3. Juli 1998

The CERA-2 Data Model

Michael Lautenschlager (DKRZ)
Frank Toussaint (PIK)
Hannes Thiemann (DKRZ)
Manfred Reinke (AWI)

Three years ago the CERA-1 version of the data model had been published. The present CERA-2 version provides more flexibility, is easier to use and is planned to be maintained for the community in order to support the concept of geographically distributed spatial database systems.

Table of Contents

Table of Contents 1
Introduction 2
Structure of CERA-2 3
Modules 5
Attribute Descriptions of CERA-2 10
Block METADATA ENTRY (10 tables) 10
Block PUBLICATION (5 tables) 12
Block CONTACT (4 tables) 13
Block ATTRIBUTE (4 tables) 15
Block SPATIAL INFORMATION (4 tables) 16
Block DISTRIBUTION (6 tables) 16
Block COVERAGE (6 tables) 18
Block STATUS (3 tables) 19
Module DATA_ORGANIZATION (12 tables) 20
Module DATA_ACCESS (4 tables) 22
Examples 24
Related Information 25
Comparison of CERA-2 with CSDGM (FGDC) and DIF (NASA) 25
References 28
Appendix A: CERA Core Scheme 29
Appendix B: Minimum CERA-2 Entry 30
Appendix C: Plain CERA-2 Entry at PIK 32
Appendix D: Hierarchical CERA-2 Entries at DKRZ 39

Introduction

CERA-1

Three years ago the first version of the CERA (Climate and Environmental data Retrieval and Archive system) metadata model had been developed and published. After three years of practical work and application to a number of data systems a re-designed version has been developed in order to provide more flexibility and to match recent international data description standards.

The first version of the CERA data model was mainly developed with the background of organizing climate model data on one side and on the other side with the aim to match international description standards at that time, namely NASA‘s DIF (Directory Interchange Format; NASA, 1998) and IEEE‘s "Reference Model for Metadata" (Bretherton, 1994). The background of climate model data led to a static two layer structure of the model. CERA-1 (Höck et al., 1995) was subdivided into an "experiment"-layer and into a "dataset"-layer. The structure was hierarchic. The "experiments" yield the overall description of data which are explored in a certain context for example a large climate modelling experiment or a field campaign of measurments. The separate "datasets" of an "experiment" contain physically the time series data of separate variables. The description content of the metadata tables was determined by the user requirements and by the integration of the international DIF standard.

Shortcommings

The application of the CERA-1 model to climate research fields which are different from climate modelling has demonstrated the shortcomings of the model. Basic criticism is the static two layers structure and the assumption that climate data are stored physically in the database system. The CERA-1 data model does not seem to fit the requirements for a data model which can be used by a large number of institutes within the concept of geographically distributed database systems.

Additional Requirements

Climate and environmental data are planned to be stored at different geographical locations at institutions where the capacity is available to maintain the data, to guarantee their quality, to provide scientific and technical advise in data usage and to ensure on-line data dissemination and long-term storage. On the other hand the scientific community wants to access the data in a so called "one-shop-approach". One graphical user interface should be available that opens the access to one virtual database system. consisting of different, interoperating database systems. Beside technical problems there exist mapping problems between the data descriptions and the data models. These mapping problems are minimized if the degree of commonality in the data models is as high as possible. The recommendation of commonality in data models can be supported by providing a model which is flexible, easy to use and maintained for the community. The aim of the CERA-2 development is to match these recommendations.

CERA-2

In order to match the requirements for geographically distributed data storage and to adapt the data model to the FGDC metadata standard the new version of the CERA model has been developed. The basics of the IEEE reference model (Bretherton, 1994) are maintained. Following the IEEE reference model the metadata should support four different interfaces to scientific data management.

- **Browse, Search and Retrieval**

The interface to a database system is related to the needs of human users which interact with the database system by for example a graphical interface. The usage will mainly be interactive which requires a response time that keeps the user engaged. Typical questions to a database system are, what data exist, is it likely to be of use to me, is it really what I want or how do I get it? The underlying datamodel is mainly dictated by the user requirements.

- **Ingest, Quality Assurance, Reprocessing**

  The interface should ensure the logical and scientific integrity of the database. The reprocessing contributes to the data quality which should be accessible within the metadata.

- **Application to Application Transfer**

  The interface places many demands on the robustness and completeness of the descriptions of the data structure. It requires standardization in terms of the datamodel and the storage formats. The interface contributes to scientific dataprocessing within the database system and to realizations of geographically distributed databases

- **Storage and Archive**

  The interface is driven by the need for efficient implementation of search and retrieval with the overall goal of total cost minimization. The user requirements to access the data must be used as a guideline. Efficient climate data access minimizes the requests to slow sequential mass storage systems.

Structure of CERA-2

The layers "experiment" and "dataset" of CERA-1 are merged into one layer named "entry". An option to describe relationships between different entities is included. Relation possibilities are tree like structures, hierarchical orders, networks and no relation between entities.

In order to reduce the complexibity the data model is divided into a core scheme and additionals. The core allows for additional modules (community defined) and local extentions (single user defined). The ER-like diagram of CERA-2 is given in Appendix A.

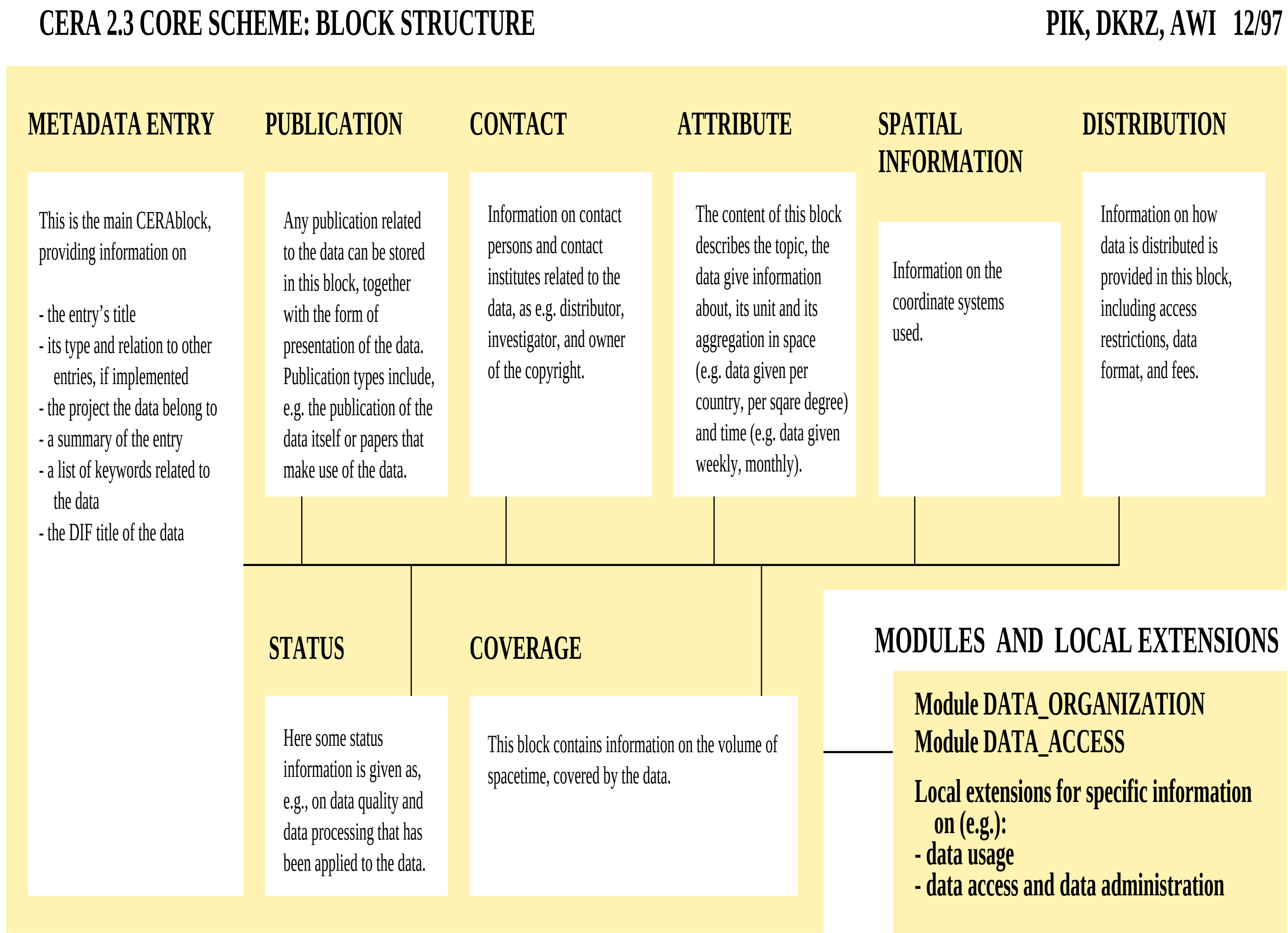

Figure 1. Block structure of CERA-2.3

The core is restricted to meta information which describes data generally and which is common to almost all data in geophysical sciences. The organisation of the block structure as well as the description elements themselves is harmonized with the first levels of the FGDC contents standard (Version 1; FDGC, 1998). The information which is availble from the CERA Core should be sufficient in order to answer the questions:

- What data are stored?
- Are the data useful to my application?
- Who can provide assistance?
- How to get the data?

The information which is contained in the CERA Core is divided into different blocks.

- The **Metadata Entry** is the central information part and provides information on the entity title, its type and its relation to other entities, a summary and a list of keywords which are related to the data.
- The **Publication** block lists any publication which is related to the data. Additionally the form and the type of the publication are included.
- The **Contact** block lists information to various types of contact points like investigator, originator, technical contact or owner of copyright. Person‘s name, address and email is stored so that the personal contact can be established by these metadata.
- The **Attribute** block contains information on topics, disciplines and parameters and variables including units and aggregation in order to identify the data by their keywords.
- The **Status** block provides information on data quality, on applied process steps and on the progress. The latter means whether the data processing is finalized or is still in progress.
- The **Location** block contains information on the temporal and spatial coverage of the data described.
- The **Spatial Information** summarizes the overall information on the coordinate system which is used the store the real data.
- The **Distribution** block delivers information about the data access methods, about the access and the use constraints and about the storage format of the real data.

In the CERA Core only little information is required in order the make the CERA data model applicable for as many institutions as possible and also the amount of mandatory information is minimized in order to make the information content as useful as possible (refere to Sec. “Attribute Descriptions of CERA-2”).

Modules

The CERA Core is open to be completed by **Modules** and by **Local Extensions**. The Modules are common developments of a certain number of institutes reflecting their needs in data management. The Local Extentions contain special requirements of basically one institute. Therefore, the CERA-2 data model shows three degrees of commonality:

1. The CERA Core, which is used exactly the same by all institutes of the CERA users group.
2. The Modules, which reflects additional data management requirements of some institutes. The Modules are common developments of these instituts.
3. The Local Extentsions reflect additional requirements of one institute which are not covered by the core or by the modules.

Presently two moduls have been developed by PIK and DKRZ, the **Data Organization** module (Fig. 2) and the **Data Access** modul (Fig. 3). The Data Organization modul determines the organization of real data in the time and in the three space dimensions. Related to the data organization is the Data Access modul which provides the detailed information about the physical storage of data, e.g. in BLOBs, in the database tables or in files. The information is complete in the sense that each individual number can be located uniquely in space and time so that a data processing can be specified by the these metadata information. This Modules are connected to the core by optional identifiers, DATA_ORG_ID and DATA_ACCESS_ID, in the table PARAMETER of the attributes‘ block (comp. Appendix A).

In module **DATA_ORGANIZATION** information is stored on the form, the time and spatial information is represented in the data set, i.e., how is the data scattered in the four dimensional space-time. The information can be given as weekly data on a three dimensional simple grid. In this case, data points are equally spaced in all four dimensions. On the other hand, measurements on an ascending balloon will result in irregular changes of at least the space coordinates. In between these two extremes any combination of regular and irregular distributed positions can be relevant for describing the data. Additionally, one has to allow for or projections, i.e. when the data has no extension along a certain dimension, e.g. surface data or data valid at one certain moment of time only. For the storage in module DATA_ORGANIZATION the four dimensions x, y, z and t have to be divided into gridlike and pointlike ones. The coordinate values of pointlike dimensions normally are only valid for one certain data point each, while values of gridlike dimensions are related to many points, that differ in at least one of the other coordinate values. The structure of gridlike dimensions is described in tables SCALE and TIME. This structure is defined by a set of positions on each axis, that make up the grid. They are given in tables POSITION and (for the time dimension) MOMENT. The positions can be equidistant, irregular, or in case of a projection just a single point; so can be the dimensions, what is described by the table DIM_TYPE for each dimension. For pointlike dimensions the coordinate values of each point (having a number of coordinates equal to the number of pointlike dimensions to describe) are stored in tables POINT and (for the time dimension) MOMENT, being combined in the POINT_SET. Remaining gridlike dimensions can be described as before.

- Simple Grid

CERA Module DATA_ORGANIZATION, Version 1.2

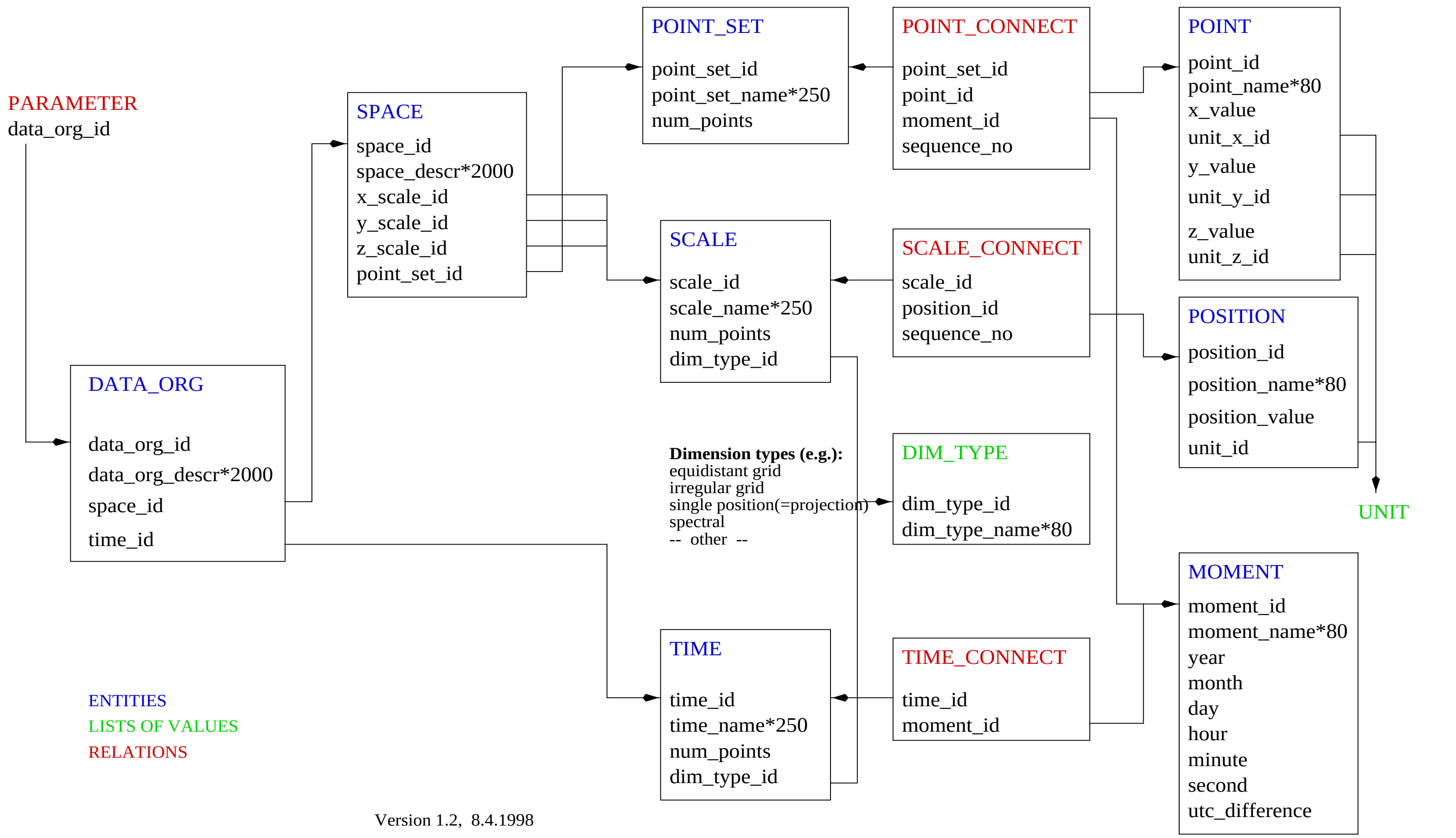

Figure 2. Data Oragnization Modul

For a simple grid in space-time the three _scale_id variables in SPACE and the time_id in DATA_ORG point to the axis' descriptions in tables SCALE and TIME. There the number of points along each axis is defined as well as the type of the grid. In case of irregularly gridded points along a certain dimension, the positions on the axis are given in POSITION or MOMENT. For regular grids, the first and last point have to be entered into POSITION/MOMENT. Their difference divided by the number of points yields the grid's spacing.

- Balloon

  Aperiodic measurements, e.g. on a flying balloon, lead to a set of points in space-time, that is described in POINT_SET. For every point the space coordinates x,y,z (table POINT) as well as the time of measurement (table MOMENT) are given. The sequence_no in POINT_CONNECT is necessary if the points are not ordered by time or if all measurement are taken at one time at different points (projection along time axis). However, the field sequence_no should be filled in anyway.

- Ocean Stations

  Imagine you have for a couple of not moving ocean measurement stations the water temperature given at surface, at 2m, 5m and 10m depths. As the stations are spread over the area, you describe their x,y (LON, LAT) position in POINT, where the altitude is set to 0 (zero). The point set is pointed to by the point_set_id in SPACE, while here x_scale_id and y_scale_id are set to 0 (zero). z_scale_id points to SCALE where the irregular depth grid is described, consisting of the four positions 0, -2, -5, -10, given in POSITION. If, e.g., all stations give data every day, the equidistant grid is given in TIME. Describing a moving station is very similar. Just, that time_id in DATA_ORG is set to 0 (zero) and moment_id in POINT_CONNECT points to the moment, a certain measurement at the point (defined in POINT) has been taken.

Information for automatted data access is provided by the module **DATA_ACCESS**. As the forms of storage can vary widely (e.g., data base, file system, tape library), DATA_ACCESS allows to define the structure of the locally used forms of storage in table ACCESS_STRUCTURE. Here the usage of the four variables storage_name in STORAGE is explained, where storage1_id to storage4_id point to. For automatted data access it is often advisable to know, in what order the single data points are stored with respect to the five dimensions time, space and topic (if more than one parameter is stored). To achieve this, in table REC_STRUCTURE dim1_type gives the type of the fastest varying variable, dim2_type of the second fastest and so on. The corresponding dim?_id points to the related record in one of the tables SCALE, TIME, POINT_SET or PARA_ORDER. Here PARA_ORDER is a locally defined table where the order of the topics can be stored, if necessary. If not all dimensions are used, their _type and _id values are set to 0.

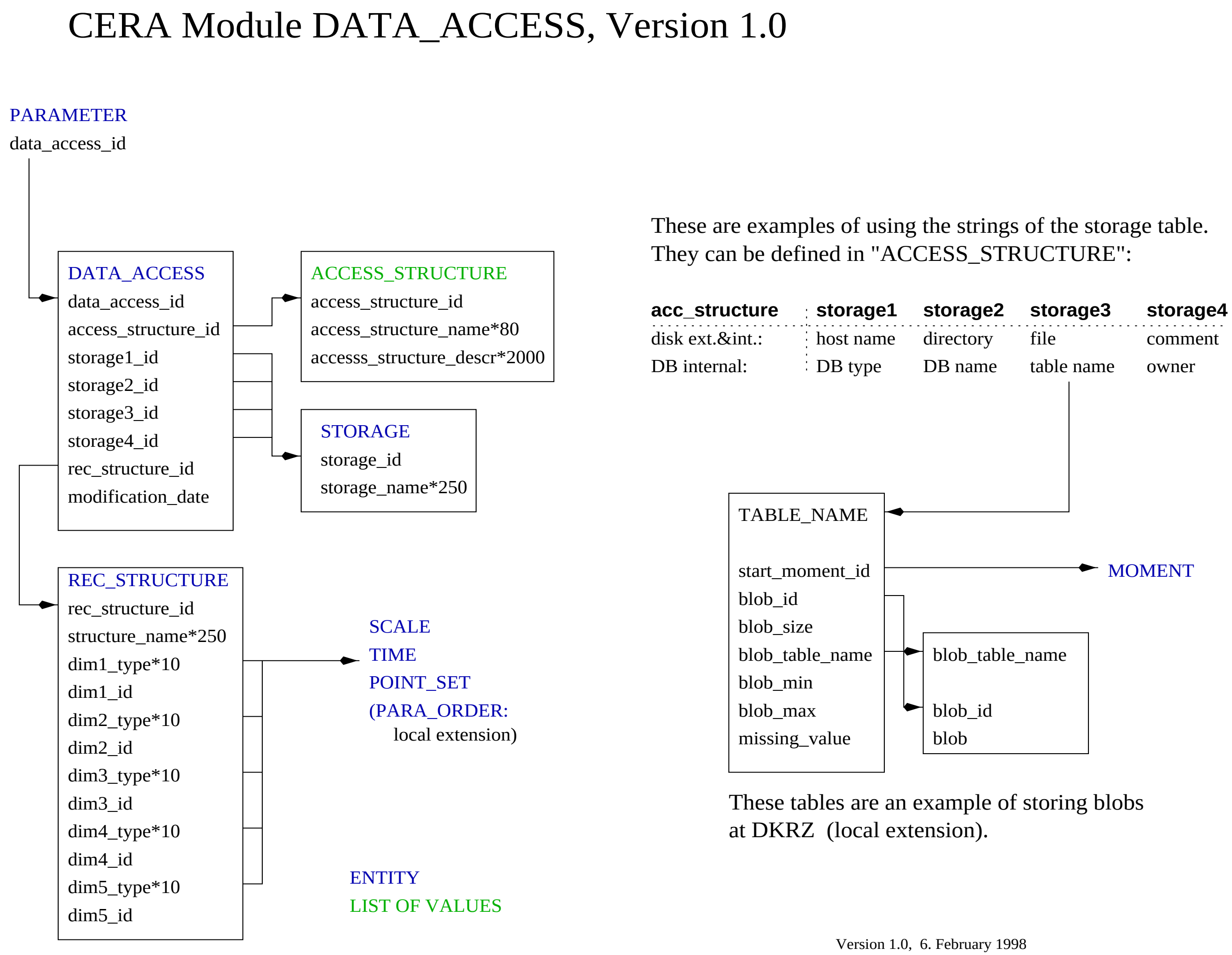

Figure 3. Data Access Modul

Among others, a simple application of module DATA_ACCESS is run at PIK: only host name, directory and file name are stored, together with an optional comment. For this, tables DATA_ACCESS, ACCESS_STRUCTURE and STORAGE are used.

- DKRZ Blob Data

  At DKRZ, binary large objects (blobs) are stored in a DBS. The storage fields contain information on the data base and the name of a locally defined data base table, providing more meta data on the specified blob.

Attribute Descriptions of CERA-2

The basic assumption for all atrributes is that 'NULL' values are not allowed. Defaults are defined for all attributes. The most unspecific ones are 'not filled' and '0' for keys. Therefore the 'NULL' value can be used to proof the correctness of the procedure to fill in the entries into the data model. If a 'NULL' value occurs an error must have appeared.

The minimal information required in CERA-2 provides information related to what is described and where is it valid. Each CERA-2 entry must contain at least the following attributes:

- Name of entry (ENTRY.entry_name),
- Keyword (GENERAL_KEY.general_key),
- Time coverage (TEMPORAL_COVERAGE.start_year, TEMPORAL_COVERAGE.stop_year),
- Space coverage (SPATIAL_COVERAGE.min_lat, SPATIAL_COVERAGEmax_lat, SPATIAL_COVERAGE.min_lon, SPATIAL_COVERAGEmax_lon).

The following sections provide the CERA-2 datamodel. Tables are divided into the CERA Core Blocks and the Modules, subdivided into “entities” (except LOVs), “lists of values” (LOVs) and “relations” and subsubordered alphabetically. The required attributes are indicated by an asterisk (" * “).

Block METADATA ENTRY (10 tables)

Table DIF_ENTRY

entry_id : (integer) Foreign key to ENTRY.

dif_title : (char*160) Title of the DIF should convey DIF content i.e. instrument, investigator, mission, parameters measured.

dif_name : (char*31) Unique identifier of the DIF for the Dataset, a collection of Datasets or part of a Dataset.

Table ENTRY (aliased as ENTRY1)

entry_id : (integer) Primary key.

entry_name : (char*160) * Entry name uniquely identifying this meta data record.

entry_acronym : (char*31) Acronym of the entry_name, uniquely identifying the entry.

entry_type_id : (integer) Foreign key to ENTRY_TYPE.

summary_id : (integer) Foreign key to SUMMARY.

quality_id : (integer) Foreign key to QUALITY.

progress_id : (integer) Foreign key to PROGRESS.

creation_date : (date) Date of data creation.

review_date : (date) Date of the latest review for accuracy of scientific content of the data.

revision_date : (date) Date of last update of data.

future_review_date : (date) Future time at which the meta data should be reviewed for accuracy of scientific or technical content.

Table SUMMARY

summary_id : (integer) Primary key.

summary : (char*2000) Contains an abstract and additional information not contained in other fields.

Table CONNECT_TYPE

connect_type_id : (integer) Primary key.

connect_type : (char*31) Type of the connect between two entries (e.g. hierarchical).

connect_type_descr : (char*2000) Description of the connect_type.

basic set of values : hierarchical / the two records are bound hierarchically.

Table ENTRY_TYPE

entry_type_id : (integer) Primary key.

entry_type : (char*31) If entry records have different types: Type of entry.

entry_type_descr : (char*2000) If entry records have different types: description of type of entry.

basic set of values : normal / the entries are all of the same type.

Table GENERAL_KEY

general_key_id : (integer) Primary key.

general_key : (char*31) * Keyword list: any words or phrases used to further describe the entry.

basic set of values : Any keywords that are appropriate for the field the institute deals with.

Table PROJECT

project_id : (integer) Primary key.

project_acronym : (char*31) Abbreviated version of project_name.

project_descr : (char*2000) Description of the Project.

project_name : (char*160) Scientific endeavour encompassing data from a number of data sources or inhose project, to which the data belongs.

basic set of values : The projects, run in the specific institute can serve as a basic set of values.

Table CAMPAIGN

entry_id : (integer) Foreign key to ENTRY.

project_id : (integer) Foreign key to PROJECT.

Table ENTRY_CONNECT

gen_entry_id : (integer) Foreign key to ENTRY.

spec_entry_id : (integer) Foreign key to ENTRY.

connect_type_id : (integer) Foreign key to CONNECT_TYPE.

Table KEY_CONNECT

entry_id : (integer) Foreign key to ENTRY.

general_key_id : (integer) Foreign key to GENERQAL_KEY.

Block PUBLICATION (5 tables)

Table CITATION

citation_id : (integer) Primary key.

title : (char*2000) Title of a publication related to the entry.

authors : (char*2000) List of the authors as given in the citation.

publication : (char*80) Name of the publication (e.g. of the journal, entology, editorial).

publisher : (char*80) Name of the publisher.

editor : (char*80) Name of the editor(s).

publication_date : (date) Date of the publication. If day and/or month unknown set to 1.

country : (char*80) Name of the country (address).

state : (char*80) Name of the state (address).

place : (char*80) Name of the place (address).

edition : (char*80) Name of the edition.

access_spec : (char*160) Specification (number) of the citation_type (the URL, the ISBN, etc.).

presentation_id : (integer) Foreign key to PRESENTATION.

citation_type_id : (integer) Foreign key CITATION_TYPE.

Table CITATION_TYPE

citation_type_id : (integer) Primary key.

citation_type : (char*80) Type of the citation (e.g. book, journal, web publication etc.).

citation_type_descr : (char*2000) Description of citation_type.

basic set of values : other; www publication; other internet access; CD publication; book; journal; unreviewed publication.

Table PRESENTATION

presentation_id : (integer) Primary key.

presentation_descr : (char*2000) Type of the data presentation, e.g. graph, picture, table...

basic set of values : atlas; diagram; globe; image: remote sensing; image: view/aerial photograph; map: political; map: thematic; map; numerical data; profile; section.

Table REF_TYPE

ref_type_id : (integer) Primary key.

ref_type_descr : (char*2000) Description of the relation, the publication has to the entry.

basic set of values : publication provides data; publication uses data.

Table REFERENCE

entry_id : (integer) Foreign key to ENTRY.

citation_id : (integer) Foreign key to CITATION.

ref_type_id : (integer) Foreign key to REF_TYPE.

Block CONTACT (4 tables)

Table INSTITUTE (aliased as INSTITUTE1)

institute_id : (integer) Primary key.

institute_name : (char*80) Name of the institute. Use the original name given by the institute to itself (if translation is necessary, translate to english).

institute_acronym : (char*31) Abbreviated version of the institute name.

department_name : (char*160) Name of the department or institute sub-division. Use multiple records for different departments.

department_acronym : (char*31) Abbreviated version of the department name.

country : (char*80) Full name of the country (address).

state_or_province : (char*80) Name of the state or province (address).

place : (char*80) Name of the place/city (address).

street : (char*80) Name of the street (address).

street_postal_code : (char*10) Postal code of the street address.

pobox : (char*80) Pobox (address).

pobox_postal_code : (char*10) Postal code of the Post box address (if different from street_postal_code).

url : (char*80) Universal Resource Locator of the institute's homepage.

additional_info : (char*400) Additional address information, or other information about the institute.

Table PERSON

person_id : (integer) Primary key.

first_name : (char*80) First name.

second_name : (char*80) Second name.

last_name : (char*80) Last name. Instead of a person's name a position or office may be given (e.g. secretary, central office).

title : (char*31) Title of the person.

institute_id : (integer) Foreign key to INSTITUTE.

telephone : (char*80) Phone.

fax : (char*80) FAX.

url : (char*80) Universal Resource Locator of the person's home page.

email : (char*80) Email address.

Table CONTACT_TYPE

contact_type_id : (integer) Primary key.

contact_type : (char*80) Type of the contact (investigator, technical info, copyright etc.).

basic set of values : investigator; originator; distributor; technical; metadata; copyright; owner; inhouse.

Table CONTACT

entry_id : (integer) Foreign key to ENTRY.

instiute_id : (integer) Foreign key to INSTITUTE.

person_id : (integer) Foreign key to PERSON.

contact_type_id : (integer) Foreign key to CONTACT_TYPE.

Block ATTRIBUTE (4 tables)

Table AGGREGATION

aggregation_id : (integer) Primary key.

aggregation_descr : (char*80) Description of the aggregation of the meeasured TOPIC, e.g. monthly mean, daily maximum.

basic set of values : none; any combination of: annual, biannual, bimonthly, daily, hourly, monthly, quarterly, weekly ---- interval, max., mean, min., scatter, std.dev., sum.

Table TOPIC (aliased as TOPIC1, TOPIC2, TOPIC3)

topic_id : (integer) Primary key.

topic_name : (char*80) Kinds of measurements/physical quantities represented by the data. A keyword list exists (e.g. sea level).

topic_acronym : (char*31) Abbreviated version of topic_name.

topic_descr : (char*2000) Description of the topic.

topic_pointer : (integer) topic_id of the preceeding topic in the hierarchical structure of the table TOPIC.

topic_level : (integer*2) Level of topic in the hierachical structure of this table.

basic set of values : The list used at PIK and DKRZ can be found via the CERA Central Page (http://www.PIK-Potsdam.de/dept/dc/e/sdm/cera/).

Table UNIT (aliased as UNIT1, UNIT2)

unit_id : (integer) Primary key.

unit_name : (char*80) Unit of the physical quantity found in table TOPIC.

unit_acronym : (char*31) Abbreviated version of unit_name.

unit_descr : (char*2000) Definition of units in measured physical quantities.

basic set of values : Any unit necessary for the data the institute deals with.

Table PARAMETER

entry_id : (integer) Foreign key to ENTRY.

topic_id : (integer) Foreign key to TOPIC.

unit_id : (integer) Foreign key to UNIT.

aggregation_id : (integer) Foreign key to AGGREGATION.

spatial_data_org_id : (integer) Foreign key to SPATIAL_DATA_ORG.

Block SPATIAL INFORMATION (4 tables)

Table HOR_COORD_SYS

hor_coord_sys_id : (integer) Primary key.

sys_descr : (char*2000) Description of the horizontal coordinate system(s).

basic set of values : LON and LAT: geographic; LON and LAT: planar; local; LON: geographic, LAT: spectral.

Table SPATIAL_DATA_ORG

spatial_dat_org_id : (integer) Primary key.

reference_method : (char*80) Form by that the data references the space (e.g. point, vector, raster, spectral...).

basic set of values : indirect spatial reference (e.g. by names of types of geographic features); point(s), non gridded; vector; points: gridded; LON:spectral/LAT:gridded.

Table VER_COORD_SYS

ver_coord_sys_id : (integer) Primary key.

sys_descr : (char*2000) Description of the vertical coordinate system(s).

basic set of values : no vertical extension; altitude: length; altitude: pressure; altitude and depth: length; altitude: pressure, depth: length; depth: length.

Table SPATIAL_REFERENCE

entry_id : (integer) Foreign key to ENTRY.

hor_coord_sys_id : (integer) Foreign key to HOR_COORD_SYS.

ver_coord_sys_id : (integer) Foreign key to VER_COORD_SYS.

Block DISTRIBUTION (6 tables)

Table ACCESS_CONSTRAINT

access_constraint_id : (integer) Primary key.

constraint_descr : (char*2000) Description of the permits of the data: Permit to see the data, permit to copy for own purpose.

basic set of values : unrestricted; institute only; work group only; owner only (supply owner in CONTACT in this case); inhouse unrestricted, externally only in collaborations.

Table ACCESS_TYPE

access_type_id : (integer) Primary key.

access_acronym : (char*31) Short version of access_descr.

access_descr : (char*2000) Description of physical access to the data.

basic set of values : remote, 3 1/2” diskette; 5 1/4” diskette; audiocassette; cartridge tape; CD-ROM; chart; computer programme; disc: DB Table; disc: DB blob pointer; disc: normal file; disc: other; electronic bulletin board; electronic mail system; filmstrip; magnetic tape; microfiche; microfilm; online: other; offline: other; paper; physical model; stable-base material; transparency; videocassette; videodisc; videotape.

Table FEES

fees_id : (integer) Primary key.

fees_acronym : (char*31) Short version of fees_descr.

fees_descr : (char*2000) Description of the fees charged for the data.

basic set of values : none; UIG GebVO; like UIG GebVO; low price (<100 Eu); medium price (100-500 Eu); high price (>500 Eu).

Table FORMAT

format_id : (integer) Primary key.

format_acronym : (char*31) Entry format, e.g. GIF, GRIB.

format_descr : (char*2000) Explanation of the format_acronym.

basic set of values : The list used at PIK and DKRZ can be found via the CERA Central Page (http://www.PIK-Potsdam.DE/dept/dc/e/sdm/cera/).

Table USE_CONSTRAINT

use_constraint_id : (integer) Primary key.

constraint_descr : (char*2000) Description of the permits of the data: Permit to use and publish the data.

basic set of values : unrestricted; institute only; work group only; owner only (supply owner in CONTACT in this case); inhouse unrestricted, externally only in collaborations.

Table DISTRIBUTION (aliased as DISTRIBUTION1)

entry_id : (integer) Foreign key to ENTRY.

access_type_id : (integer) Foreign key to ACCESS_TYPE.

data_size : (integer*8) Size of data set in Bytes.

format_id : (integer) Foreign key to FORMAT.

fees_id : (integer) Foreign key to FEES.

access_constraint_id : (integer) Foreign key to ACCESS_CONSTRAINT.

use_constraint_id : (integer) Foreign key to USE_CONSTRAINT.

Block COVERAGE (6 tables)

Table SPATIAL_COVERAGE

spatial_coverage_id : (integer) Primary key.

min_lat : (float*4) * Southernmost latitude of the area covered by the data (south of Equator negative).

max_lat : (float*4) * Northernmost latitude of the area covered by the data (south of Equator negative).

min_lon : (float*4) * Westernmost longitude of the area covered by the data (west of Greenwich negative).

max_lon : (float*4) * Easternmost longitude of the area covered by the data (west of Greenwich negative).

min_altitude : (float*4) Lowest depth/altitude of the vertical coverage of the data.

min_alt_unit_id : (integer) Foreign key to UNIT.

max_altitude : (float*4) Highest depth/altitude of the vertical coverage of the data.

max_alt_unit_id : (integer) Foreign key to UNIT.

Table TEMPORAL_COVERAGE

temporal_coverage_id : (integer) Primary key.

start_year : (integer*8) * Start year of the time covered by the data (YYYYYY...).

start_month : (integer*1) Start month (MM), NULL if unknown.

start_day : (integer*1) Start day (DD), NULL if unknown.

stop_year : (integer*8) * End year of the time covered by the data (YYYYYY...), NULL if unknown.

stop_month : (integer*1) End month (MM), NULL if unknown.

stop_day : (integer*1) End day (DD), NULL if unknown.

currentness_ref_id : (integer) Foreign key to CURRENTNESS_REF.

Table CURRENTNESS_REF

currentness_ref_id : (integer) Primary key.

currentness_ref_descr : (char*2000) Basis, for determining the time period of content information (meta data: temporal coverage).

basic set of values : ground condition; publication date; arbitrary numbered years; calendrical; climate model time; unchecked data; time before present (1950=0).

Table LOCATION (aliased as LOCATION1, LOCATION2, LOCATION3)

location_id : (integer) Primary key.

location_name : (char*80) Names of places/regions.

location_level : (integer*2) Level of location_id in the hierachical structure of this table.

ref_date : (char*15) A date or year, when the information of this record was valid.

min_lat : (float*4) Southernmost latitude of the location area (south of Equator negative).

max_lat : (float*4) Northernmost latitude of the location area (south of Equator negative).

min_lon : (float*4) Westernmost longitude of the location area (west of Greenwich negative).

max_lon : (float*4) Easternmost longitude of the location area (west of Greenwich negative).

basic set of values : The list used at PIK can be found via the CERA Central Page (http://www.PIK-Potsdam.DE/dept/dc/e/sdm/cera/).

Table COVERAGE

entry_id : (integer) Foreign key to ENTRY.

temporal_coverage_id : (integer) Foreign key to TEMPORAL_COVERAGE.

spatial_coverage_id : (integer) Foreign key to SPATIAL_COVERAGE.

Table LOCATION_CONNECT

entry_id : (integer) Foreign key to ENTRY.

location_id : (integer) Foreign key to LOCATION.

Block STATUS (3 tables)

Table QUALITY

quality_id : (integer) Primary key.

accuracy_report : (char*2000) Quantitative accuracy of the attributes described by the data (e.g. information on error bars).

consistency_report : (char*2000) An explanation of the fidelity of the relationships in the data set and the tests used for their proof.

completeness_report : (char*2000) Information about omissions, selection criteria, generalization, definitions used and other rules to derive the data set.

horizontal_acc_report : (char*2000) Quantitative horizontal position accuracy.

vertical_acc_report : (char*2000) Quantitative vertical position accuracy.

Table PROGRESS

progress_id : (integer) Primary key.

progress_descr : (char*31) The state of the data set in the institute (e.g. complete, in work, planned).

basic set of values : remote; complete; in work; planned; complete, original; complete, processed.

Table PROCESS_STEP

entry_id : (integer) Foreign key to ENTRY.

process_descr : (char*2000) If data has been processed/changed (not just format changes): description of processing.

process_date : (date) If data has been processed/changed (not just format changes): date of processing.

person_id : (integer) Foreign key to PERSON.

Module DATA_ORGANIZATION (12 tables)

additionally into PARAMETER

data_org_id : (integer) Foreign key to DATA_ORG.

Table DATA_ORG

data_org_id : (integer) Primary key.

data_org_descr : (char*2000) Description of the data organization in space and time, described by the information in this record.

space_id : (integer) Foreign key to SPACE.

time_id : (integer) Foreign key to TIME.

Table SPACE

space_id : (integer) Primary key.

space_descr : (char*2000) Description of the spatial data organization, described by the information in this record.

x_scale_id : (integer) Foreign key to SCALE, pointing to the scale of coordinate x.

y_scale_id : (integer) Foreign key to SCALE, pointing to the scale of coordinate y.

z_scale_id : (integer) Foreign key to SCALE, pointing to the scale of coordinate z.

point_set_id : (integer) Foreign key to POINT_SET.

Table POINT_SET

data_org_id : (integer) Primary key.

point_set_name : (char*250) Name of the group of points.

num_points : (integer*4) Number of points in the point set.

Table POINT

point_id : (integer) Primary key.

point_name : (char*80) Name of the point.

x_value : (float*4) x position of the point, longitude if applicable. East positive.

unit_x_id : (integer) Foreign key to UNIT.

y_value : (float*4) y position of the point, latitude if applicable. North positive.

unit_y_id : (integer) Foreign key to UNIT.

z_value : (float*4) z position of the point, altitude if applicable. Depth negative.

unit_z_id : (integer) Foreign key to UNIT.

Table SCALE (aliased as SCALE1, SCALE2)

scale_id : (integer) Primary key.

scale_name : (char*250) Name or explication of the scale.

num_points : (integer*4) Number of points defining the scale in this dimension.

dim_type_id : (integer) Foreign key to DIM_TYPE.

Table POSITION

position_id : (integer) Primary key.

position_name : (char*80) Name or explication of this position on the axis.

position_value : (float*4) Coordinate value of this position on the axis.

unit_id : (integer) Foreign key to UNIT.

Table TIME

time_id : (integer) Primary key.

time_name : (char*250) Name or explication of the time scale.

num_points : (integer*4) Number of points defining the time scale.

dim_type_id : (integer) Foreign key to DIM_TYPE.

Table MOMENT

moment_id : (integer) Primary key.

moment_name : (char*80) Name or explication of this moment (position on the time axis).

year : (integer*8) Year of this moment.

month : (integer*1) Month of this moment.

day : (integer*1) Day of this moment.

hour : (integer*1) Hour of this moment.

minute : (integer*1) Minute of this moment.

second : (integer*1) Second of this moment.

utc_difference : (float*4) Difference of the scale to UTC in hours as UTC minus local hour, i.e. being west of Greenwich results in positive values of utc_difference.

Table DIM_TYPE

dim_type_id : (integer) Primary key.

dim_type_name : (char*80) Type of the scale in this dimension.

basic set of values : equidistant grid; irregular grid; single point(s); spectral.

Table POINT_CONNECT

point_set_id : (integer) Foreign key to POINT_SET.

point_id : (integer) Foreign key to POINT.

moment_id : (integer) Foreign key to MOMENT.

sequence_no : (integer*4) Sequence number of a point within one point set.

Table SCALE_CONNECT

scale_id : (integer) Foreign key to SCALE.

position_id : (integer) Foreign key to POSITION.

sequence_no : (integer*4) Sequence number of a scale position in this dimension.

Table TIME_CONNECT

time_id : (integer) Foreign key to TIME.

moment_id : (integer) Foreign key to MOMENT.

Module DATA_ACCESS (4 tables)

additionally into PARAMETER

data_access_id : (integer) Foreign key to DATA_ACCESS.

Table DATA_ACCESS

data_access_id : (integer) Primary key.

access_structure_id : (integer) Foreign key to ACCESS_STRUCTURE.

storage1_id : (integer) Foreign key to STORAGE1.

storage2_id : (integer) Foreign key to STORAGE2.

storage3_id : (integer) Foreign key to STORAGE3.

storage4_id : (integer) Foreign key to STORAGE4.

rec_structure_id : (integer) Foreign key to REC_STRUCTURE.

modification_date : (date) Date when the information of this table was valid.

Table REC_STRUCTURE

rec_structure_id : (integer) Primary key.

structure_name : (char*250) Name, describing the structure given for the five dimensions in this record.

dim1_type : (char*10) Type of the fastest varying dimension in storage (= name of the table discribing this dimension).

dim1_id : (integer) Foreign key to POINT_SET, SCALE, TIME, or PARA_ORDER according to the description in dim1_type. PARA_ORDER can be a local table to specify the storage sequence of a number of different topics.

dim2_type : (char*10) Type of the second fastest varying dimension in storage (= name of the table discribing this dimension).

dim2_id : (integer) Foreign key to POINT_SET, SCALE, TIME, or PARA_ORDER according to the description in dim2_type.

dim3_type : (char*10) Type of the third fastest varying dimension in storage (= name of the table discribing this dimension).

dim3_id : (integer) Foreign key to POINT_SET, SCALE, TIME, or PARA_ORDER according to the description in dim3_type.

dim4_type : (char*10) Type of the fourth fastest varying dimension in storage (= name of the table discribing this dimension).

dim4_id : (integer) Foreign key to POINT_SET, SCALE, TIME, or PARA_ORDER according to the description in dim4_type.

dim5_type : (char*10) Type of the fifth fastest varying dimension in storage (= name of the table discribing this dimension).

dim5_id : (integer) Foreign key to POINT_SET, SCALE, TIME, or PARA_ORDER according to the description in dim5_type. PARA_ORDER can be a local table to specify the storage sequence of a number of different topics.

Table STORAGE (aliased as STORAGE1, STORAGE2, STORAGE3, STORAGE4)

storage_id : (integer) Primary key.

storage_name : (char*250) The storage information assigned to this field according the description in table ACCESS_STRUCTURE.

Table ACCESS_STRUCTURE

access_structure_id : (integer) Primary key.

access_structure_name : (char*80) Name of the access structure described in access_structure_descr.

access_structure_descr : (char*2000) Description of the form by which the data access information is stored in STORAGE1 to STORAGE4.

basic set of values : Any description how the data are arranged in the storage.

Examples

In order to provide an impression, how to fill the data model, a few practical examples are given in the Appendix: the minimum CERA-2 entry, a PIK observational data entry and a DKRZ climate model data entry.

Minimum CERA-2 Entry

The minimum entry as given in Appendix B is possible within the CERA-2 data model but it is not recommended. The minmum entry provides only coarse information on the content and on the coverage. No information is provided on contact, on publication, on attributes and on distribution. The only information provided in the example is that some kind of global ocean data are stored for the time between 1854 and 1989. Nothing else is available from the entry given in Appendix B.

Plain CERA-2 Entry at PIK

The PIK entry describes as complete as possible a dataset on soil degradation, that has been fetched as input to ecological model calculations. Data are gathered through a questionaire. In addition to the minimum entry, information on responsible persons, dataset contents, storage and access is provided. The tables and attributes are listed in detail in Appendix C.

As in the minimum entry the PIK entry defines no relations to other entries. A plain structure is implemented.

Hierarchical CERA-2 Entry at DKRZ

The climate model data at DKRZ are organized in a two-layer, hierarchical structure: the experiment and the related datasets. The experiment contains the overall description of the related datasets. Experimental descriptions are realized for climate model calculations and for observational data. The datasets contain the description and the data themselves. Therefore the CERA database at DKRZ contains not only the complete data descriptions (meta data) but also the climate data (numbers).

The DKRZ example provides information on the data which are produced in the atmopshere re-analysis project of the European Centre for Medium-Range Weather Forecasts. The experiment description contains metadata for the compiled monthly mean surface data in T42 climate model resolution. The data provide a consistent global data set of the the last 15 years. The hierarchical structure is defined in the table CONNECT_TYPE and the related dataset(s) is/are specified in table ENTRY_CONNECT.

As part of the experiment "ERA15_SFCFC00_MONTHLY_T42" the subordinated dataset "ERA15_SFCFC00_MSLP" gives a complete description of the mean sea level data. Not only the contents information, but also technical information like the definition of the 5-dimensional data space, the storage and the access information is included. The DKRZ example is given in detail in Appendix D.

Related Information

CERA Central Page

A central WWW page is maintained at PIK in order to disseminate not only this report but also related information about the status and the progress of the CERA environment:

URL: http://www.PIK-Potsdam.DE/dept/dc/e/sdm/cera/

The exact table attribute definitions together with their lists of values, as far as they exist, are given at the above location, too.

CERA users can download SQL scripts to carry out the data base installation and to facilitate loading and handling of CERA. If CERA is run on an ORACLE^{TM} data base management system (DBMS), only few institute specific changes of the scripts are necessary to adapt them to a certain location. In this case, existing graphical user interfaces (GUI) on base of ORACLE Forms or Java can be used.

SQL scripts which are extracting from the CERA-2 datamodel metadata in NASA's directory interchange format (DIF) have been developed at AWI and will are available via the CERA Central web page, as well.

Platform Independent GUIs

Grapical user interfaces (GUIs) which are independent from specific machine architectures were more and more required in the recent past. First specfic realizations were available. As A Flexible Retrieval Interface (AFRI), a more general solution has been developed at PIK. The GUI bases on the WWW and uses Java and JDBC to realize the database connectivity. It is applicable with any common WWW browser and provides scaleable access from world wide use (Internet) to inhouse use (Intranet). Alternatively, AFRI can be run locally, thus avoiding the need for an internet server.

Database Interoperability

Within the concept of geographically distributed data storage the interoperabilty of different local data(base) systems is strongly required. Furthermore, the substitution of plain data retrieval (numbers and ASCII text) by a more general information retrieval (graphics, data processing and evaluation) is discussed and planned. The traditional 2-layer architecture (client and server) will be sustituted by a 3-layer architecture consisting of graphical based client, application server and data server. Parts of the original client and server tasks are separated and concentrated in the application server layer. This middle ware layer is responsible for the database connectivity and for the data processing. The entire concept is scalable in dependence of the size of problems and of available hardware. The realization by using CORBA is discussed. First steps are planned in cooperation between Research Centre Geesthacht, University Hamburg and DKRZ.

Comparison of CERA-2 with CSDGM (FGDC) and DIF (NASA)

Implementations of meta data models are evaluated in the context, as they are sufficient for international standards. Many of them have evolved within organizations and on national levels during the last years. Examples are DIF (Directory Interchange Format, NASA 1998a), FGDC Content Standard (Federal Geographic Data Committee Metadata Content Standards, FGDC 1996), ANZLIC (Australian New Zealand Land Information Council, ANZLIC 1998), INFOClima (1996), Dublin Core (1998) and CEN TC287 (1998). Today a new ISO standard (TC211 geomatics standardization activity (ISO/TC211 1998) , ISO15046-15) is under development to satisfy demands from the scientific community, the environmental agencies and the economy for a common standard to maintain relevant descriptions of geospatial data. Since the ISO standard is not adopted at present and uses the drafts and the final standard commercially, the CERA data model cannot be judged in its context. The implementation of CERA is therefore compared with DIF (NASA) and CSDGM (FGDC).

Basic principles of meta data descriptions are: All standards provide a common set of terminology and definitions for the documentation of geospatial data. All standards establish the names of data elements and groups of data elements, and information about the values that are to be provided for the data elements. The standards provide information about terms that are mandatory, mandatory under certain conditions, and optional. All standards do not specify how this information is organized in a computer system or in a data transfer, nor the means by which this information is transmitted or communicated to the user.

The first implementation of CERA (CERA-1) oriented itself at the definition of the NASA DIF standard at that time (Hoeck et al. 1995). During the last years DIF and CERA have developed independently. The advancement of DIF was influenced by the national standardisation of meta data with geographical reference within the USA (FGDC Metadata Content Standard), while CERA was adapted to practical needs in the AWI, DKRZ and PIK.

To map the implementation of CERA-2 with NASA DIF, SQL procedures have been developed, mapping the results to the formal description of DIF (NASA 1998b). PIK presents this page at the CERA Central Page.

CERA-2 fulfills the mandatory requests of DIF. With the optional fields there are the following deviations:

- The field Data_Resolution belongs not to the CERA Core but to the CERA Module DATA_ORGANIZATION.
- Some access related DIF fields can be included into the CERA Module DATA_ACCESS: IDN_Node, Data_Set_ID, Multimedia_Sample.
- As CERA regards depth as negative altitude, Minimum_Depth und Maximum_Depth are covered by the altitude limits.
- Parent_DIF is accessible, if a hierarchical entry structure is implemented. In this case, it is the DIF of the parent entry.

- The DIF fields Sensor_Name, Source, Distribution_Media and Data_Set_Language are not mappable to CERA Core and Modules.
- Issue_Identification can be stored in CERA, but it has no own attribute.

Some of the terms (e.g., Source, Sensor) were less relevant in the context of AWI, DKRZ and PIK because they were commonly used in remote sensing.

Local implementations of CERA-2 have the opportunity to fill this gap to DIF by implementing the missing terms in modules or own local extensions.

Looking at CSDGM of FGDC (NASA 1998b), the following fields are missing in CERA-2:

- 1.10 Browse Graphic
- The more detailed information of CSDGM, part 4., Spatial Reference Information, will partly be included in a CERA Module DATA_STRUCTURE, which has still to be developed; some general spatial reference information is included in CERA-2.

The new ISO Standard 15046-15 will certainly include new mandatory fields, which are at present not fulfilled by the CERA-2 model. A minor adjustment of the data model may therefore be necessary within the next years.

References

- ANZLIC, 1998: Core Metadata Elements for Land and Geographic Directories in Australia and New Zealand
(URL: http://www.anzlic.org.au/metaelem.htm)
- Bretherton, Francis, 1994: Refrence Model for Metadata: A Strawman. Whitepaper, University Wisconsin.
(URL: http://www.llnl.gov/liv_comp/metadata/papers/whitepaper-bretherton.ps)
- CEN TC287:
(URL: http://forum.afnor.fr/afnor/WORK/AFNOR/GPN2/Z13C/PUBLIC/WEB/ENGLISH/index.htm)
- Hoeck, H., H. Thiemann, M. Lautenschlager, I. Jessel, B. Marx and M. Reinke, 1995: The CERA Metadata Model.
Technical Report No. 9, pp. 55, DKRZ, Hamburg
(URL: http://www.dkrz.de/forschung/reports.html)
- DUBLIN CORE, 1998: Dublin Core Metadata
(URL: http://purl.oclc.org/metadata/dublin_core/)
- FGDC, 1998: Content Standard for Digital Geospatial Metadata (CSDGM), Vers. 2. Ferderal Geographic data Committee. Reston, VA, USA.
(URL: http://fgdc.er.usgs.gov/Metadata/ContStan.html)
- FGDC, 1996 Content Standard for Digital Geospatial Metadata, Version 1.0
(URL: http://www.fgdc.gov/Metadata/metav1-0.html)
- INFOCLIMA, 1996: WMO INFOCLIMA Project
(URL: http://www.wmo.ch/web/wcp/wcdmp/infoclim/infoclim.html)
- ISO/TC211, 1998: Geographic information/Geomatics
(URL: http://www.statkart.no/isotc211/welcome.html)
- NASA, 1998a: Directory Interchange Format (DIF) Writer‘s Guide, Vers. 6.0. Global Interchange Master Directory. National Aeronautic and Space Administration.
(URL: http://gcmd.nasa.gov/difguide/)
- NASA, 1998b: FGDC Metadata Standard to GCMD DIF.
(URL: http://gcmd.gsfc.nasa.gov/dif_mapping/fgdc_dif_map6_0.html)

Appendix A


CERA CORE SCHEMA 2.3
DISTRIBUTION
USE_CONSTRAINT
use_constraint_id
constraint_descr
ACCESS_CONSTRAINT
access_constraint_id
constraint_descr
FEES
fees_id
fees_acronym
fees_descr
FORMAT
format_id
format_acronym
format_descr
ACCESS_TYPE
access_type_id
access_acronym
access_descr
DISTRIBUTION
entry_id
access_type_id
data_size
format_id
fees_id
access_constraint_id
use_constraint_id
SPATIAL INFORMATION
SPATIAL_DATA_ORG
spatial_data_org_id
reference_method
HOR_COORD_SYS
hor_coord_sys_id
sys_descr
VER_COORD_SYS
ver_coord_sys_id
sys_descr
SPATIAL_REFERENCE
entry_id
hor_coord_sys_id
ver_coord_sys_id
ATTRIBUTE
TOPIC
topic_id
topic_name
topic_acronym
topic_descr
topic_pointer
topic_level
UNIT
unit_id
unit_name
unit_acronym
unit_descr
PARAMETER
entry_id
topic_id
unit_id
spatial_data_org_id
aggregation_id
(data_org_id)
(data_access_id)
AGGREGATION
aggregation_id
aggregation_descr
CURRENTNESS_REF
currentness_ref_id
currentness_ref_descr
TEMPORAL_COVERAGE
temporal_coverage_id
start_year
start_month
start_day
stop_year
stop_month
stop_day
currentness_ref_id
SPATIAL_COVERAGE
spatial_coverage_id
min_lat
max_lat
min_lon
max_lon
min_altitude
max_altitude
min_alt_unit_id
max_alt_unit_id
CONTACT
PERSON
person_id
first_name
second_name
last_name
title
institute_id
telephone
fax
email
url
INSTITUTE
institute_id
institute_name
institute_acronym
department_name
department_acronym
country
state_or_province
place
street
street_postal_code
pobox
pobox_postal_code
url
additional_info
CONTACT
entry_id
institute_id
person_id
contact_type_id
CONTACT_TYPE
contact_type_id
contact_type
COVERAGE
LOCATION_CONNECT
location_id
entry_id
LOCATION
location_id
location_name
location_pointer
location_level
ref_date
min_lat
max_lat
min_lon
max_lon
COVERAGE
entry_id
temporal_coverage_id
spatial_coverage_id
PUBLICATION
REFERENCE
citation_id
entry_id
ref_type_id
REF_TYPE
ref_type_id
ref_type_descr
CITATION
citation_id
title
authors
publication
publisher
publication_date
editor
country
state
place
edition
access_spec
presentation_id
citation_type_id
PRESENTATION
presentation_id
presentation_descr
CITATION_TYPE
citation_type_id
citation_type
citation_type_descr
STATUS
QUALITY
quality_id
accuracy_report
consistency_report
completeness_report
horizontal_acc_report
vertical_acc_report
PROCESS_STEP
entry_id
process_descr
process_date
person_id
PROGRESS
progress_id
progress_descr
METADATA ENTRY
GENERAL_KEY
general_key_id
general_key
KEY_CONNECT
general_key_id
entry_id
PROJECT
project_id
project_name
project_acronym
project_descr
CAMPAIGN
project_id
entry_id
SUMMARY
summary_id
summary
ENTRY
entry_id
entry_name
entry_acronym
entry_type_id
summary_id
quality_id
progress_id
creation_date
review_date
revision_date
future_review_date
ENTRY_TYPE
entry_type_id
entry_type
entry_type_descr
ENTRY_CONNECT
gen_entry_id
spec_entry_id
connect_type_id
CONNECT_TYPE
connect_type_id
connect_type
connect_type_descr
DIF_ENTRY
entry_id
dif_title
dif_name
MODULES AND LOCAL EXTENSIONS
LOCAL_CONNECT
..._id
...
ENTITIES
LISTS OF VALUES
RELATIONS
PIK, DKRZ, AWI 12/97

Appendix B: Minimum CERA-2 Entry

The attribute value "n/a" stands for "not applicable“ or "not available" and "not filled" means that no information is provided. In this case the data model default for "NOT NULL" fields is used for the attribute. Keys which have no value are indicated by "0" (zero).

Block: METADATA ENTRY	
Table ENTRY	.
ENTRY_ID	111
ENTRY_NAME	COADS data
ENTRY_ACRONYM	not filled
ENTRY_TYPE_ID	0
SUMMARY_ID	0
QUALITY_ID	0
PROGRESS_ID	0
CREATION_DATE	08.06.1998
REVIEW_DATE	08.06.1998
REVISION_DATE	08.06.1998
FUTURE_REVIEW_DATE	08.06.1998
Table SUMMARY	not filled
Table ENTRY_TYPE	not filled
Table ENTRY_CONNECT	not filled
Table CONNECT_TYPE	not filled
Table KEY_CONNECT	.
GENERAL_KEY_ID	5
ENTRY_ID	111
Table GENERAL_KEY	.
GENERAL_KEY_ID	5
GENERAL_KEY	ocean surface
Table CAMPAIGN	not filled
Table PROJECT	not filled
Table DIF_ENTRY	not filled
Block: STATUS	
Table QUALITY	not filled
Table PROGRESS	not filled
Table PROCESS_STEP	not filled
Block: ATTRIBUTE	
Table PARAMETER	not filled
Table TOPIC	not filled
Table UNIT	not filled
Table AGGREGATION	not filled
Block: PUBLICATION	
Table REFERENCE	not filled
Table CITATION	not filled
Table CITATION_TYPE	not filled

Table PRESENTATION	not filled
Block: CONTACT	
Table CONTACT	not filled
Table CONTACT_TYPE	not filled
Table PERSON	not filled
Table INSTITUTE	not filled
Block: COVERAGE	
Table LOCATION_CONNECT	not filled
Table COVERAGE	.
ENTRY_ID	111
TEMPORAL_COVERAGE_ID	10
SPATIAL_COVERAGE_ID	10
Table TEMPORAL_COVERAGE	.
TEMPORAL_COVERAGE_ID	10
START_YEAR	1854
START_MONTH	not filled
START_DAY	not filled
STOP_YEAR	1989
STOP_MONTH	not filled
STOP_DAY	not filled
CURRENTNESS_REF_ID	0
Table CURRENTNESS_REF	not filled
Table SPATIAL_COVERAGE	.
SPATIAL_COVERAGE_ID	10
MIN_LAT	-90
MAX_LAT	90
MIN_LON	-180
MAX_LON	180
MIN_ALTITUDE	0
MAX_ALTITUDE	0
MIN_ALT_UNIT_ID	0 (not filled)
MAX_ALT_UNIT_ID	0 (not filled)
Block: SPATIAL INFORMATION	
Table SPATIAL_DATA_ORG	not filled
Table SPATIAL_REFERENCE	not filled
Table VER_COORD_SYS	not filled
Table HOR_COORD_SYS	not filled
Block: DISTRIBUTION	
Table DISTRIBUTION	not filled
Table ACCESS_TYPE	not filled
Table FEES	not filled
Table USE_CONSTRAINT	not filled

Appendix C: Plain CERA-2 Entry at PIK

The attribute value "n/a" stands for "not applicable“ or "not available" and "not filled" means that no information is provided. In this case the data model default for "NOT NULL" fields is used for the attribute. Keys which have no value are indicated by "0" (zero).

Block: METADATA ENTRY	
Table ENTRY	.
ENTRY_ID	193
ENTRY_NAME	Global Assessment of Human-Induced Soil Degradation GLASOD
ENTRY_ACRONYM	GLASOD
ENTRY_TYPE_ID	1
SUMMARY_ID	193
QUALITY_ID	60
PROGRESS_ID	4
CREATION_DATE	18.03.1998
REVIEW_DATE	18.03.1998
REVISION_DATE	18.03.1998
FUTURE_REVIEW_DATE	18.03.2003
Table SUMMARY	.
SUMMARY_ID	193
SUMMARY	Database contains information on soil degradation within map units as reported by soil experts around the world through a questionnaire. Documentation available at Otto TestMan.
Table ENTRY_TYPE	.
ENTRY_TYPE_ID	1
ENTRY_TYPE	normal
ENTRY_TYPE_DESCR	no different types used until now
Table ENTRY_CONNECT	n/a
Table CONNECT_TYPE	n/a
Table KEY_CONNECT	.
GENERAL_KEY_ID	2404
ENTRY_ID	193
Table GENERAL_KEY	.
GENERAL_KEY_ID	2404
GENERAL_KEY	soil degradation
Table CAMPAIGN	.
ENTRY_ID	193
PROJECT_ID	1
Table PROJECT	.
PROJECT_ID	1
PROJECT_ACRONYM	n/a
PROJECT_NAME	n/a
PROJECT_DESCR	The project the data belongs to is unknown
Table DIF_ENTRY	no entries

Block: STATUS	
Table QUALITY	.
QUALITY_ID	60
ACCURACY_REPORT	not checked
CONSISTENCY_REPORT	no inconsistencies found
COMPLETENESS_REPORT	48% of the globe, depending on the area
HORIZONTAL_ACC_REPORT	not checked
VERTICAL_ACC_REPORT	no vertical extension
Table PROGRESS	.
PROGRESS_ID	4
PROGRESS_DESCR	complete, processed
Table PROCESS_STEP	.
ENTRY_ID	193
PROCESS_DESCR	Data has been scaled to continent size and converted to .ps files, stored in the same directory.
PROCESS_DATE	24.04.1998
PERSON_ID	144
Block: ATTRIBUTE	
Table PARAMETER	.
ENTRY_ID	193
TOPIC_ID	0
UNIT_ID	1
AGGREGATION_ID	1
SPATIAL_DATA_ORG_ID	4
(DATA_ORG_ID)	0
(DATA_ACCESS_ID)	188
Table TOPIC	.
TOPIC_ID	562
TOPIC_NAME	degradation
TOPIC_ACRONYM	n/a
TOPIC_DESCR	n/a
TOPIC_POINTER	44
TOPIC_LEVEL	4
Table UNIT	.
UNIT_ID	1
UNIT_NAME	unknown
UNIT_ACRONYM	unknown
UNIT_DESCR	the unit is not known
Table AGGREGATION	.
AGGREGATION_ID	1
AGGREGATION_DESCR	none
Block: PUBLICATION	
Table REFERENCE	.
ENTRY_ID	193
CITATION_ID	188
REF_TYPE_ID	1

Table REF_TYPE	.
REF_TYPE_ID	1
REF_TYPE_DESCR	publication provides entry data
Table CITATION	.
CITATION_ID	188
TITLE	Global Assessment of Human-Induced Soil Degradation GLASOD
AUTHORS	n/a
PUBLICATION	n/a
PUBLISHER	UNEP / GRID
EDITOR	n/a
PUBLICATION_DATE	01.01.1991
COUNTRY	n/a
STATE	n/a
PLACE	Wageningen
EDITION	n/a
ACCESS_SPEC	n/a
PRESENTATION_ID	8
CITATION_TYPE_ID	1
Table CITATION_TYPE	.
CITATION_TYPE_ID	1
CITATION_TYPE	unknown
CITATION_TYPE_DESCR	Type of the publication is not known
Table PRESENTATION	.
PRESENTATION_ID	8
PRESENTATION_DESCR	map: thematic
Block: CONTACT	
Table CONTACT	(this and following four tables have two records)
ENTRY_ID	193
.	193
INSTITUTE_ID	2
.	79
PERSON_ID	18
.	144
CONTACT_TYPE_ID	8
.	3
Table CONTACT_TYPE	(table has two records)
CONTACT_TYPE_ID	8
.	3
CONTACT_TYPE	inhouse (in this institute)
.	distributor (where to get data)
Table PERSON	(table has two records)
PERSON_ID	18
.	144
FIRST_NAME	Erwin
.	unknown
SECOND_NAME	n/a

.	n/a
LAST_NAME	Exampleman
.	name of person unknown
TITLE	Dr.
.	n/a
INSTITUTE_ID	2
.	79
TELEPHONE	0331-288-9999
.	04122-999 999
FAX	0331-288-9998
.	04122-999 998
URL	www.pik-potsdam.de/~exman
.	n/a
EMAIL	ExMan@pik-potsdam.de
.	n/a
Table INSTITUTE	(table has two records)
INSTITUTE_ID	2
.	79
INSTITUTE_NAME	Potsdam Institut fuer Klimafolgenforschung
.	UNEP Global Resource Information Database
INSTITUTE_ACRONYM	PIK
.	UNEP GRID
DEPARTMENT_NAME	n/a
.	n/a
DEPARTMENT_ACRONYM	n/a
.	n/a
COUNTRY	Germany
.	Switzerland
STATE_OR_PROVINCE	Brandenburg
.	n/a
PLACE	Potsdam
.	Crouge
STREET	Telegrafenberg C4
.	6, rue de la Gabelle
STREET_POSTAL_CODE	n/a
.	1227
POBOX	60 12 03
.	n/a
POBOX_POSTAL_CODE	14412
.	n/a
URL	www.pik-potsdam.de
.	n/a
ADDITIONAL_INFO	n/a
.	n/a
Alias Table INSTITUTE1	Alias link to the INSTITUTE, the person is affilated to.
Block: COVERAGE	

Table LOCATION_CONNECT	.
ENTRY_ID	193
LOCATION_ID	1
Table LOCATION	.
LOCATION_ID	1
LOCATION_NAME	World
LOCATION_POINTER	0
LOCATION_LEVEL	1
REF_DATE	20.03.1960
MIN_LAT	-90
MAX_LAT	90
MIN_LON	-180
MAX_LON	180
Table COVERAGE	.
ENTRY_ID	193
TEMPORAL_COVERAGE_ID	69
SPATIAL_COVERAGE_ID	82
Table TEMPORAL_COVERAGE	.
TEMPORAL_COVERAGE_ID	69
START_YEAR	1991
START_MONTH	0
START_DAY	0
STOP_YEAR	1991
STOP_MONTH	0
STOP_DAY	0
CURRENTNESS_REF_ID	5
Table CURRENTNESS_REF	.
CURRENTNESS_REF_ID	5
CURRENTNESS_REF_DESCR	
Table SPATIAL_COVERAGE	.
SPATIAL_COVERAGE_ID	82
MIN_LAT	-90
MAX_LAT	90
MIN_LON	-180
MAX_LON	180
MIN_ALTITUDE	0
MAX_ALTITUDE	0
MIN_ALT_UNIT_ID	0
MAX_ALT_UNIT_ID	0
AliasTable UNIT1	Alias link to the UNIT of MIN_ALTITUDE
AliasTable UNIT2	Alias link to the UNIT of MAX_ALTITUDE
Block: SPATIAL INFORMATION	
Table SPATIAL_DATA_ORG	.
SPATIAL_DATA_ORG_ID	4
REFERENCE_METHOD	vector
Table SPATIAL_REFERENCE	.

ENTRY_ID	193
HOR_COORD_SYS_ID	4
VER_COORD_SYS_ID	5
Table HOR_COORD_SYS	.
HOR_COORD_SYS_ID	2
SYS_DESCR	LON and LAT: geographic
Table VER_COORD_SYS	.
VER_COORD_SYS_ID	2
SYS_DESCR	no vertical extension
Block: DISTRIBUTION	
Table DISTRIBUTION	.
ENTRY_ID	193
ACCESS_TYPE_ID	11
DATA_SIZE	0
FORMAT_ID	47
FEES_ID	2
ACCESS_CONSTRAINT_ID	2
USE_CONSTRAINT_ID	2
Table FORMAT	.
FORMAT_ID	47
FORMAT_ACRONYM	ARC
FORMAT_DESCR	ARC/INFO
Table ACCESS_TYPE	.
ACCESS_TYPE_ID	11
ACCESS_ACRONYM	disc: normal file
ACCESS_DESCR	data are files in the file system
Table FEES	.
FEES_ID	2
FEES_ACRONYM	none
FEES_DESCR	The distribution of the data is free
Table ACCESS_CONSTRAINT	.
ACCESS_CONSTRAINT_ID	2
CONSTRAINT_DESCR	unrestricted
Table USE_CONSTRAINT	.
USE_CONSTRAINT_ID	2
CONSTRAINT_DESCR	unrestricted
Modul: DATA_ACCESS	
Table DATA_ACCESS	.
DATA_ACCESS_ID	188
ACCESS_STRUCTURE_ID	3
STORAGE1_ID	47
STORAGE2_ID	375
STORAGE3_ID	376
STORAGE4_ID	0
REC_STRUCTURE_ID	0

MODIFICATION_DATE	18.03.1998
Table ACCESS_STRUCTURE	.
ACCESS_STRUCTURE_ID	3
ACCESS_STRUCTURE_NAME	Conventional storage description: storage 1 - hostname, storage 2 - directory, storage 3 - file, storage 4 - comment
Alias Table STORAGE1	.
STORAGE_ID	47
STORAGE_NAME	machine.domain.de
Alias Table STORAGE2	.
STORAGE_ID	375
STORAGE_NAME	/u/erwin/global/GRID/land_surface/processes/glasod
Alias Table STORAGE3	.
STORAGE_ID	376
STORAGE_NAME	file.dat
Alias Table STORAGE4	.
STORAGE_ID	0
STORAGE_NAME	not filled

Appendix D: Hierarchical CERA-2 Entries at DKRZ

The attribute value "n/a" stands for "not applicable“ or "not available" and "not filled" means that no information is provided. In this case the data model default for "NOT NULL" fields is used for the attribute. Keys which have no value are indicated by "0" (zero).

Table: Experiment Metadata

Block: METADATA ENTRY	
Table ENTRY	.
ENTRY_ID	144
ENTRY_NAME	ERA15_SFCFC00_MONTHLY_T42
ENTRY_ACRONYM	n/a
ENTRY_TYPE_ID	2
SUMMARY_ID	144
QUALITY_ID	144
PROGRESS_ID	6
CREATION_DATE	11-NOV-97
REVIEW_DATE	11-NOV-97
REVISION_DATE	11-NOV-97
FUTURE_REVIEW_DATE	not filled
Table SUMMARY	.
SUMMARY_ID	144
SUMMARY	Operational analyses are affected by the major changes in models, analysis technique, assimilation and observation usage which are an essential product of research and progress. They also can make use only of those observations which become available within near real time. The ECMWF (European Centre for Medium-Range Weather Forecasts) performed a consistent re-analysis of atmospheric data using a „frozen“ production system. The ECMWF Re-Analysis (ERA) Project produced a new, validated 15 year data set of assimilated data for the period 1979 to 1993. The assimilation scheme consisted mainly of the Integrated Forecast System (IFS) version of the ECMWF forecast model with T106 resolution on 31 vertical hybrid levels. An intermittent statistical (optimum interpolation) analysis with 6 hour cycling was performed. It includes a diabatic, non-linear normal mode initialisation (5 vertical modes). The current data set consists monthly means of initialized surface fields.
Table ENTRY_TYPE	.
ENTRY_TYPE_ID	2
ENTRY_TYPE	Experiment
ENTRY_TYPE_DESCR	Contains the overall description of related datasets
Table ENTRY_CONNECT	.
GEN_ENTRY_ID	144
SPEC_ENTRY_ID	6187
CONNECT_TYPE_ID	1
Table CONNECT_TYPE	.
CONNECT_TYPE_ID	1
CONNECT_TYPE	hierarchical

Table: Experiment Metadata

CONNECT_TYPE_DESCR	SPEC_ENTRY_ID specifies the subsequent entries (e.g. datasets) which follow GEN_SPEC_ID (e.g. experiment)
Table KEY_CONNECT	.
GENERAL_KEY_ID	1
ENTRY_ID	144
GENERAL_KEY_ID	2
ENTRY_ID	144
Table GENERAL_KEY	
GENERAL_KEY_ID	1
GENERAL_KEY	Atmosphere Re-Analysis
GENERAL_KEY_ID	2
GENERAL_KEY	Data Assimilation
Table CAMPAIGN	.
ENTRY_ID	144
PROJECT_ID	41
Table PROJECT	.
PROJECT_ID	41
PROJECT_ACRONYM	ERA15
PROJECT_NAME	ECMWF Re-Analysis Project
PROJECT_DESCR	Production of a new, validated 15 year data set.
Table DIF_ENTRY	
ENTRY_ID	144
DIF_TITLE	ERA15 initialized surface fields monthly means
DIF_NAME	ERA15_SFCFC00_MONTHLY_T42
Block: STATUS	
Table QUALITY	.
QUALITY_ID	144
ACCURACY_REPORT	assimilated data
CONSISTENCY_REPORT	not filled
COMPLETENESS_REPORT	not filled
HORIZONTAL_ACC_REPORT	not filled
VERTICAL_ACC_REPORT	not filled
Table PROGRESS	.
PROGRESS_ID	6
PROGRESS_DESCR	complete, processed
Table PROCESS_STEP	n/a
Block: ATTRIBUTE	
Table PARAMETER	n/a
Table TOPIC	n/a
Table UNIT	n/a
Table AGGREGATION	n/a
Block: PUBLICATION	
Table REFERENCE	.
ENTRY_ID	144
CITATION_ID	103

Table: Experiment Metadata

REF_TYPE_ID	17
Table REF_TYPE	.
REF_TYPE_ID	17
REF_TYPE_DESCR	Publication provides information on the methods which are used to produce the data.
Table CITATION	.
CITATION_ID	103
TITLE	ERA Description
AUTHORS	J.K. Gibson, P. Kallberg, S. Uppala, A. Hernandez, A. Nomura, E. Serrano
PUBLICATION	ECMWF Re-Analysis Project Report Series
PUBLISHER	European Centre for Medium-Range Weather Forecasts
EDITOR	n/a
PUBLICATION_DATE	July 1997
COUNTRY	United Kingdom
STATE	England
PLACE	Reading
EDITION	n/a
ACCESS_SPEC	n/a
PRESENTATION_ID	10
CITATION_TYPE_ID	8
Table CITATION_TYPE	.
CITATION_TYPE_ID	8
CITATION_TYPE	unreviewed publication
CITATION_TYPE_DESCR	internal report series at ECMWF
Table PRESENTATION	.
PRESENTATION_ID	10
PRESENTATION_DESCR	numerical data
Block: CONTACT	
Table CONTACT	.
ENTRY_ID	144
INSTITUTE_ID	196
PERSON_ID	299
CONTACT_TYPE_ID	1
Table CONTACT_TYPE	.
CONTACT_TYPE_ID	1
CONTACT_TYPE	investigator
Table PERSON	.
PERSON_ID	299
FIRST_NAME	Rex
SECOND_NAME	n/a
LAST_NAME	Gibson
TITLE	Dr.
INSTITUTE_ID	196
TELEPHONE	not filled
FAX	not filled

Table: Experiment Metadata

URL	n/a
EMAIL	not filled
Table INSTITUTE	
INSTITUTE_ID	196
INSTITUTE_NAME	European Centre for Medium-Range Weather Forecasts
INSTITUTE_ACRONYM	ECMWF
DEPARTMENT_NAME	n/a
DEPARTMENT_ACRONYM	n/a
COUNTRY	United Kingdom
STATE_OR_PROVINCE	England
PLACE	Reading
STREET	Shinfield Park
STREET_POSTAL_CODE	RG2 9AX
POBOX	n/a
POBOX_POSTAL_CODE	n/a
URL	http://www.ecmwf.int/
ADDITIONAL_INFO	n/a
Block: COVERAGE	
Table LOCATION_CONNECT	.
ENTRY_ID	144
LOCATION_ID	728
Table LOCATION	.
LOCATION_ID	728
LOCATION_NAME	global
LOCATION_POINTER	0
LOCATION_LEVEL	1
REF_DATE	n/a
MIN_LAT	-90
MAX_LAT	90
MIN_LON	-180
MAX_LON	180
Table COVERAGE	.
ENTRY_ID	144
TEMPORAL_COVERAGE_ID	9
SPATIAL_COVERAGE_ID	9
Table TEMPORAL_COVERAGE	.
TEMPORAL_COVERAGE_ID	9
START_YEAR	1978
START_MONTH	12
START_DAY	01
STOP_YEAR	1994
STOP_MONTH	02
STOP_DAY	28
CURRENTNESS_REF_ID	5
Table CURRENTNESS_REF	.

Table: Experiment Metadata

CURRENTNESS_REF_ID	5
CURRENTNESS_REF_DESCR	calendrical
Table SPATIAL_COVERAGE	.
SPATIAL_COVERAGE_ID	9
MIN_LAT	-90
MAX_LAT	90
MIN_LON	-180
MAX_LON	180
MIN_ALTITUDE	0
MAX_ALTITUDE	0
MIN_ALT_UNIT_ID	0 (not filled)
MAX_ALT_UNIT_ID	0 (not filled)
Block: SPATIAL INFORMATION	
Table SPATIAL_DATA_ORG	n/a
Table SPATIAL_REFERENCE	.
ENTRY_ID	144
HOR_COORD_SYS_ID	2
VER_COORD_SYS_ID	0
Table HOR_COORD_SYS	
HOR_COORD_SYS_ID	2
SYS_DESCR	LON and LAT: geographic
Table VER_COORD_SYS	not filled
Block: DISTRIBUTION	
Table DISTRIBUTION	.
ENTRY_ID	144
ACCESS_TYPE_ID	9
DATA_SIZE	n/a
FORMAT_ID	16
FEES_ID	2
ACCESS_CONSTRAINT_ID	8
USE_CONSTRAINT_ID	8
Table FORMAT	.
FORMAT_ID	26
FORMAT_ACRONYM	GRIB
FORMAT_DESCR	WMO format GRIdded Binary data
Table ACCESS_TYPE	.
ACCESS_TYPE_ID	9
ACCESS_ACRONYM	DB
ACCESS_DESCR	disc: DB tables
Table FEES	.
FEES_ID	2
FEES_ACRONYM	none
FEES_DESCR	n/a
Table ACCESS_CONSTRAINT	.
ACCESS_CONSTRAINT_ID	8

Table: Experiment Metadata

CONSTRAINT_DESCR	Access is restricted to German scientists.
Table USE_CONSTRAINT	.
USE_CONSTRAINT_ID	8
CONSTRAINT_DESCR	Data use is only permitted for science.
Modul: DATA_ORGANIZATION	

Table: Dataset Metadata

Block: METADATA ENTRY	
Table ENTRY	.
ENTRY_ID	6187
ENTRY_NAME	ERA15_SFCFC00_MSLP
ENTRY_ACRONYM	n/a
ENTRY_TYPE_ID	3
SUMMARY_ID	5
QUALITY_ID	7
PROGRESS_ID	6
CREATION_DATE	19-MAY-98
REVIEW_DATE	not filled
REVISION_DATE	19-MAY-98
FUTURE_REVIEW_DATE	not filled
Table SUMMARY	.
SUMMARY_ID	5
SUMMARY	Operational analysis are affected by major changes in models, analysis technique, asimilation and observation usage, which are an essential product of research and progress. The ECMWF (European Centre for Medium-Range Weather Forecasts) performed a consistent re-analysis of atmospheric data using a „frozen“ production system. The ECMWF Re-Analysis (ERA) Project produced a new, validated 15 year dataset of assimilated data for the period 12/1978-02/1994. The assimilation scheme consisted mainly of the Integrated Forecast System (IFS) version of the ECMWF forecast model with T106 resolution on 31 vertical hybrid levels. An intermitted statistical (optimum interpolation) analysis with 6 hour cycling was performed. It includes a diabatic, non-linear normal mode initialisation (5 vertical modes). This data set consists of an interpolated dataset (T42, originally T106) of monthly means of initialized surface fields.
Table ENTRY_TYPE	.
ENTRY_TYPE_ID	3
ENTRY_TYPE	Dataset
ENTRY_TYPE_DESCR	Contains data (description) which belongs to a superior experiment.
Table ENTRY_CONNECT	n/a
Table CONNECT_TYPE	n/a
Table KEY_CONNECT	.

Table: Dataset Metadata

GENERAL_KEY_ID	1
ENTRY_ID	144
GENERAL_KEY_ID	2
ENTRY_ID	144
Table GENERAL_KEY	
GENERAL_KEY_ID	3
GENERAL_KEY	Atmopsheric Pressure
Table CAMPAIGN	n/a
Table PROJECT	n/a
Table DIF_ENTRY	.
ENTRY_ID	6187
DIF_TITLE	n/a
DIF_NAME	ERA15_SFCFC00_MSLP
Block: STATUS	
Table QUALITY	.
QUALITY_ID	7
ACCURACY_REPORT	assimilated data
CONSISTENCY_REPORT	data was checked
COMPLETENESS_REPORT	no gaps in records
HORIZONTAL_ACC_REPORT	T42 model accuracy
VERTICAL_ACC_REPORT	T42 model accuracy
Table PROGRESS	.
PROGRESS_ID	6
PROGRESS_DESCR	complete, processed
Table PROCESS_STEP	.
ENTRY_ID	6187
PROCESS_DESCR	Monthly average and transformation to T42 model grid have been performed.
PROCESS_DATE	n/a
Block: ATTRIBUTE	
Table PARAMETER	.
ENTRY_ID	6187
TOPIC_ID	1003
UNIT_ID	4
SPATIAL_DATA_ORG_ID	3
AGGREGATION_ID	1
DATA_ORG_ID	2
DATA_ACCESS_ID	1
Table TOPIC	.
TOPIC_ID	1003
TOPIC_NAME	Mean Sea Level Pressure
TOPIC_ACRONYM	MSLP
TOPIC_DESCR	Surface pressure which is reduced to the height of normal zero (sea level)
TOPIC_POINTER	17
TOPIC_LEVEL	2
Table UNIT	.

Table: Dataset Metadata

UNIT_ID	4
UNIT_NAME	Pascal
UNIT_ACRONYM	Pa
UNIT_DESCR	SI unit for pressure
UNIT_ID	8
UNIT_NAME	Degree
UNIT_ACRONYM	deg
UNIT_DESCR	Unit for geographical length
UNIT_ID	10
UNIT_NAME	Metre
UNIT_ACRONYM	m
UNIT_DESCR	SI unit for length
Table AGGREGATION	.
AGGREGATION_ID	1
AGGREGATION_DESCR	Monthly Mean
Block: PUBLICATION	
Table REFERENCE	n/a
Table REF_TYPE	n/a
Table CITATION	n/a
Table CITATION_TYPE	n/a
Table PRESENTATION	n/a
Block: CONTACT	
Table CONTACT	.
ENTRY_ID	6187
INSTITUTE_ID	196
PERSON_ID	299
CONTACT_TYPE_ID	2
ENTRY_ID	6187
INSTITUTE_ID	dkrz
PERSON_ID	2
CONTACT_TYPE_ID	3
ENTRY_ID	6187
INSTITUTE_ID	dkrz
PERSON_ID	1
CONTACT_TYPE_ID	5
Table CONTACT_TYPE	.
CONTACT_TYPE_ID	2
CONTACT_TYPE	Originator
CONTACT_TYPE_ID	3
CONTACT_TYPE	Distributor
CONTACT_TYPE_ID	5
CONTACT_TYPE	Metadata
Table PERSON	.
PERSON_ID	299
FIRST_NAME	Rex

Table: Dataset Metadata

SECOND_NAME	n/a
LAST_NAME	Gibson
TITLE	Dr.
INSTITUTE_ID	196
TELEPHONE	not filled
FAX	not filled
URL	n/a
EMAIL	not filled
PERSON_ID	2
FIRST_NAME	Michael
SECOND_NAME	n/a
LAST_NAME	Lautenschlager
TITLE	Dr.
INSTITUTE_ID	1
TELEPHONE	+49 40 41173 297
FAX	+49 40 41173 400
URL	n/a
EMAIL	lautenschlager@dkrz.de
PERSON_ID	1
FIRST_NAME	Ulrich
SECOND_NAME	n/a
LAST_NAME	Cubasch
TITLE	Dr.
INSTITUTE_ID	1
TELEPHONE	not filled
FAX	not filled
URL	n/a
EMAIL	not filled
Table INSTITUTE	
INSTITUTE_ID	196
INSTITUTE_NAME	European Centre for Medium-Range Weather Forecasts
INSTITUTE_ACRONYM	ECMWF
DEPARTMENT_NAME	n/a
DEPARTMENT_ACRONYM	n/a
COUNTRY	United Kingdom
STATE_OR_PROVINCE	England
PLACE	Reading
STREET	Shinfield Park
STREET_POSTAL_CODE	RG2 9AX
POBOX	n/a
POBOX_POSTAL_CODE	n/a
URL	http://www.ecmwf.int/
ADDITIONAL_INFO	n/a
INSTITUTE_ID	1
INSTITUTE_NAME	Deutsches Klimarechenzentrum GmbH

Table: Dataset Metadata

INSTITUTE_ACRONYM	DKRZ
DEPARTMENT_NAME	Climate Model and Data Support
DEPARTMENT_ACRONYM	n/a
COUNTRY	Germany
STATE_OR_PROVINCE	Hamburg
PLACE	Hamurg
STREET	Bundesstrasse 55
STREET_POSTAL_CODE	D-20146
POBOX	n/a
POBOX_POSTAL_CODE	n/a
URL	http://www.dkrz.de/
ADDITIONAL_INFO	n/a
Block: COVERAGE	
Table LOCATION_CONNECT	.
ENTRY_ID	6187
LOCATION_ID	728
Table LOCATION	.
LOCATION_ID	728
LOCATION_NAME	global
LOCATION_POINTER	0
LOCATION_LEVEL	1
REF_DATE	n/a
MIN_LAT	-90
MAX_LAT	90
MIN_LON	-180
MAX_LON	180
Table COVERAGE	.
ENTRY_ID	6187
TEMPORAL_COVERAGE_ID	9
SPATIAL_COVERAGE_ID	9
Table TEMPORAL_COVERAGE	.
TEMPORAL_COVERAGE_ID	9
START_YEAR	1978
START_MONTH	12
START_DAY	01
STOP_YEAR	1994
STOP_MONTH	02
STOP_DAY	28
CURRENTNESS_REF_ID	5
Table CURRENTNESS_REF	.
CURRENTNESS_REF_ID	5
CURRENTNESS_REF_DESCR	calendrical
Table SPATIAL_COVERAGE	.
SPATIAL_COVERAGE_ID	9
MIN_LAT	-90

Table: Dataset Metadata

MAX_LAT	90
MIN_LON	-180
MAX_LON	180
MIN_ALTITUDE	0
MAX_ALTITUDE	0
MIN_ALT_UNIT_ID	10
MAX_ALT_UNIT_ID	10
Block: SPATIAL INFORMATION	
Table SPATIAL_DATA_ORG	.
SPATIAL_DATA_ORG_ID	3
REFERENCE_METHOD	Raster
Table SPATIAL_REFERENCE	.
ENTRY_ID	144
HOR_COORD_SYS_ID	2
VER_COORD_SYS_ID	3
Table HOR_COORD_SYS	.
HOR_COORD_SYS_ID	2
SYS_DESCR	LON and LAT: geographic
Table VER_COORD_SYS	.
VER_COORD_SYS_ID	3
SYS_DESCR	Altitude: length
Block: DISTRIBUTION	
Table DISTRIBUTION	.
ENTRY_ID	6187
ACCESS_TYPE_ID	9
DATA_SIZE	3012180
FORMAT_ID	16
FEES_ID	2
ACCESS_CONSTRAINT_ID	8
USE_CONSTRAINT_ID	8
Table FORMAT	.
FORMAT_ID	26
FORMAT_ACRONYM	GRIB
FORMAT_DESCR	WMO format GRIdded Binary data
Table ACCESS_TYPE	.
ACCESS_TYPE_ID	9
ACCESS_ACRONYM	DB
ACCESS_DESCR	disc: DB tables
Table FEES	.
FEES_ID	2
FEES_ACRONYM	none
FEES_DESCR	n/a
Table ACCESS_CONSTRAINT	.
ACCESS_CONSTRAINT_ID	8
CONSTRAINT_DESCR	Access is restricted to German scientists.

Table: Dataset Metadata

Table USE_CONSTRAINT	.
USE_CONSTRAINT_ID	8
CONSTRAINT_DESCR	Data use is only permitted for science.
Modul: DATA_ORGANIZATION	
Table DATA_ORG	.
DATA_ORG_ID	2
DATA_ORG_DESCR	Data are stored as real-time series of horizontal layers at fixed height.
SPACE_ID	2
TIME_ID	1
Table SPACE	.
SPACE_ID	2
SPACE_DESCR	Space storage is organized horizontal layers of the ECHAM(T42) grid at fixed height, longitude: 128 equidistant grid points, latitude: 64 points in Gaussian grid.
X_SCALE_ID	1
Y_SCALE_ID	2
Z_SCALE_ID	3
POINT_SET_ID	0
Table SCALE	.
SCALE_ID	1
SCALE_NAME	cyclic, global longitude scale
NUM_POINTS	128
DIM_TYPE__ID	1
SCALE_ID	2
SCALE_NAME	Gaussian, global latitude scale
NUM_POINTS	64
DIM_TYPE__ID	2
SCALE_ID	3
SCALE_NAME	vertical level
NUM_POINTS	1
DIM_TYPE__ID	3
Table SCALE_CONNECT	.
SCALE_ID	1
POSITION_ID	1
SEQUENCE_NO	1
SCALE_ID	1
POSITION_ID	2
SEQUENCE_NO	128
SCALE_ID	2
POSITION_ID	3
SEQUENCE_NO	1
SCALE_ID	2
POSITION_ID	4
SEQUENCE_NO	2

Table: Dataset Metadata

The SCALE_ID + POSITION_ID + SEQUENCE_ID have to be repeated for the given positions.	
POSITION_ID	5 65
SEQUENCE_NO	3 63
SCALE_ID	2
POSITION_ID	66
SEQUENCE_NO	64
SCALE_ID	3
POSITION_ID	67
SEQUENCE_NO	1
Table POSITION	.
POSITION_ID	1
POSITION_NAME	Start point of equidistant grid
POSITION_VALUE	0
UNIT_ID	8
POSITION_ID	2
POSITION_NAME	End point of equidistant grid
POSITION_VALUE	357.1875
UNIT_ID	8
POSITION_ID	3
POSITION_NAME	Start point of Gaussian latitude grid (most northern)
POSITION_VALUE	87.863
UNIT_ID	8
POSITION_ID	4
POSITION_NAME	2. point of Gaussian latitude grid (most northern)
POSITION_VALUE	85.096
UNIT_ID	8
The POSITION_ID + POSITION_VALUE + ... have to be repeated for the given positions.	
POSITION_ID	5 65
POSITION_VALUE	82.312 ...... -85.096
POSITION_ID	66
POSITION_NAME	End point of Gaussian latitude grid (most northern)
POSITION_VALUE	-87.863
UNIT_ID	8
POSITION_ID	67
POSITION_NAME	The vertical level.
POSITION_VALUE	0
UNIT_ID	10
Table DIM_TYPE	.
DIM_TYPE__ID	1
DIM_TYPE_NAME	equidistant grid
DIM_TYPE__ID	2
DIM_TYPE_NAME	irregular grid
DIM_TYPE__ID	3

Table: Dataset Metadata

DIM_TYPE_NAME	single position
Table TIME	.
TIME_ID	1
TIME_NAME	real time series of monthly means
NUM_POINTS	183
DIM_TYPE__ID	1
Table TIME_CONNECT	.
TIME_ID	1
MOMENT_ID	1
TIME_ID	1
MOMENT_ID	2
Table MOMENT	.
MOMENT_ID	1
MOMENT_NAME	start month of time series
YEAR	1978
MONTH	12
DAY	n/a
HOUR	n/a
MINUTE	n/a
SECOND	n/a
UTC_DIFFERENCE	0
MOMENT_ID	2
MOMENT_NAME	end month of time series
YEAR	1994
MONTH	02
DAY	n/a
HOUR	n/a
MINUTE	n/a
SECOND	n/a
UTC_DIFFERENCE	0
Modul: DATA_ACCESS	
Table DATA_ACCESS	.
DATA_ACCESS_ID	1
ACCESS_STRUCTURE_ID	2
STORAGE1_ID	1
STORAGE2_ID	2
STORAGE3_ID	3
STORAGE4_ID	4
REC_STRUCTURE_ID	1
MODIFICATION_DATE	n/a
Table ACCESS_STRUCTURE	.
ACCESS_STRUCTURE_ID	2
ACCESS_STRUCTURE_NAME	DB internal
ACCESS_STRUCTURE_DESCR	Data are stored as BLOB‘s in DB table.

Table: Dataset Metadata

Table STORAGE	.
STORAGE_ID	1
STORAGE_NAME	RDBMS ORACLE
STORAGE_ID	2
STORAGE_NAME	cldb.dkrz.de
STORAGE_ID	3
STORAGE_NAME	ERA15_SFCFC00_MSLP
STORAGE_ID	4
STORAGE_NAME	cera
Table RECORD_STRUCTURE	.
RECORD_STRUCTURE_ID	1
STRUCTURE_NAME	Time series of horizontal fields (=BLOB storage) at specified level for specified parameter
DIM1_TYPE	SCALE
DIM1_ID	1
DIM2_TYPE	SCALE
DIM2_ID	2
DIM3_TYPE	TIME
DIM3_ID	1
DIM4_TYPE	SCALE
DIM4_ID	3
DIM5_TYPE	TOPIC
DIM5_ID	1003