

THE CEOP MODEL DATA ARCHIVE

AS PART OF THE WORLD DATA CENTRE FOR CLIMATE

Frank Toussaint, Michael Lautenschlager*, Hans Luthardt
World Data Centre for Climate, Hamburg, Germany

1. INTRODUCTION

The CEOP (Coordinated Enhanced Observing Period) data management responsibility is covered by three thematically distinguished archives: The University of Tokyo stores satellite data, the CEOP Data Archive (CDA) at the University Corporation for Atmospheric Research/Joint Office for Science Support (UCAR/JOSS) stores the reference site data and the World Data Centre for Climate (WDCC) archives the model output data for CEOP (www.joss.ucar.edu/ghp/ceopdm/). CEOP is defined for the period from July 2001 until December 2004. Model data from ten different National Weather Prediction Centres (NWP) is collected, harmonised and disseminated by web access (www.mad.zmaw.de/CEOP/ceop_overview.html) by the WDCC.

2. THE WDCC DATA ARCHIVE

The World Data Centre Climate (WDCC, wdcc.dkrz.de) is hosted by the Model & Data Group of the Max-Planck-Institute for Meteorology. Financed by the German



* *Corresponding author address:* M.Lautenschlager, M&D, MPI-Met, Bundesstr.53, 20146 Hamburg, Germany, Lautenschlager@dkrz.de

government, this department uses the computers and mass storage facilities of the German Climate Computing Centre (Deutsches Klimarechenzentrum, DKRZ).

The WDCC provides web access to 120 TB climate data, the total mass storage archive contains nearly 4 PB. Although the majority of the data is model output datasets, some satellite and observational data is accessible as well. The underlying relational database is distributed on five data servers.

3. THE CERA DATA MODEL

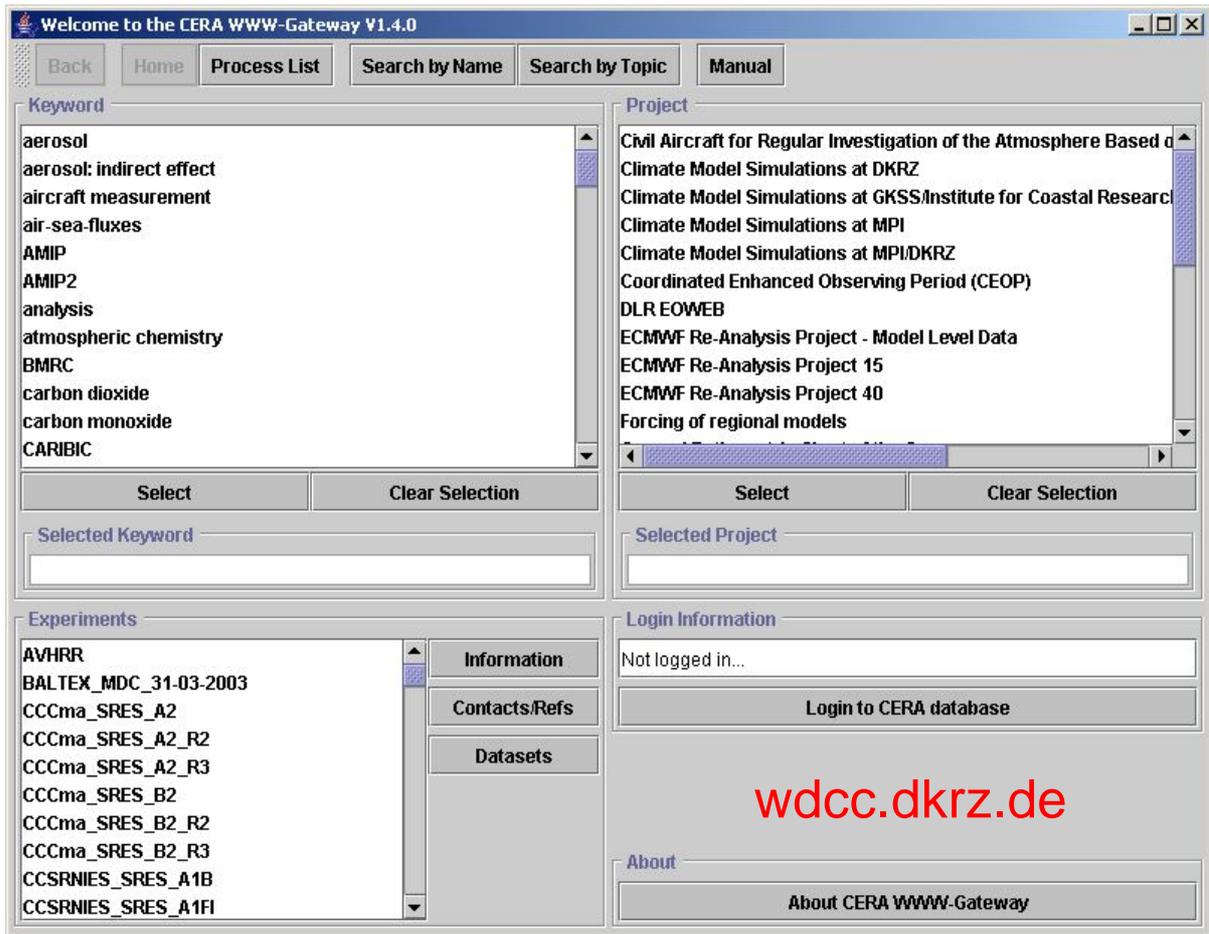
To integrate catalogue data and mass data, the CERA relational data model was designed. Its flexibility allows for very different types of data and metadata.

To facilitate the inclusion of supplements and modules CERA has a simple structure: its tables are block wise grouped and the groups have similar structures. More difficult structures go into CERA Modules, which are table groups attached to the core data model.

In various tables CERA also allows for definable contents (fields, tables, data types), pointers between different entries, and extensible lists of values.

The CERA metadata catalogue not only links to the content of the CERA database. It also can give easy access to other data on the web, such as, e.g., datasets offered by DODs/OPeNDAP servers.

Visit www.pik-potsdam.de/cera for additional information on the CERA data model.



Above: The open access user interface to the CERA metadata catalogue allows for data selection by keyword, project, and topic.

4. THE CEOP DATA

The CEOP model data archive is integrated into the standard CERA archive environment of the World Data Centre for Climate. All meta data will be available as xml files on http request. The model data itself can be retrieved in different formats: grib, ASCII, netCDF, and binary (IEEE), an online facility for geographic selections from the global model data is prototypically ready and will be operational within the next months.

5. THE WEB ACCESS

Presently, the DB web access is mainly by browser and applet. The WDCC gradually transfers the functionalities to Servlets, as they reduce the security and firewall problems. A java command line interface (jblob) for script access is available, too.

The access to the CERA metadata catalogue is open without registration. However, if you want to retrieve bulk data, you will have to check in and apply for a userid and password.

An access control by browser certificates is under development. Especially for distributed data storage this will make the access control much more user friendly. Within a group of trusted data centres, the user will be able to change between the internal websites without further personal identification.