

# Facilitating global access to a high-volume flagship climate model dataset: the MPI-M Grand Ensemble experience

EGU 2020, Session ESSI3.6, Live Chat on 7 May 2020

**Karsten Peters**<sup>1</sup>, Michael Botzet<sup>2</sup>, Veronika Gayler<sup>2</sup>, Estefania Montoya Duque<sup>2</sup>, Nicola Maher<sup>2</sup>, Sebastian Milinski<sup>2</sup>, Katharina Berger<sup>1</sup>, Fabian Wachsmann<sup>1</sup>, Laura Suarez-Gutierrez<sup>2</sup>, Dirk Olonscheck<sup>2</sup>, and Hannes Thiemann<sup>1</sup>

<sup>1</sup>Deutsches Klimarechenzentrum (DKRZ), Hamburg, Germany <sup>2</sup>Max-Planck-Institut für Meteorologie, Hamburg, Germany



**DKRZ**

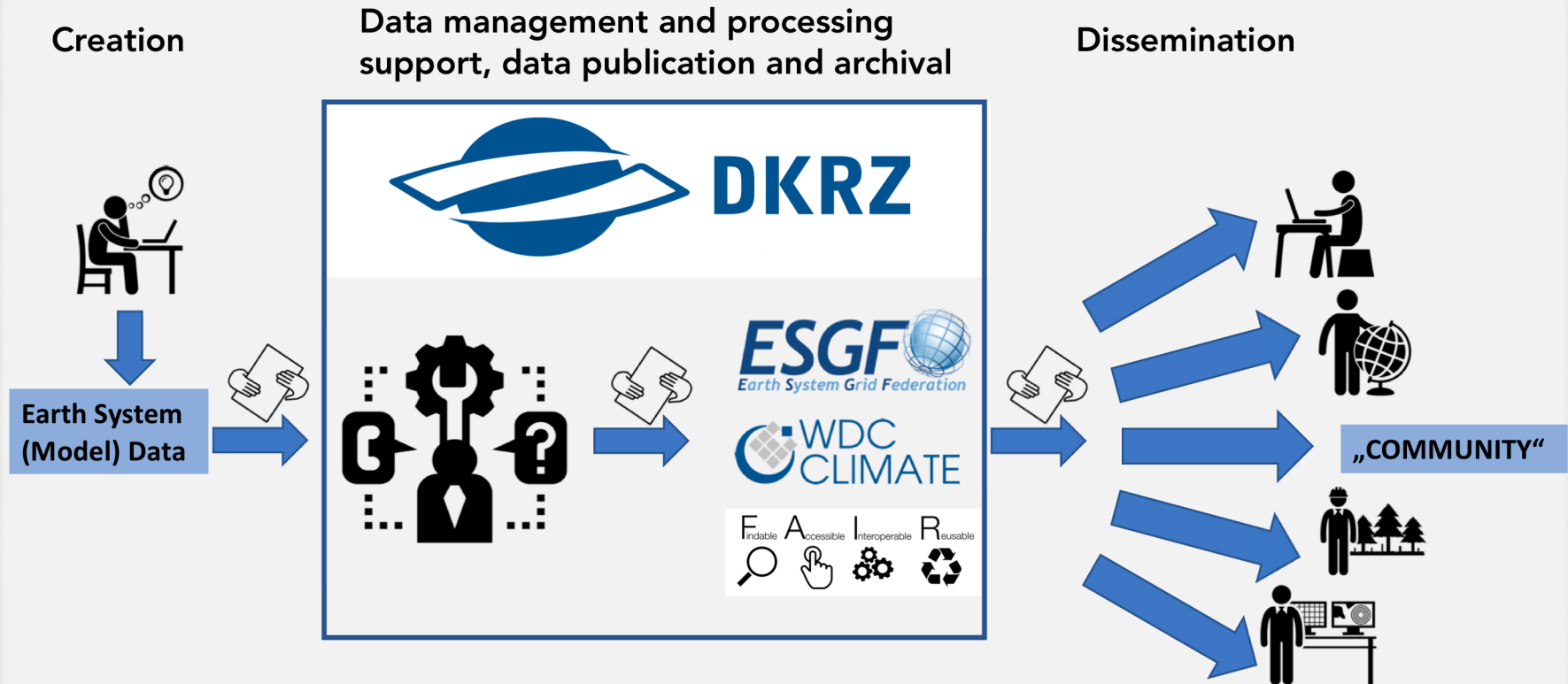


Max-Planck-Institut  
für Meteorologie



# DKRZ Data Management (DM)

Your one-stop-shop for dissemination, publication and preservation of FAIR Earth System (Model) Data

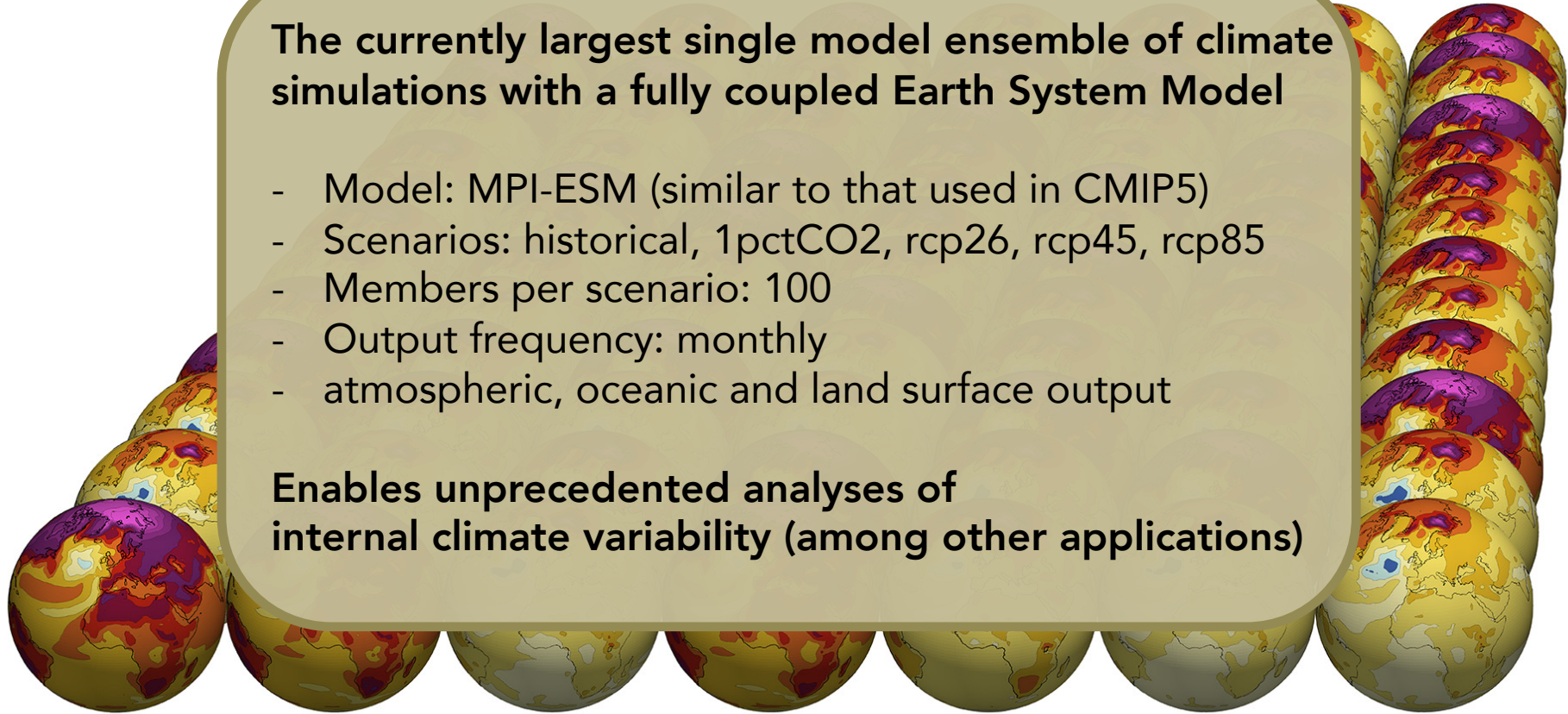


# Use case: MPI-M Grand Ensemble (MPI-GE)

**The currently largest single model ensemble of climate simulations with a fully coupled Earth System Model**

- Model: MPI-ESM (similar to that used in CMIP5)
- Scenarios: historical, 1pctCO<sub>2</sub>, rcp26, rcp45, rcp85
- Members per scenario: 100
- Output frequency: monthly
- atmospheric, oceanic and land surface output

**Enables unprecedented analyses of internal climate variability (among other applications)**



# Use case: MPI-M Grand Ensemble (MPI-GE)

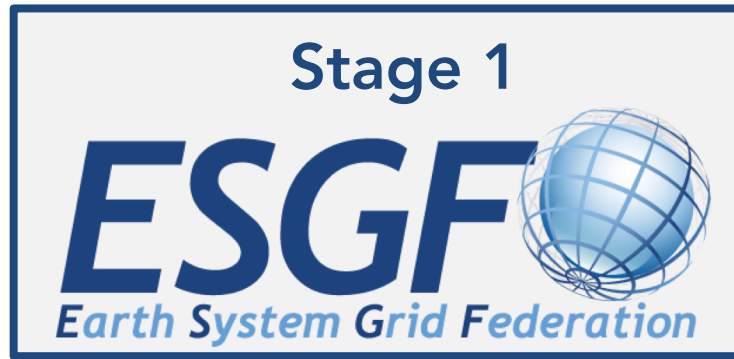


A flagship climate model data set of high community interest produced >4 years ago



# How to get the MPI-GE data to the community?

DKRZ data managers and MP-M scientists **jointly established** a two-stage dissemination/preservation process



- 👍 Effective global dissemination of standardised climate model data
- 👎 No long-term preservation option, no back-up
- 👎 Counts towards user's quota on shared file system

- 👍 FAIR long-term preservation and reusability (>10 years)
- 👍 CoreTrust Seal certified repository
- 👍 Archived data does not count towards any user quota
- 👎 Retrieval via ESGF portal, but slower



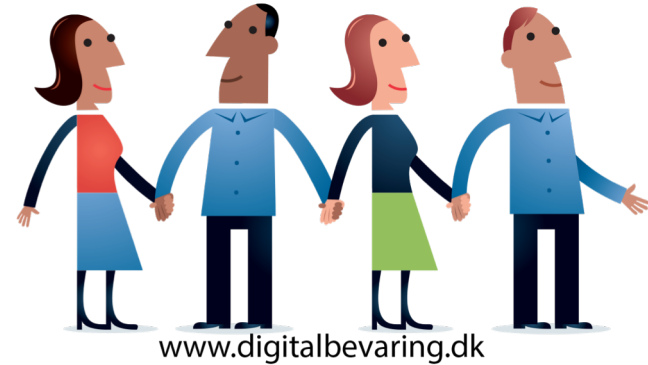
# ESGF publication of the MPI-GE

DKRZ DM experts and MPI-M scientists collaborated closely to accomplish the MPI-GE publication via ESGF

Together,

- a subset of variables was selected
- data and metadata standards were agreed on (close to CMIP5)
- raw data were standardised using software tools provided by DKRZ

MPI-GE data were published by June 2019, currently consisting of >55k files with a volume >70TB.



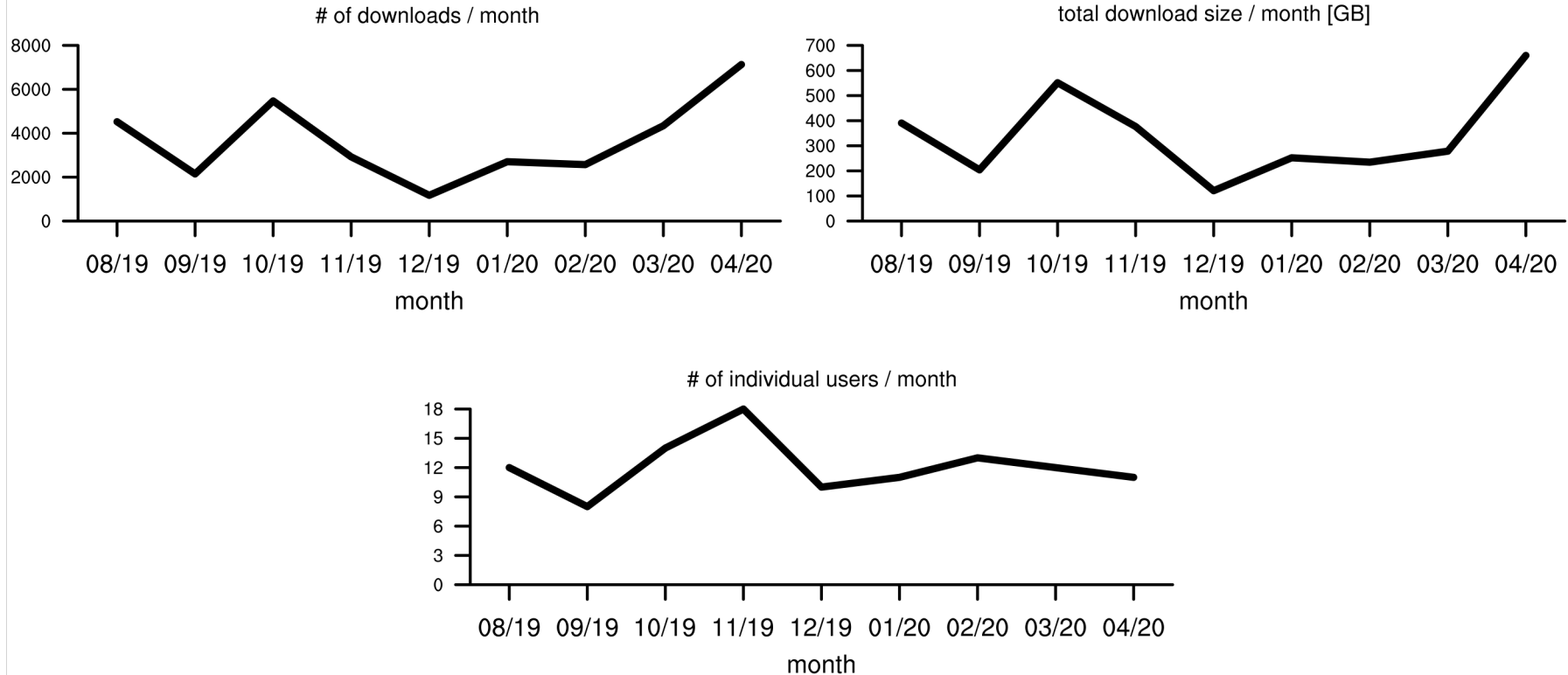
**MPI-GE@ESGF**  
Earth System Grid Federation

<https://esgf-data.dkrz.de/projects/mpi-ge/>

MPI-GE license:  
[CC BY-SA 4.0](#)

# Access statistics – is reuse happening?

## MPI-M Grand Ensemble ESGF Download Statistics



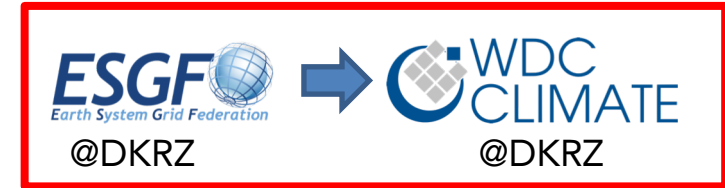
**Access statistics are stable with no decline in interest since publication 😊**

# Next steps – FAIR long-term archival in WDCC

planned for Q3/Q4 2020

Long-term archival is relatively straightforward, because the

- the workflow is part of DKRZ's in-house DM portfolio
- MPI-GE data are already standardised and associated with extensive metadata
- metadata can be harvested from the ESGF-index



Efficient process



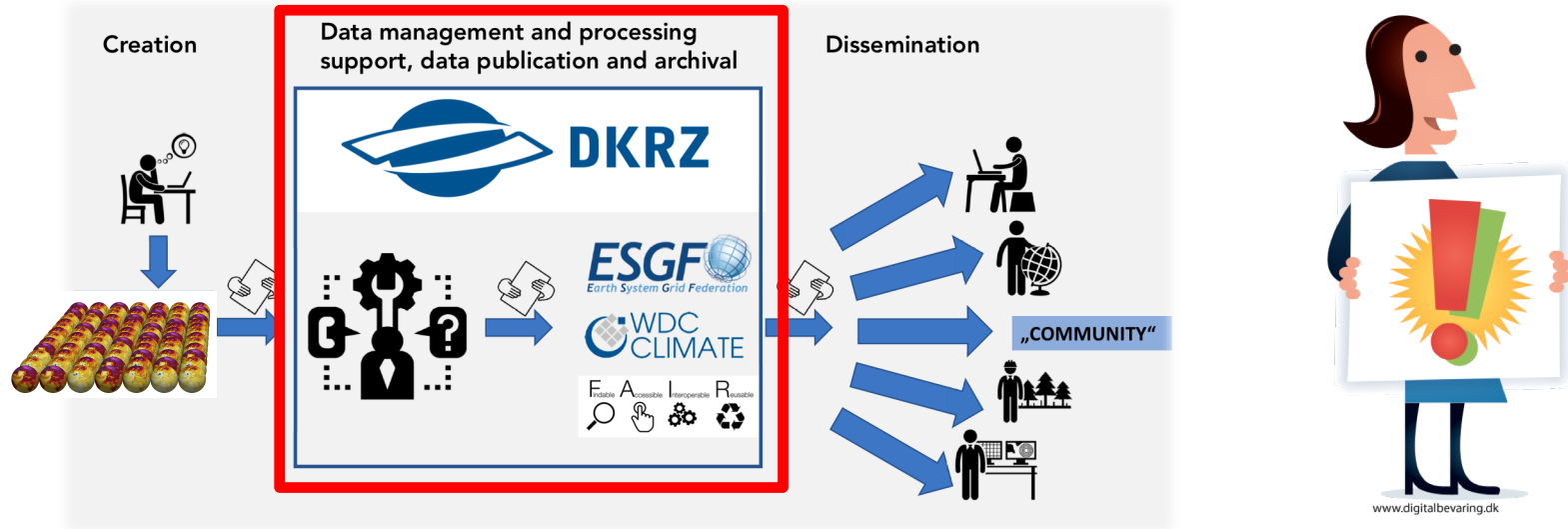
Post-processing completed



Little further input required



# Key points to take home



DKRZ's domain knowledgeable DM-Team supported the entire process of the MPI-GE publication process and will lead the FAIR long-term archival in the WDCC.

Providing all required DM services under the same roof proves invaluable.

Data access statistics are promising – data reuse is happening! 😊

# Further reading

## DKRZ Data Management

- <https://www.dkrz.de/about-en/staff/data-management>
- <https://www.dkrz.de/up/services/data-management>

## MPI-GE

- <https://www.mpimet.mpg.de/en/grand-ensemble/>
- <https://esgf-data.dkrz.de/projects/mpi-ge/>

## ESGF and corresponding services at DKRZ

- <https://www.dkrz.de/up/services/data-management/esgf-services-1>

## WDCC

- <https://cera-www.dkrz.de/WDCC/ui/cersearch/>

## Data processing tools for standardisation developed at DKRZ

- [https://code.mpimet.mpg.de/projects/cdo/wiki/CDO\\_CMOR\\_Operator](https://code.mpimet.mpg.de/projects/cdo/wiki/CDO_CMOR_Operator)