



## **Preservation and Long Term Access of Data at the World Data Centre for Climate**

Frank Toussaint, Peter Drakenberg, Heinke Höck, Stephan Kindermann, Michael Lautenschlager, Hans Luthardt, Martina Stockhause, and Hannes Thiemann

Deutsches Klimarechenzentrum, World Data Centre for Climate, Hamburg, Germany (toussaint@dkrz.de)

Founded in 1987 and owned by several German research facilities, the German Climate Computing Centre (Deutsches Klimarechenzentrum GmbH, DKRZ) is a service centre for earth system researchers. It is responsible for operating a high performance computer and storage system for basic as well as applied research in the field of climatology and related disciplines. Since 1995 it gives online access to Terabytes of climate model data via the internet. Later observational data were added, too. In 2003 the Data Management department was approved as World Data Centre for Climate Data (WDCC) by the International Council for Science (ICSU).

With increasing need for interoperability, the flexibility of all IO processes was enhanced and different standards were implemented. Today several catalogues are openly accessible (e.g., [cera.wdc-climate.de](http://cera.wdc-climate.de)) or in the process of installation and a variety of metadata standards can be handled.

Over the last few years, long term storage of project results has become increasingly important and is today required by many funding agencies. Especially for these data, processes of quality management have been implemented. Furthermore, these quality assured data can be published according to the STD-DOI ([www.std-doi.de](http://www.std-doi.de)) data publication standard. Thus they are assigned a persistent Digital Object Identifier (DOI) to make them citable. A publication in a data journal like, e.g., Earth System Science Data (ESSD, [earth-system-science-data.net](http://earth-system-science-data.net)) is supported by WDCC.

Huge amounts of climate data are stored on tapes and made accessible via the internet. To account for data access of finest granularity, a software system (LOBSTER) was developed and installed which allows for well performing download of files of different sizes, spanning over more than six orders of magnitude from some Kilobytes to many Gigabytes. LOBSTER comprises the management of file packing into container files as well as caching and many aspects of data recovery after a possible file damage. All of these actions are parameterised and tuneable. At the WDCC, this adjustment phase is nearly finished and since some months a LOBSTER pilot version is running. Interested third parties with similar requirement profiles are welcome to take part in its usage and development.

WDCC provides a data archiving service and offers data management consulting for climate data over the whole life time of data to earth system researchers. This includes advice in data storage structure as well as support in metadata layout and creation. Data storage in the long term archive including the necessary pre-processing is offered, too. All data can be made accessible via internet. Here data comprises input quantities like forcing data to feed the model runs, work data during the lifetime of the project and long term archiving of basic and final results.