

# Data-centric Science: New challenges for long-term archives and data publishers

cmip6cite.wdc-climate.de

M. Stockhause and M. Lautenschlager  
German Climate Computing Center (DKRZ)

EGU2016-11315 

## Role of Data Long-Term Archives in Connecting Data and Documentations

Data and related documentations are no longer stored in one place but distributed over multiple archives. Often even data for a single project cannot be held at one data center but is disseminated via many data nodes.

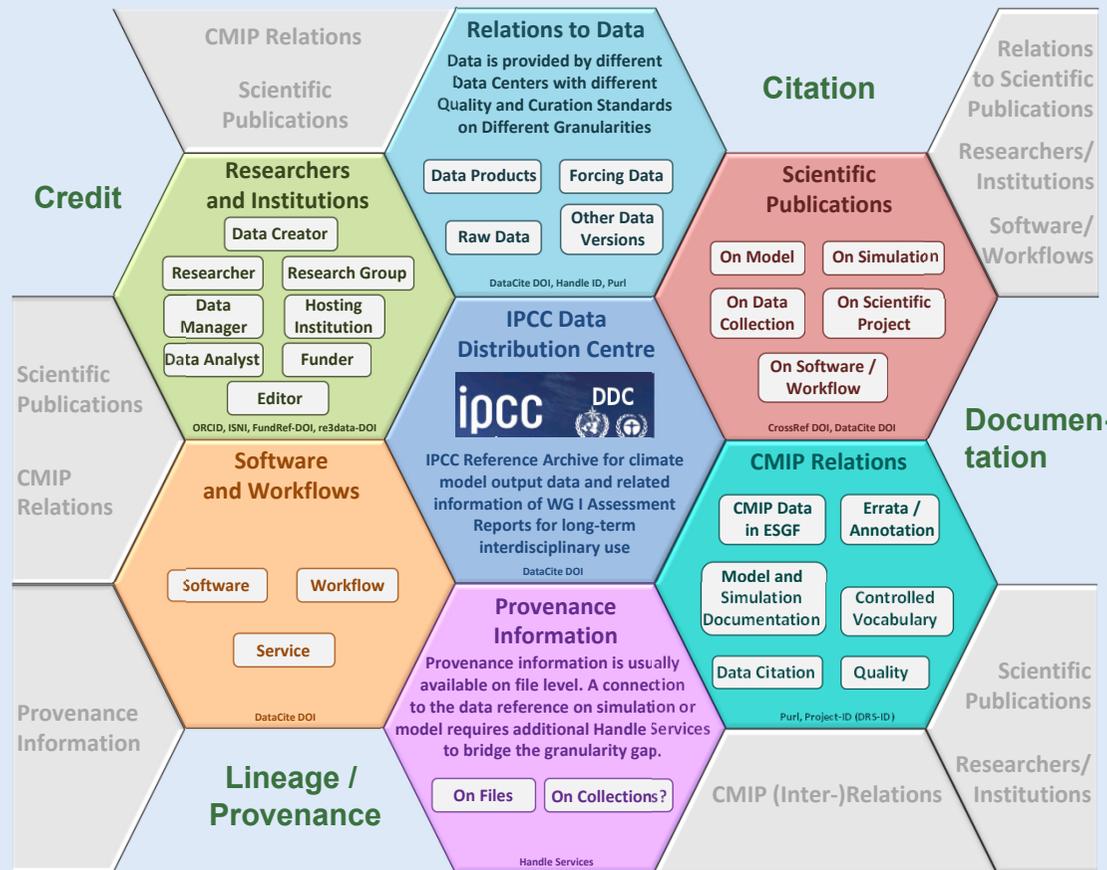
Long-term archives are the place where data and documentations out of various sources with different quality standards and on a wide range of granularities come together for long-term interdisciplinary use.

Some of these data and information out of project resources or new data infrastructures are persistent. However, most of them vanish some time after project termination.

Other than for scientific publications, project data and information are not referenced but usually stored in the long-term archive along with the data.

The aim for a future research environment is to leave data at the external data center. Then established certification systems are required to evaluate data centers, e.g. World Data System (WDS) certificate or Data Seal of Approval (DSA).

## Relations of the IPCC Reference Data Archive to external PID Systems



## Role of Data Long-Term Archives in Scientific Publications

Long-term archives provide persistent data, which is suitable to be cited in scientific publications with DataCite DOIs. The function of data DOIs in reference lists is primarily to give access to the data and its documentation. Secondly, data DOIs serve as entry points to related further information.

Data Publications cite scientific publications, related data and different project resources. They also connect the data to persons and institutes with various roles.

As data citations get more and more incorporated in the scientists' research environment, data citations are requested by scientists or scientific organizations. E.g. WGCM CMIP6 explicitly asked for a citation possibility for CMIP6 data entities.

The main function of data citations is to give credit to other scientists for the data underlying the research.

Apart from the established data citations for stable data, requests to cite dynamic project data have been formulated. In case of CMIP6 and IPCC-AR6 an overall citation concept has been developed (cmip6cite.wdc-climate.de).



stockhause@dkrz.de  
id orcid.org/0000-0001-6636-4972  
www.wdc-climate.de  
www.dkrz.de

