CERA:
Data Structure and User Interface

Frank Toussaint
Michael Lautenschlager

World Data Center for Climate
(M&D/MPI-Met, Hamburg)

NINTH WORKSHOP ON
METEOROLOGICAL OPERATIONAL SYSTEMS
ECMWF, Reading/Berks., 10 - 14 November 2003
Content

• Starting Point: The Requirements

• The CERA Characteristics

• CERA: Additional Features

• The Running System: Data streams

• The Running System: Inclusion of Foreign Sources

• The Graphical User Interface
What are the Requirements of a Meta Database?

Capabilities
- meta information storage
- web enabled
- linked to data sources
- good Performance

Efficiency
- easy to change/extend in data structure
- ... and, of course, in content
- little effort for additional software (tools, etc.)
- easy to adapt to different user interfaces

--- KEEP THINGS SIMPLE ---

--- KEEP THINGS SIMPLE ---
CERA-2 Development

How does CERA-2 meet these requirements?

✓ by flexible structure: data structure not completely fixed
  - table groups form modules and submodules
  - local extensions for local integration and local needs

✓ by flexible content: definable fields and tables
  - definable entry types, contact information types, & various other
    i.e., definable contents of table groups
  - flexible lists of values (LOV): extensible but controlled
  - necessary: LOVs include attributes “acronym” & “description”
  - free text fields of user-defined contents: e.g., for automated data access

✓ by simple structure: blockwise tablegroups
  - CERA-2 Blocks have similar structure
  - more difficult structures go into CERA Blocks
  - some basic rules:
    no <null>s allowed
    negative topics in the LOVs: “none”, “n/a”, “not filled”
The Block Structure

Metadata Entry
This is the central CERA Block, providing information on
- the entry’s title
- type and relation to other entries
- the project the data belong to
- a summary of the entry
- a list of general keywords related to data
- creation and review dates of the metadata

Coverage
Information on the volume of space-time covered by the data

Parameter
Block describes data topic, variable and unit

Spatial Reference
Information on the coordinate system used

Contact
Data related to contact persons and institutes like distributor, investigator, and owner of copyright

Status
Status information like data quality, processing steps, etc.

Reference
Any publication related to the data together with the publication form

Distribution
Distribution information including access restrictions, data format and fees if necessary

Additionally: Modules and Local Extensions
Module DATA_ORGANIZATION (grid structure)
Module DATA_ACCESS (physical storage)
Local extension for specific information on (e.g.)
- data usage
- data access and data administration
Some additional features

- allows for pointers between CERA Entries: directed or symmetrical relations

- various views on CERA Modules for the
  - automatted data access and
  - axes description of multidimensional equidistant gridded data
  - including axes consisting of different equidistant patches

- information on in house data processing can be added to every entry
A CERA Module

CERA Module DATA_ORGANIZATION, Version 1.3

PARAMETER

DATA_Org
data_org_id
data_org_acronym*31
data_org_name*250
data_org_desc*2000

time_id

SPACE

space_id
space_acronym*31
space_name*250
space_desc*2000
x_scale_id
y_scale_id
z_scale_id
point_set_id

POINT_SET

point_set_id
point_set_acronym*31
point_set_name*250
point_set_desc*2000
num_points

POINT_CONNECT

point_set_id
point_id
moment_id
sequence_no

POSITION

position_id
position_name*80
sequence_no

UNIT

UNIT

DIM_TYPE

dim_type_id
dim_type_acronym*31
dim_type_name*250
dim_type_desc*2000

TIME

time_id
time_acronym*31
time_name*250
time_desc*2000
num_points
dim_type_id

TIME_CONNECT

time_id
moment_id
sequence_no

SCALE

scale_id
scale_acronym*31
scale_name*250
scale_desc*2000
num_points
dim_type_id

SCALE_CONNECT

scale_id
position_id
sequence_no

SCALE_PATCHES

scale_id
patch_id
specifier*int

2 ENTITIES

5 LISTS OF VALUES

3 RELATIONS

1 LISTS OF VALUES

2 RELATIONS

Submodule PATCHED_DIMENSION 1.0

to allow for patchwise equidistant gridded dimensions

SCALE_PATCHES

scale_id
patch_id
specifier*int

2 RELATIONS

Version 1.0, 2001-03-24

Version 1.3, 2001-03-22

Ein Institut der Max-Planck-Gesellschaft
An Institute of the Max Planck Society

F. Toussaint (WDCC, Hamburg) / 11.11.03 / 8
Data Streams

Access Client realised as web-based Java Applet.

Middleware layer provides applet and DB connection

DB-Server for catalogue operations and climate data retrieval.

F. Toussaint (WDCC, Hamburg) / 11.11.03 / 9
Inclusion of other Data Sources

Client applet receives foreign data URI from CERA-2 DB

Foreign server provides DB data by http: German Aerospace Centre
Selection via CERA meta data:

- selection of the experiment (=model run)
- display of meta data: experiment, quality, datasets
- selection of the dataset
- display of dataset information
- add datasets to “process list”
- download from tape archive to data server
- dowload to the client
User Needs

- Browser
- Java 1.3: signed Java Applet
- account / password

START: http://cera-www.dkrz.de/CERA/