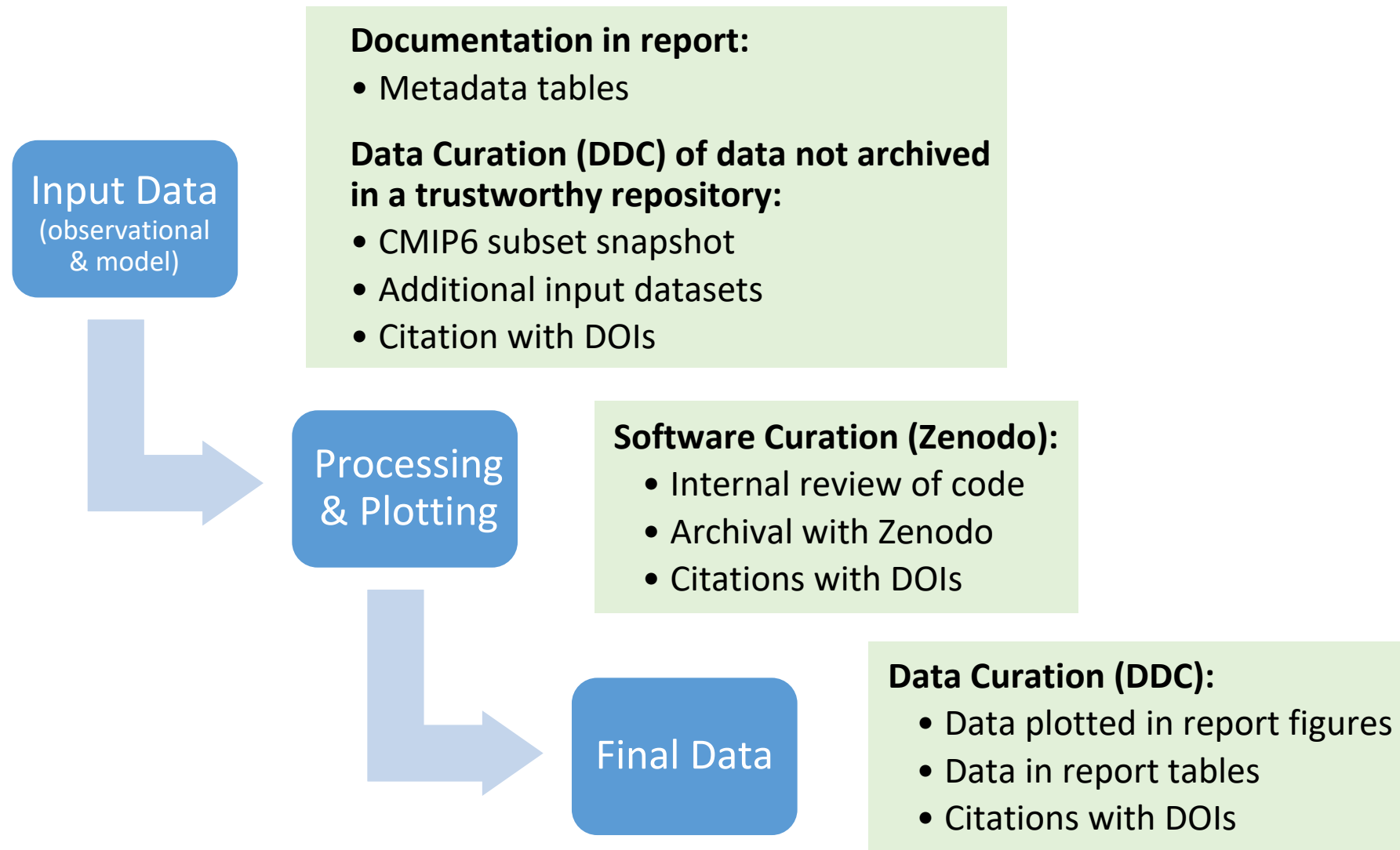
- Credit for data creators
- Enhanced discoverability of IPCC results

Reproducibility

- Archival of final data and software

- Enables re-creation of figures
- Enables repurposing of data
- Enables external checking
- Credit for data creators

I. FAIR Guidelines: Approaches to FAIR in AR6 (WGI Example)



I. FAIR Guidelines: Approaches to FAIR

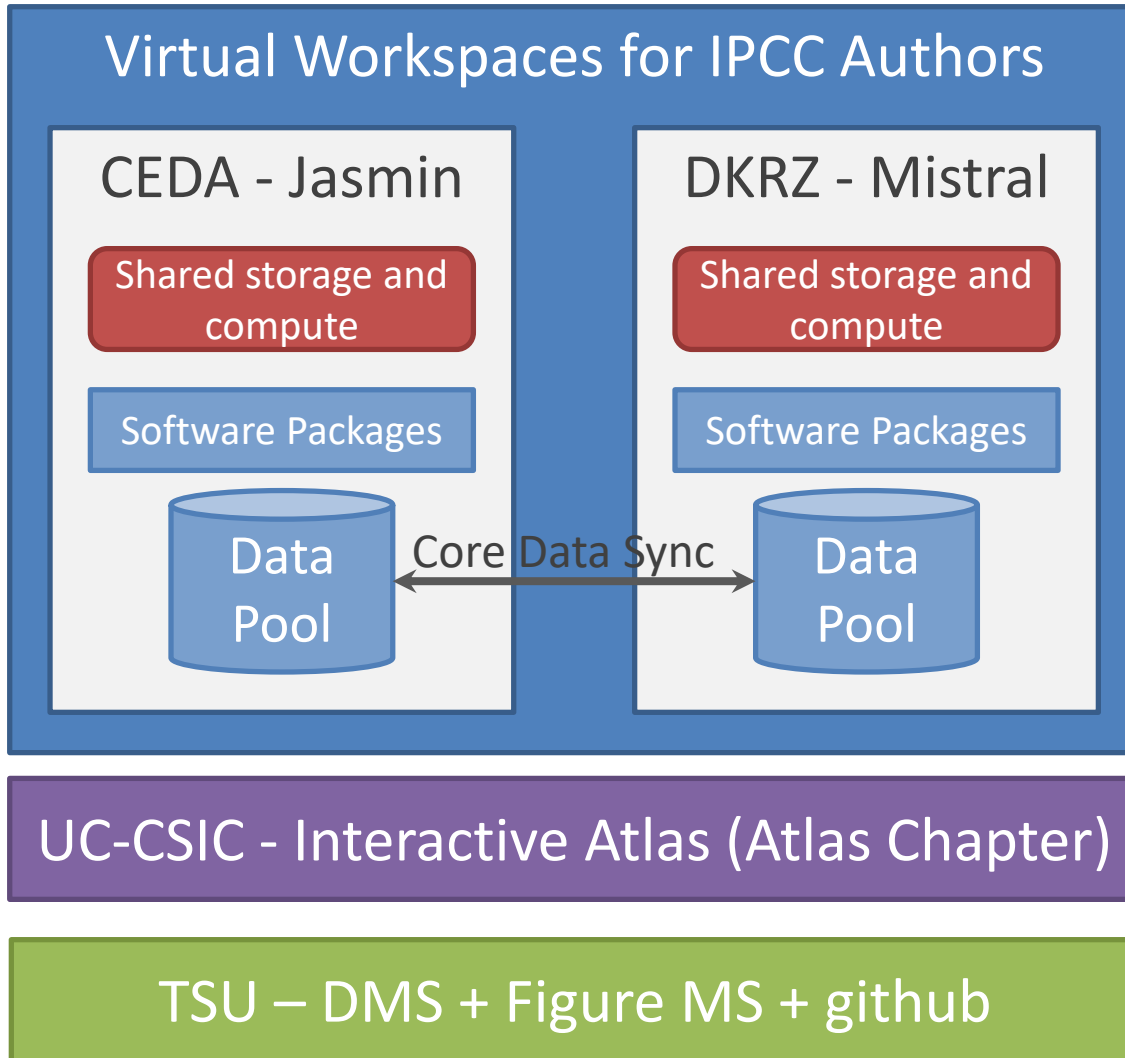
WG-Specific Differences in Approaching FAIR

	Focus in AR6	Challenges
WGI	Final data, software, selected and prioritized input datasets	Input data volume
WGII	final data and software (first priority SPM and TS), few input data	
WGIII	final data and software (focus on SPM and TS), no input data	License issues with input data

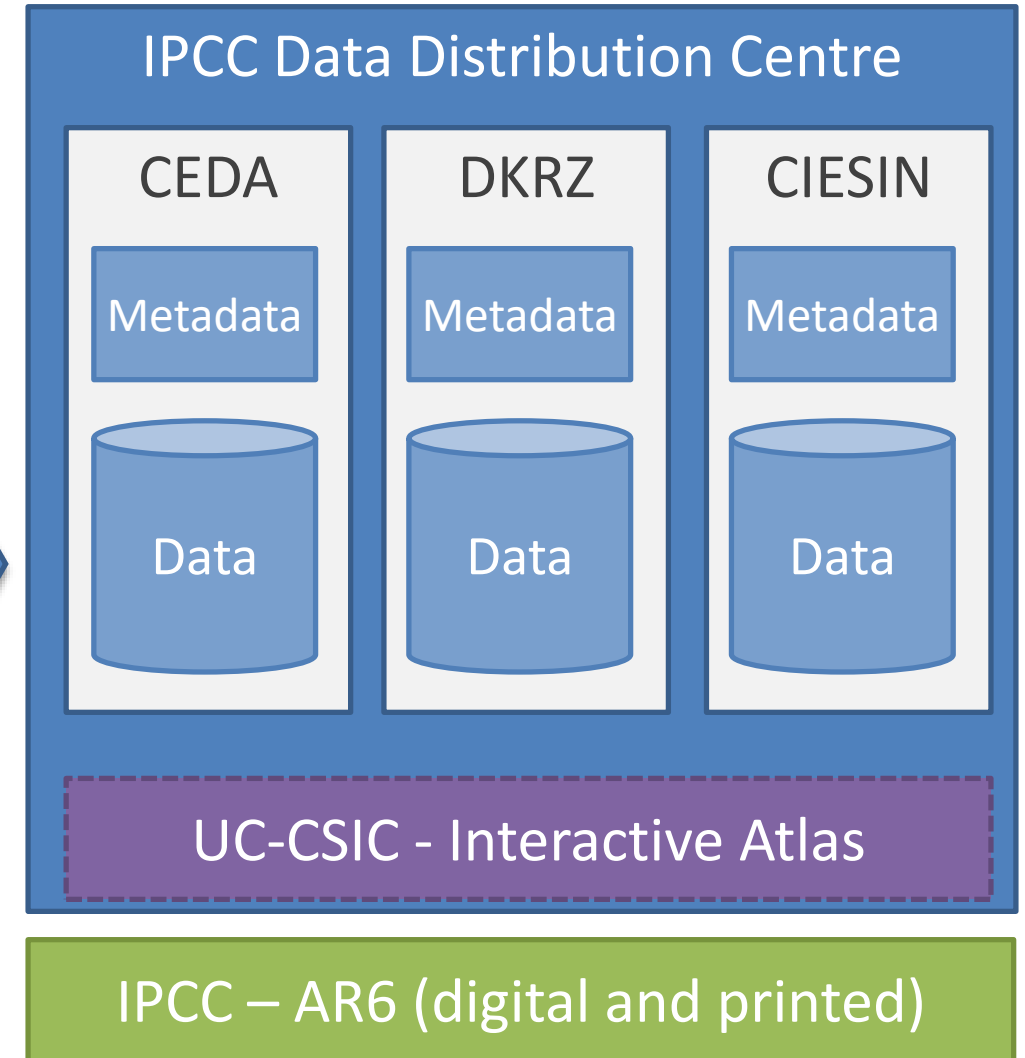
III. Long-Term Data Stewardship

I. FAIR Guidelines and Long-Term Stewardship

Preparing WGI AR6



Long-Term Stewardship



III. Long-Term Data Stewardship: TRUST and the DDC (1)

Data Stewardship in DDC:

DDC Partners are established trustworthy repositories:
DKRZ and CIESIN are WDS Regular Members
and CEDA is a WDS Network Member.

The **jointly managed DDC** adds to these individual TRUST-compliant data center operations:

Transparency:

IPCC's DDC Guidance describes DDC objectives, further materials are provided on or accessible from the DDC webpages.

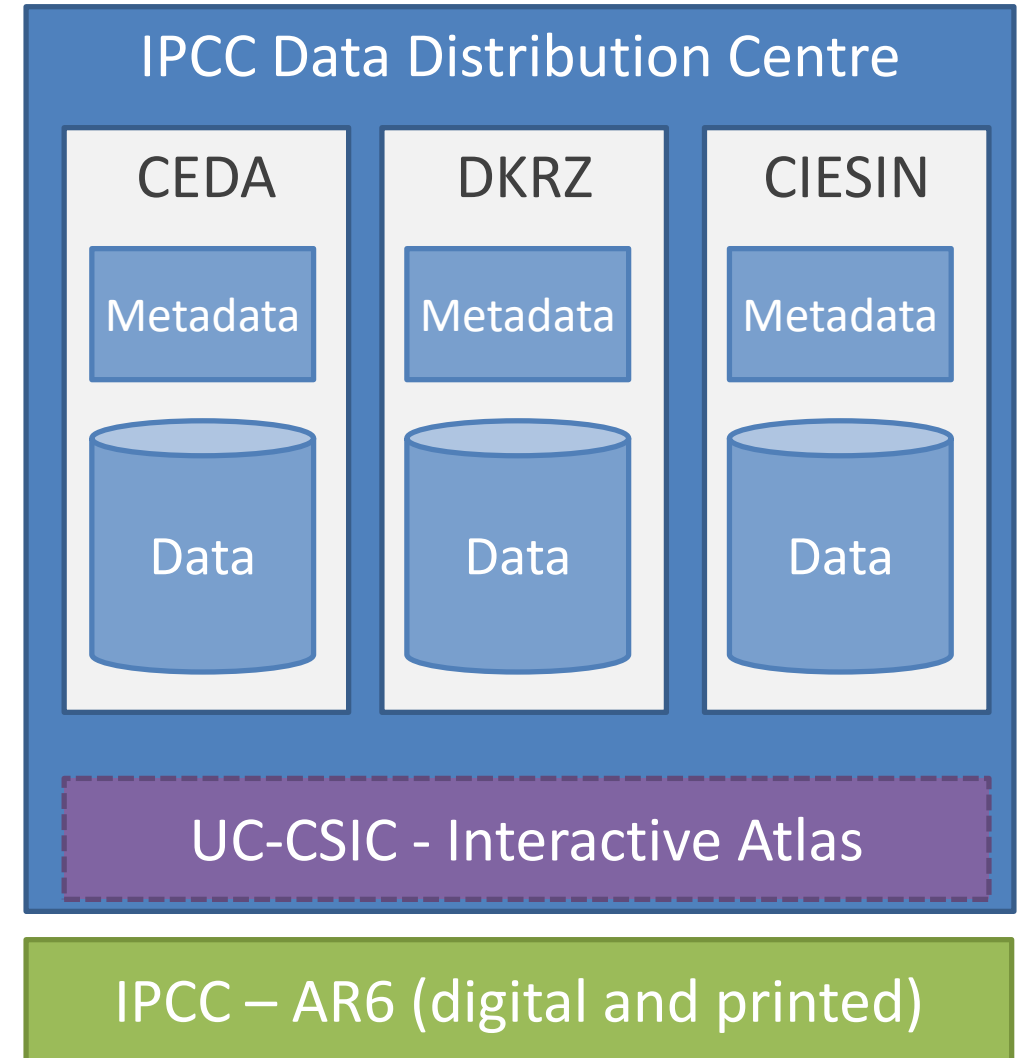
Responsibility:

In the DDC's MoU the responsibilities of the Partners are described.

User Focus:

The collaboration within IPCC ensure the user focus.

Long-Term Stewardship



III. Long-Term Data Stewardship: TRUST and the DDC (2)

Data Stewardship in DDC:

DDC Partners are established trustworthy repositories:
DKRZ and CIESIN are WDS Regular Members
and CEDA is a WDS Network Member.

The **jointly managed DDC** adds to these individual TRUST-compliant data center operations:

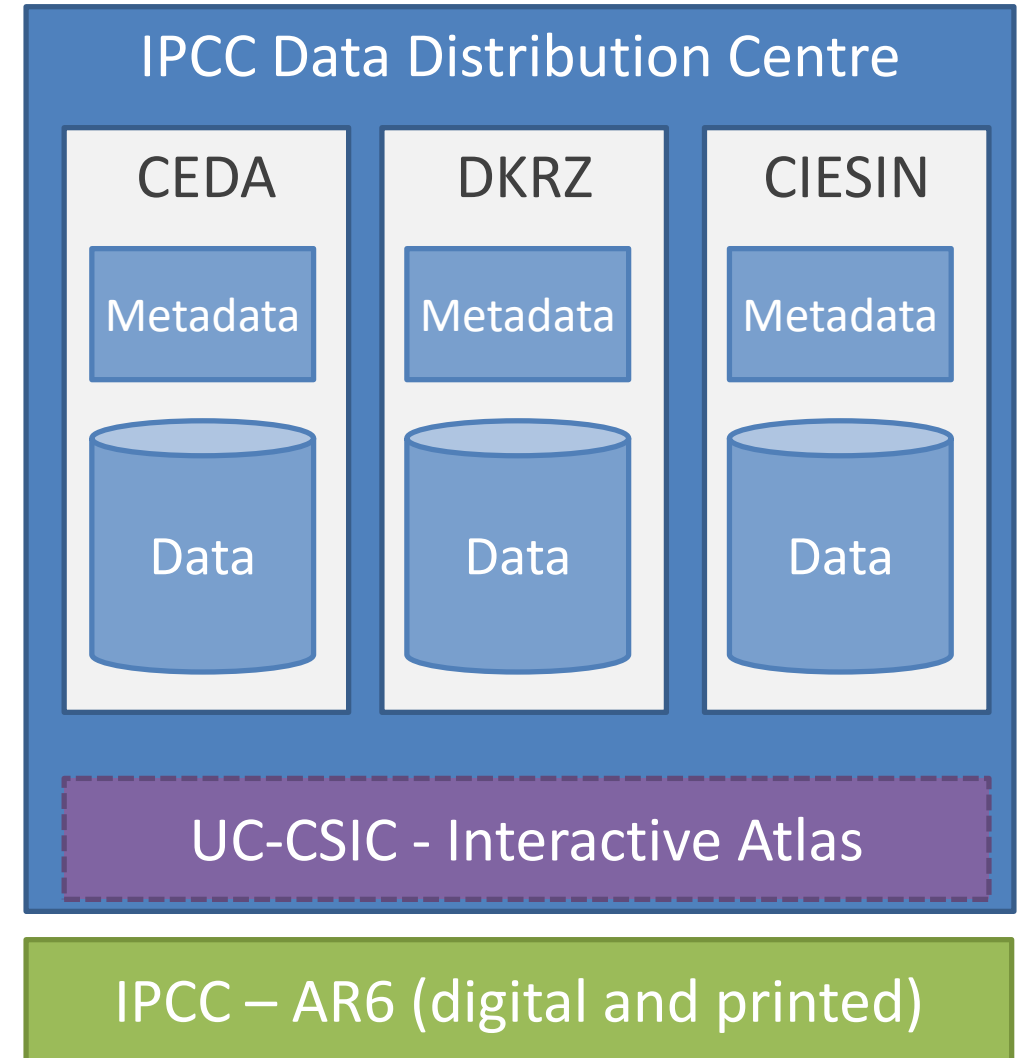
Sustainability:

DDC Partners have committed in a MoU to preserve data holdings on the long-term and sustain **core** services.

Technology:

DDC Partners use institutional infrastructure and staff to maintain their services.

Long-Term Stewardship



Summary

- In AR6, data and software get integrated into the Assessment via the FAIR Guidelines and become IPCC products.
 - Digital products get integrated into established IPCC protocols and procedures.
 - The traceability, transparency, and accessibility of IPCC products is enhanced.
- The presented approach towards FAIR data and software with long-term stewardship in the DDC provides the first step of a long-term effort.

Open Issues / Gap Analysis

1. Exhaustive IPCC data archival:

- Recruit new DDC partners to improve regional data access
- Establish partnerships to bring in new expertise related to data discovery, access and analysis

2. Integration into IPCC procedures and products:

- Implementation of IPCC Errata Protocol for assessed digital products
- Open licensing for IPCC data products
- Two-way integration of AR with DDC data holdings


3. DDC service improvements:

- Single access point for data discovery
- Support for data users in Annex 1 list of UNFCCC by e.g. server-side processing
- Common API for IPCC DDC data

Special thanks to
Jose M. Gutierrez Llorente and Santander MG staff,
IPCC Authors, TG-Data Members, TSU staff, DDC staff,
data and metadata providers, and many more contributors...

ipcc.ch

ipcc-data.org

 orcid.org/0000-0001-6636-4972