Long-term Archiving of Climate Model Data at WDC Climate and DKRZ

Michael Lautenschlager [(1)], Wolfgang Stahl [(2)]

[(1)] ***World Data Center Climate (WDCC)***
Max-Planck-Institute for Meteorology
Bundesstrasse 53, D-20146 Hamburg, Germany
EMail: michael.lautenschlager@zmaw.de

[(2)] ***German Climate Computing Centre (DKRZ)***
Bundesstrasse 55, D-20146 Hamburg, Germany
EMail: stahl@dkrz.de

ABSTRACT

The computing capabilities for production of Earth system model data are growing faster than the prices for mass storage media sink. If the archive philosophy left unchanged during the migration to the next compute server generation consequently the amount of money for long-term archiving rises and the total amount of money for archiving tends to exceed the money which is left for compute services. At WDCC (World Data Center Climate) and DKRZ (German Climate Computing Centre) a new concept for long-term archiving has been developed which addresses this problem and improves the overall confidence in the long-term archive.

The new archive concept separates data storage with expiration date at the scientific project level and the documented long-term archive. The transition process to the new archive concept already started and at the end we expect to have a completely documented long-term archive with a searchable data catalogue. This archive concept is supported by a four level storage hierarchy which reflects the lifetimes of the different data categories.

Keywords: climate model data, long-term archiving, data preservation, quality assurance, usability

INTRODUCTION

The magnitude of data output from climate models depends on spatial resolution, temporal output intervals, number of variables and the data storage format. Increase of installed compute power motivates finer spatial (comp. Fig. 2) and temporal model resolution and integration of additional physical and chemical processes into the models (increasing number of output variables, comp. Fig. 1).

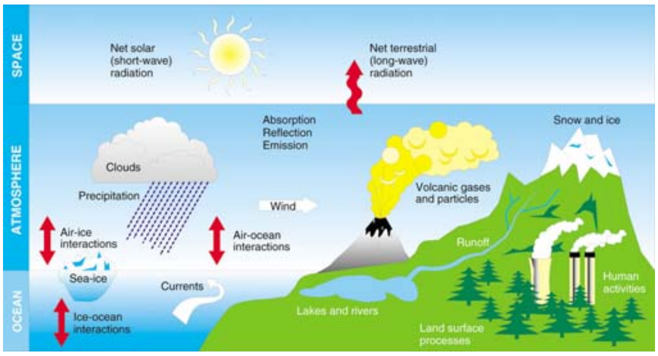

Figure 1: Processes in climate system which are taken into account in climate models

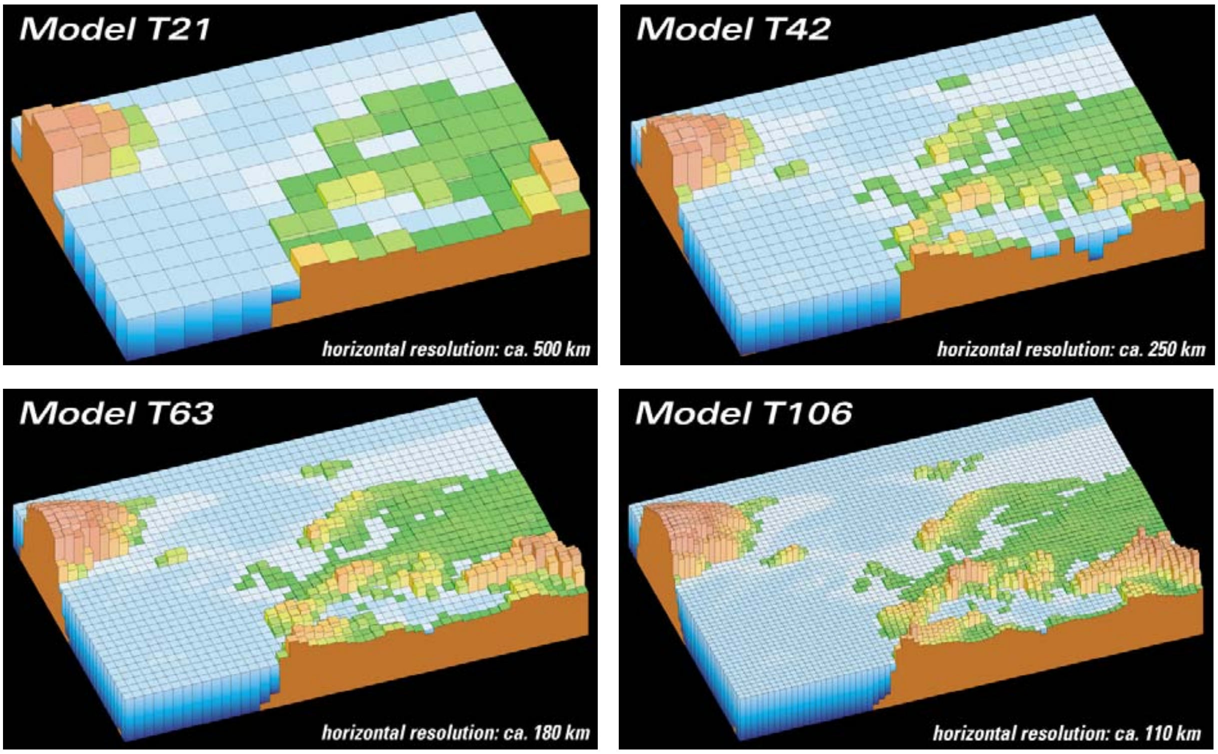

Figure 2: North Atlantic sector of the land-ocean distribution of the Hamburg climate model in four different horizontal resolutions providing mesh sizes of 500 km, 250 km, 180 km and 110 km.

In the past the model data output at DKRZ increased linearly with the installed compute power and there is no reason for change in future. Presently at DKRZ all produced data migrate almost directly into the archive system and finally they are stored on tapes. Data for the long-term archive and from project data management are currently not separated. All data are archived on tapes without expiration date (comp. Fig. 3). Continuation of this strategy will result in management, financial and acceptance problems. Data service costs which supersede the costs for compute services are not accepted by DKRZ shareholders and by the scientific community.

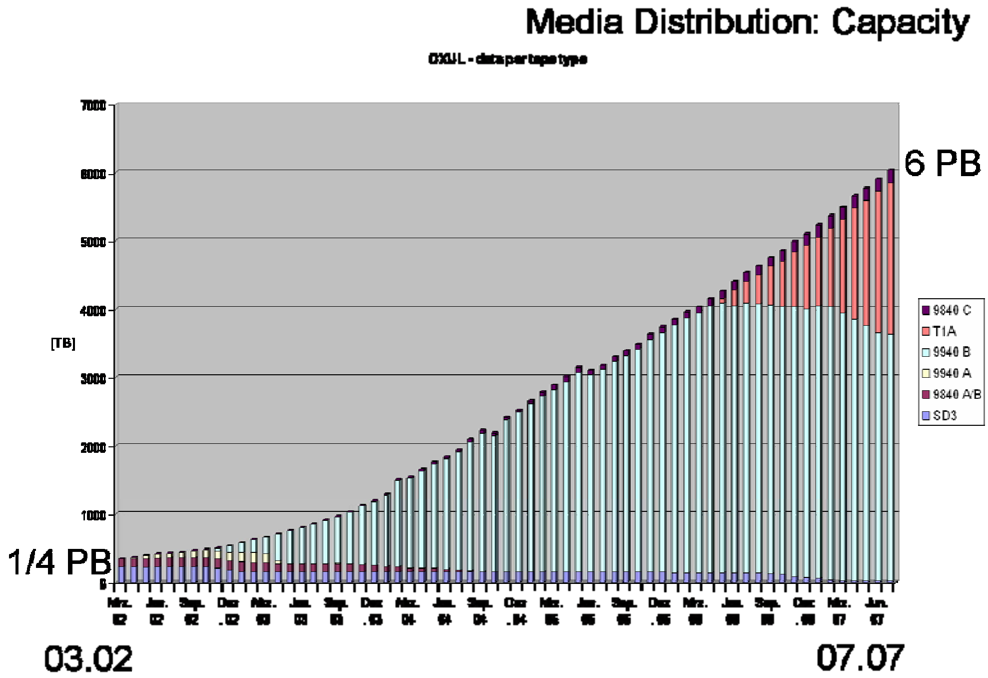

Figure 3: Development of the mass storage at DKRZ from February 2002 until July 2007 including the migration of tape technology.

The next generation of compute power at DKRZ is projected conservatively as an increase by a factor of 10 – 15 or optimistically as an increase by a factor of 30. Therefore the data archive rate will grow linearly from 1 PB/year to 13 PB/year if the data archive policy at DKRZ left unchanged (comp. Fig. 4).

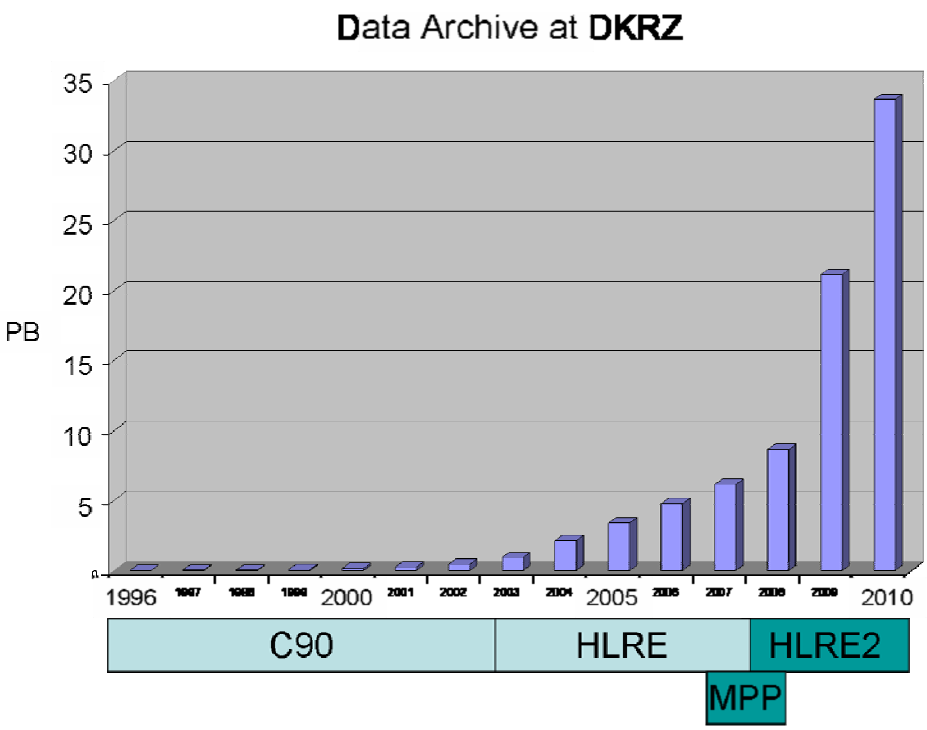

Figure 4: Development of the total data archive at DKRZ from 1996 extrapolated to 2010. Below the graph the related compute server architectures are given: Cray C90 (C90), NEC SX-6 (HLRE), Sun Cluster (MPP) and the next generation (HLRE2).

The technical challenges of data management can be handled theoretically but the financial effort for data services will not be balanced any longer with the effort for compute services. Although the media costs per TB storage capacity decrease continuously the total media costs increase because of stronger increase in compute power and connected data production rates (comp. Fig. 5). The fraction for data services in the operational costs of DKRZ tends to grow from presently 25% to 50% or more. The total amount of envisaged archive data overcompensate the continuously decreasing unit costs for storage media.

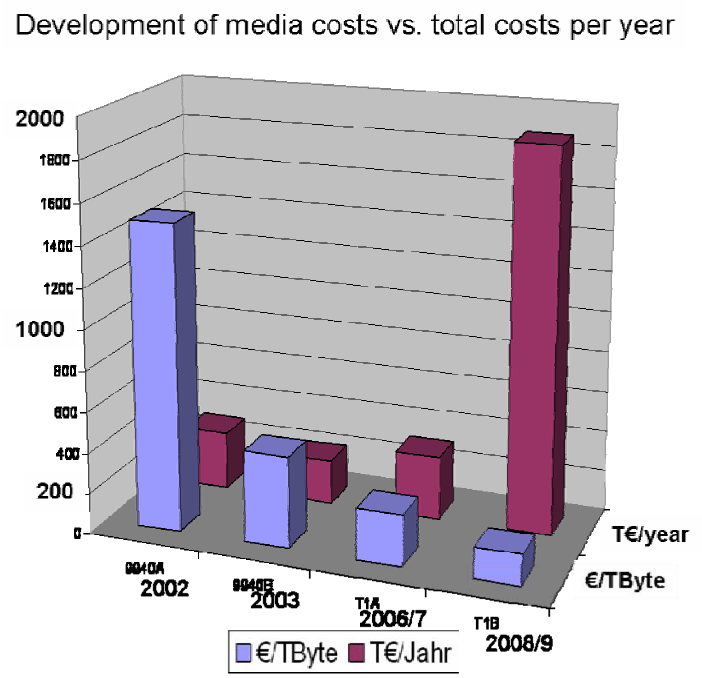

Figure 5: Development of media costs per TB against development of total media costs at DKRZ in 2002, 2003, 2006/7 and 2008/9. The years correspond with the tape media notification.

Additionally to the financial aspect the effort for data preservation, quality assurance and securing usability increases with the total size of the mass storage archive and requires more and more effort for the case of linear data increase with installed compute power. This development cannot be balanced with the envisaged fraction of costs between compute and data services. Therefore the long-term archiving strategy has to be modified with the installation of the of the next generation compute service at DKRZ. Basic idea is to separate project data with a limited lifetime and data suitable for long-term archiving and to impose an aware, scientific decision for long-term archiving which takes into account the rules for good scientific practise [1].

LONG-TERM ARCHIVING STARTEGY FOR THE 50 TFLOP/S – CLASS COMPUTE SERVICE

The activity focus of DKRZ is Earth system modeling namely model development and simulations of the past, present and future climate. Scientists from DKRZ shareholders and from various research institutes are working on the system. The produced data at DKRZ can be divided into three classes which reflect also the different life cycles: test data, project data and final results.

- Test data with life cycle of weeks to months

  Test data are connected with model code development (implementation of new numerical schemes or integration of new parameterizations and processes). A large number of short models runs are performed in order to verify the model code modifications. These test data have only a short life time in the order of weeks or months until the model code modifications have been tested for technical correctness. Next development step is then the consideration of the scientific question behind the imposed model changes. The work is organized in projects at DKRZ and resources are allocated to these projects after review.

- Project data with life cycle of 3-5 years, i.e. run time of a scientific project

  The modified numerical model is used after technical verification for answering the scientific question which motivated the implied code changes. In a scientific project the number of model runs decreases compared to the test phase but the model run time increases together with the turn around time of the numerical model on the compute server. Typical run times for scientific projects are 3 – 5 years and this is also the life cycle of most of the project data. But not all data will probably go into the final results category for long-term archiving.

- Final results life cycle: 10 years and longer

  Final results from DKRZ projects or data which are especially produced for international projects (like climate scenarios calculations for the IPCC assessment reports) are usually part of scientific publications. Following the rules of good scientific practice these data have to be available and accessible for at least ten years in order to allow for verification of published results. These data are part of the general knowledge creation process and they will be used as well in interdisciplinary scientific fields. These two application aspects of long-term archiving especially the interdisciplinary usage impose stronger requirements on data preservation, quality assurance and usability enabling than for the other two data categories. Final results and data products cannot be easily reproduced after a number of years. Data users from the wider audience do not have the background how to deal with numerical model data. This imposes documentation requirements on the long-term archive.

Access constraints of the three standard data categories are reflected in a four step data hierarchy. The expected life cycles are represented in data expiration dates. Data which crosses the expiration date will be removed from the system after warning. The four levels of the DKRZ data hierarchy will be (comp. Fig. 6 for today's distribution):

- Temp(orary)

  The "temp" level storage consists of fast discs and built the scratch space for running calculations. Data have to be moved or deleted after end of the calculation. This level presently exists with the scratch disc space of the compute server.

- Work

  The "work" level provides limited disc space for each individual DKRZ project for the project run time. The administration has to be done by the project itself. This level is presently not implemented at DKRZ in full complexity. The basic idea is to provide enough disc space to a project that actual calculations and evaluations can be performed without migration of these data to the tapes of the mass storage archive. This level presently exists in a small version with the GFS (NEC Global File System).

- Arch(ive)

  The "arch" level assigns tape archive space to the projects. Single copies of project data files are stored here. It represents an undocumented file based tape archive for work group application.

  The idea is to provide mass storage archive space beyond the disc quota to projects for those data which are not in the direct evaluation process but which have to be available for later use in the project. These data are stored in the mass storage archive with expiration date "project end plus one year". This data will be (automatically) removed from the system after the expiration date is crossed and if no reaction to the warning message(s) occurred. If these data should be archived longer they must be moved to the long-term archive level.

- Docu(mentation)

  The "docu" level represents the long-term data archive with life cycles of ten years and longer according to library rules. Data will be stored on the tapes of the DKRZ mass storage archive together with a complete documentation and a second copy for security. The WDC Climate with its searchable data catalogue builds presently the core of this archive level. For the next compute server generation the WDCC documentation concept will be enlarged to all data in this hierarchy level either in files or in the database system. In the future DKRZ will have a completely documented long-term mass storage archive. These data in the "docu" hierarchy level are fixed and no longer matter of change.

  The concept is to provide an infrastructure for those model results which are part of scientific publications and for data products which are generated for and disseminated to a wider audience. The service levels for data preservation, quality assurance and usability enabling are higher than for the other hierarchy levels because data cannot easily reproduced and an inexperienced user audience has to be expected.

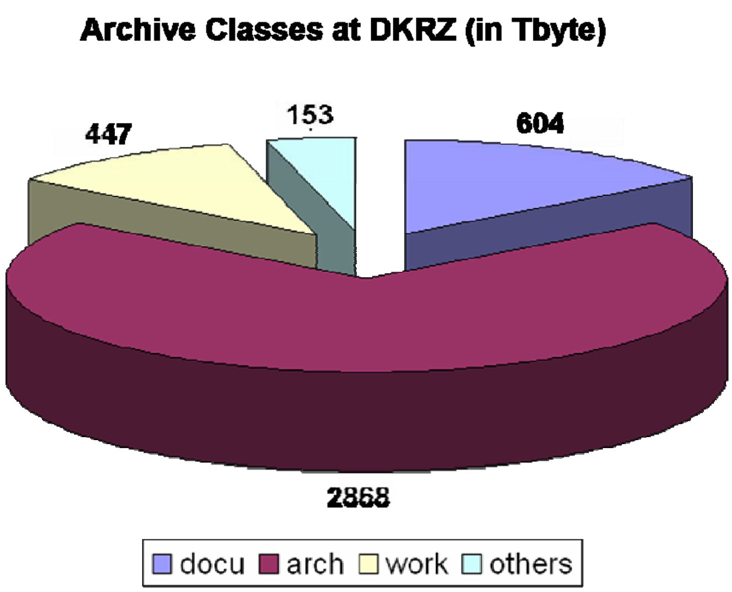

Figure 6: Tape archive data distribution on DKRZ archive classes at begin of 2007

New in the DKRZ mass storage archive concept are the separation between project data and long-term archiving, the expiration date for files in the project data space ("arch" hierarchy level) and the data documentation requirements for the long-term data archive ("docu" hierarchy level). These fit the experience that data without documentation are only numbers and contain no information when they left the work group regime and the unstructured documentation disappeared.

Documentation requirements are accomplished by using the metadata model and the catalogue implementation of the WDC Climate. The CERA metadata model (Climate and Environmental data Retrieval and Archiving) [2] contains information to browse, to identify and to use climate data which are stored in the database system and outside in flat files without attaching the climate data physically. The use metadata are part of the data model and allow for the specification of data processing without downloading data from tapes. The complete information on variables, spatial-temporal resolution as well as information on experimental design, responsible persons and data quality can be searched and retrieved from the WDCC data catalogue.

The CERA data model fits the ISO 19115 metadata standard [3] and the WDCC data catalogue is therefore suitable to be harvested in data portals and international data federations which are based on this description standard. Metadata for CERA are expected as XML file which is loaded into the database tables of the catalogue system. Missing information has to be completed by hand during the metadata review process. The XML scheme for CERA metadata is provided with the data catalogue input interface (http://input.wdc-climate.de). Metadata export is as well implemented by XML. The corresponding XML file can be transformed with an XML style sheet into the metadata format which is required by the data federation.

Climate data in the docu storage hierarchy level are archived as part of the CERA database system or as flat files which are referenced by the CERA database system. Data products which are suitable for international research activities are stored as BLOB entries in the database tables together with their corresponding metadata (comp. Fig. 7 for details on climate model data structure). These data are open for web-based data access and can be downloaded by a standard web-browser at the granularity of one, a few or many table entries. Data which are stored in the long-term archive for reference are archived as flat files. These files are documented in the CERA catalogue of the WDCC but cannot be directly accessed by the graphical user interface and standard web technology. In any case data in the “docu” hierarchy level are no longer matter of change because they are freely available at least for the scientific audience and at the archive level we have no control of their usage in for example scientific publications. Changing data in this long-term archive contradicts the rules of good scientific practice. For the case that data errors imply changes of a data entity this can be realized with a new documented version of the corresponding data entity.

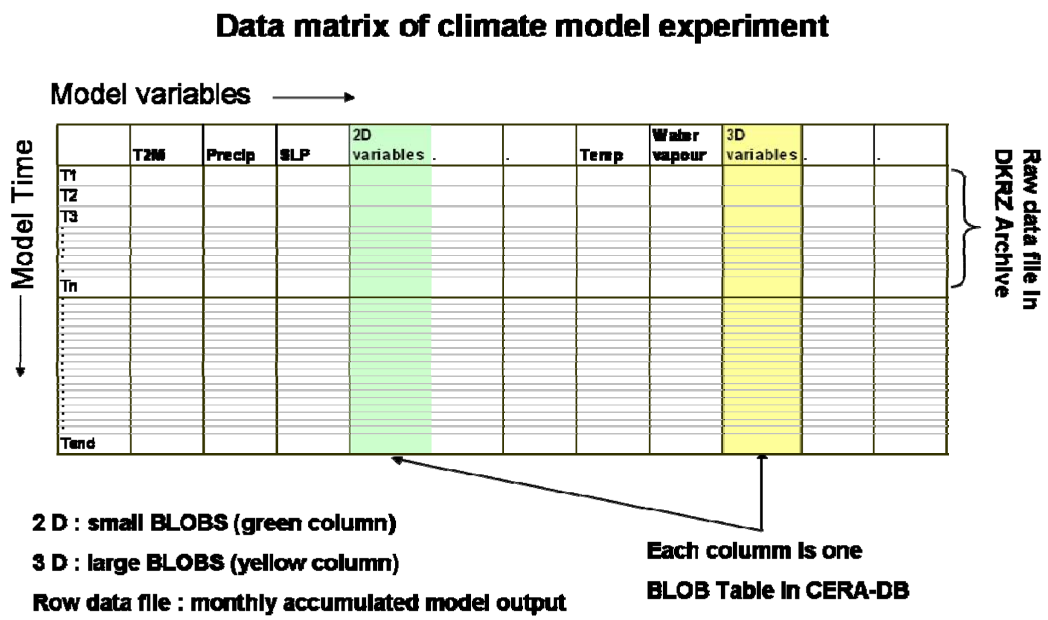

Figure 7: In the climate model data matrix the rows represent the direct model raw data output and the columns give the WDCC database table storage of time series of individual variables.

The primary data publication service is offered by the WDCC as an additional service for data of general interest which should be referenced directly in scientific publications and which are then searchable in standard library catalogues together with scientific publications. The primary data publication process has been developed together with the Technical Information Library, Hannover (TIB) and has been implemented in the STD-DOI profile (Scientific and Technical Data - Digital Object Identifier, http://www.std-doi.de) [4], [5]. Primary data which are identified as independent data entities in the context of scientific literature are suitable for the STD-DOI publication process (comp. Fig. 8 for workflow). These data together with their metadata pass a review process and a process of quality assurance before the metadata for electronic publication and a persistent identifier (DOI) are assigned. The metadata for electronic publication and the corresponding identifier are registered in the library catalogue of the TIB. Now the published primary data can be searched and accessed together with standard scientific publications. The metadata of the STD-DOI profile are open for harvesting and for integration in alternative information systems.

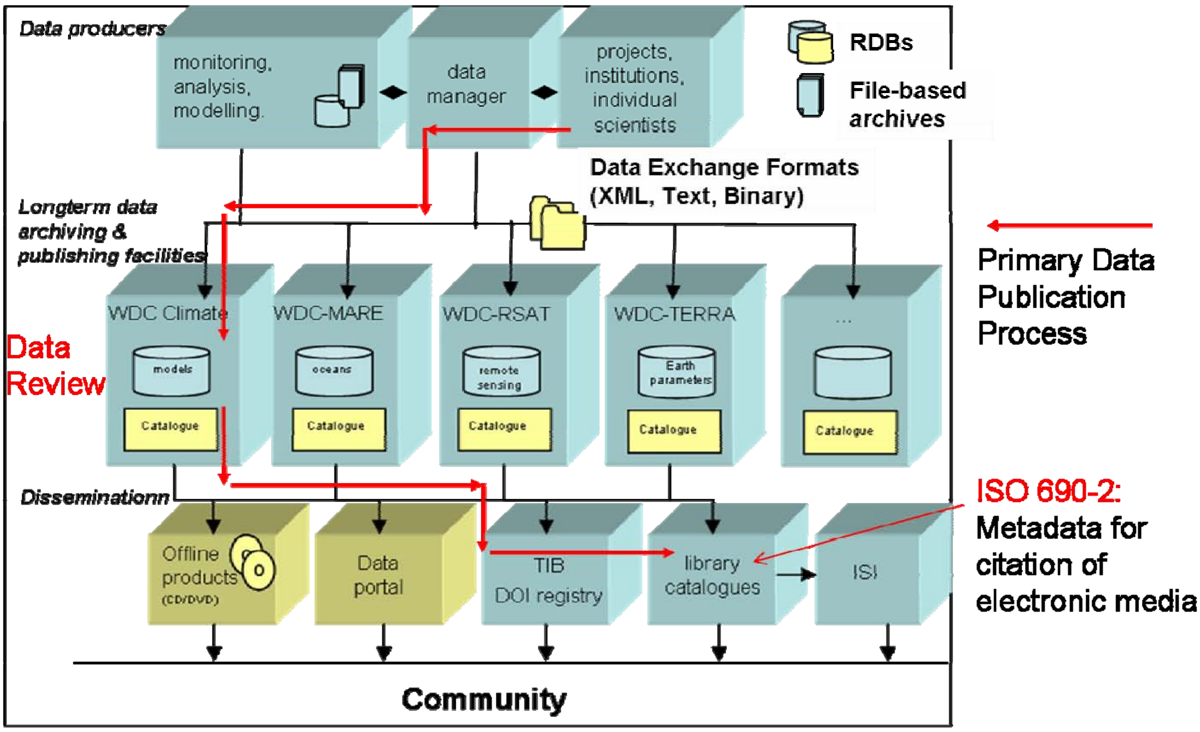

Figure 8: Data and process flow in the STD-DOI publication process

The STD-DOI concept integrates scientific primary data in more general e-science information systems. The data entities of the WDCC which are currently registered in the TIB catalogue and can be obtained from TIBORDER (URL: http://tiborder.gbv.de/psi/DB=2.63/LNG=DU/) using the keyword “WDCC”. The search results are presented in Fig. 9. The corresponding DOI can be resolved directly in the global handle system and allows for direct data access independent from the specific URL which provides the data storage location.

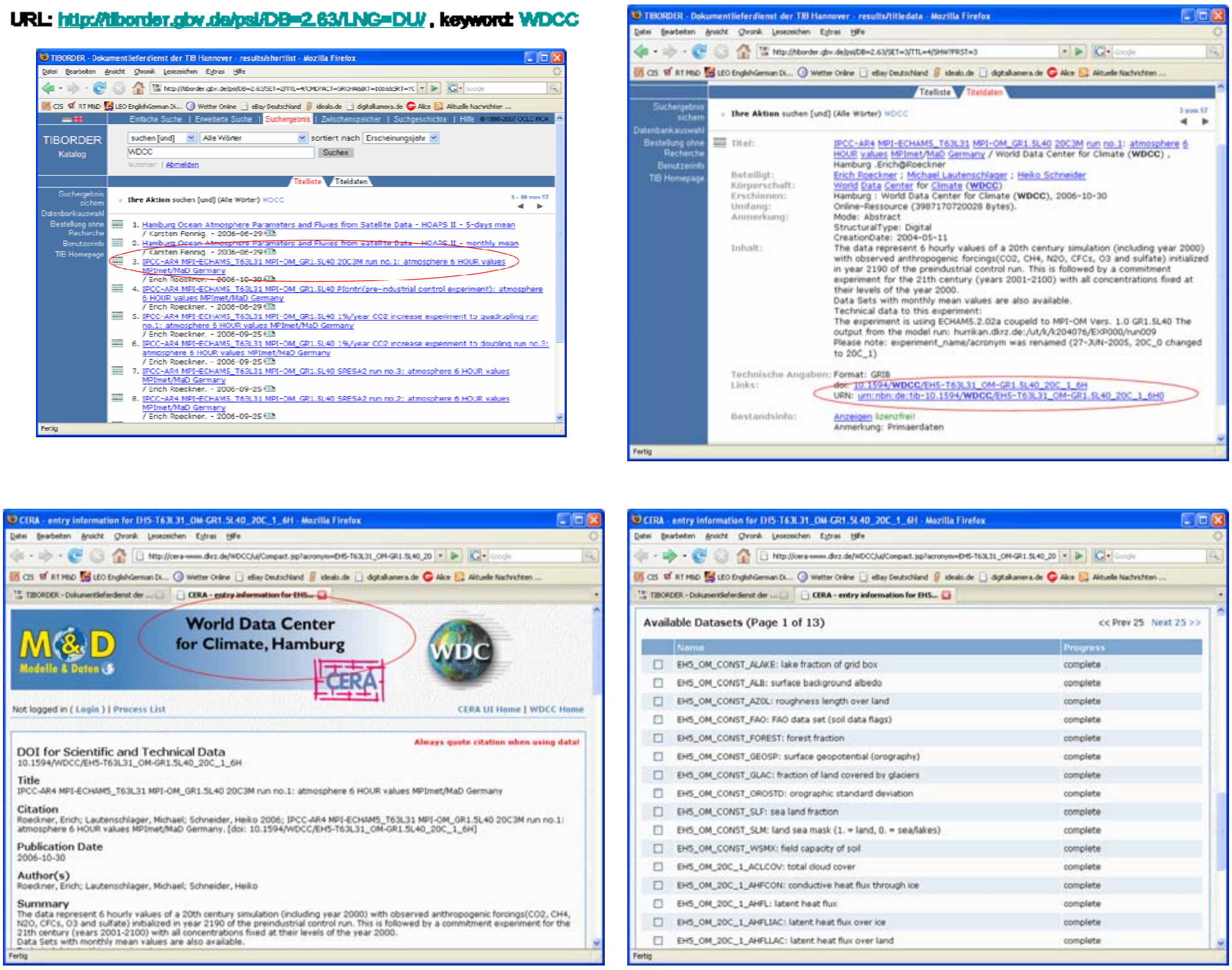

Figure 9: The top two panels provide the WDCC representation in the TIB library catalogue, the bottom two panels show the scientific metadata representation in the WDCC metadata catalogue.

DATA PRESERVATION, QUALITY ASSURANCE AND USABILITY ENABLING

The new and elaborated long-term archiving concept of WDCC and DKRZ is connected with more intensive stewardship of data in the “docu” hierarchy level. This naturally limits the amount of data which can be handled.

Bit stream preservation will be secured by secondary tape copies on independent tapes at a separate location. The secondary copies should employ technology of an independent vendor. Additionally the total number of tape accesses is recorded. If the maximum number of accesses is reached the old tape is copied to new one and the old tape is removed from the silo. Presently secondary tape copies using identical technology and tape refreshment is installed. This will be integrated into the DKRZ data flows (comp. Fig. 10)

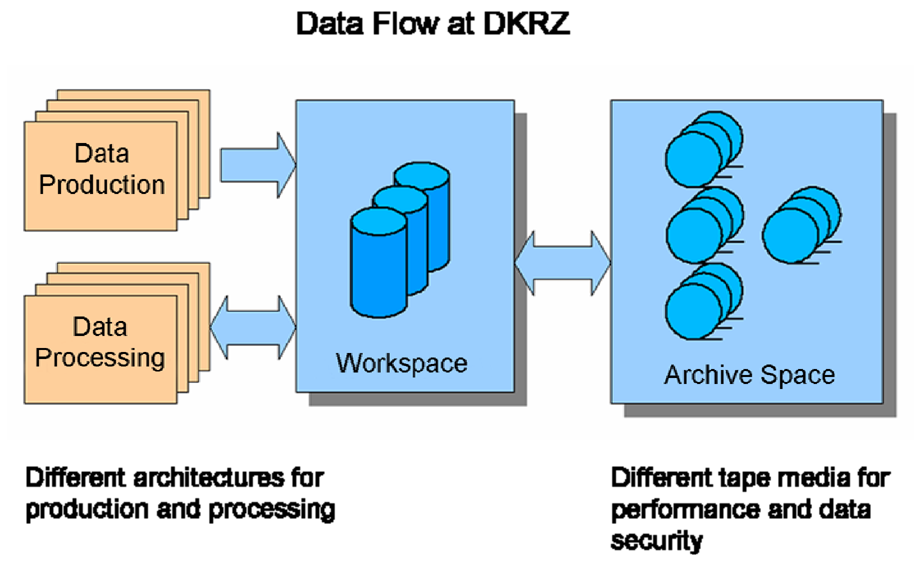

Figure 10: DKRZ data flow adapted to the envisaged data archive hierarchy levels

Quality assurance of numerical model results is complex because of the huge amount of data which are produced by high performance computers. Not every single number in the output can be inspected. Alternatively spot tests are performed at different levels of complexity.

Quality assurance of numerical model output basically contains semantic and syntactic examinations. Semantic examination means considering the behavior of a numerical model itself compared to observations and to other models. This is mainly part of the scientific evaluation process. Results are normally documented in scientific literature. Syntactic examination means considering formal aspects of data archiving and ensure that the archiving is free of errors as far as possible. This includes for model data the examination of

- consistency between metadata and climate data,
- completeness of climate data,
- standard range of values and
- spatial and temporal ordering.

These proofs are time consuming but in the majority of cases they can be performed automatically. They have to be accomplished in order to gain confidence into the data archive.

Quality assurance in the WDCC and the DKRZ long-term archive is organized as a three stages process.

1. The semantic check is performed at the scientific level during the runtime of the corresponding scientific project in order to decide on the validity and usefulness of the actual model results. A positive decision is the basic criterion to migrate the data from the project data management (hierarchy level “arch”) into the long-term data archive (hierarchy level “docu”).
2. Data documentation and syntactic proof routines are conducted during the data integration process into the long-term archive. This is presently realized for data integration into the WDCC as spot tests of consistency, completeness, standard range and spatial and temporal order for the most important variables.
3. The most complete data examination is accomplished in connection with the STD-DOI data publication process. The publication process includes review of metadata and climate data. Presently the data review is performed in the WDCC as complete syntactic proof (stage two) of all time series which belong to the data entity under review. Additionally semantic proofs of

core variables which are often downloaded from the WDCC like the near surface climatology are examined in cooperation with the data originators (STD-DOI data authors). Normally two dimensional plots of model fields are compared by the data reviewers with the expected spatial distribution. This data review will be documented in the corresponding metadata of the WDCC as part of the review process itself.

The usability of the long-term archive will be improved by a complete searchable documentation of the climate data entities (database tables and flat files) in the catalogue system of the WDCC. Additionally the WDCC offers web-based data access in small granules for those climate data which are stored as individual two dimensional fields in the database tables. Presently the WDCC offers web-access to, more than 300 TBytes of climate data which are organized in 950 climate (model) data experiments. The climate data are stored in 120,000 database tables which contain 5.6 billion individual table entries. The average size of a single table entry (these BLOBs correspond to the data access granules) is then calculated to 60 kBytes.

Data usability has additionally to be supported on the technical level. Archive technology transfer must be downward compatible to keep old data technically readable in future. The related software has also to be migrated to new platforms. Data processing tools and data format access libraries are continuously needed to work even with older data from the long-term archive.

CONCLUSION

The new archive strategy at DKRZ includes expiration dates for project data. Only selected data are migrated into the long-term archive. The long-term archive integrates the scientific data management principles of the WDCC. Even with an expected compute power increase of a factor of 30 or more we expect to limit the file archive increase to a factor of 10 and the database increase of the WDCC to a factor of 5 due to the application of our modified long-term data archiving concept.

Limitations in long-term archiving will also improve the long-term data archive reliability because more effort can be spend on data stewardship instead on technical handling of huge amount of data. Stewardship includes archive documentation, bit-stream preservation, quality assurance and usability enabling.

REFERENCES

[1] Senate of the Max Planck Society: Rules for Good Scientific Practice. November 2000

[2] Lautenschlager, M., F. Toussaint, H. Thiemann and M. Reinke: The Cera-2 Data Model. Technical Report No. 15, DKRZ, Hamburg, 1998

[3] Technical Committee Geographic Information (TC 211): ISO 19115 – Geographic Information – Metadata. International Standards Oragnisation, 2003.

[4] Lautenschlager, M. und I. Sens: Konzept zur Zitierfähigkeit wissenschaftlicher Primärdaten. Information – Wissenschaft & Praxis, 54(2003) 463-466

[5] Klump, J., R. Bertelmann, J. Brase, M. Diepenbroek. H. Grobe, H. Höck, M. Lautenschlager, U. Schindler, I. Sens and J. Wächter: Data Publication in the Open Access Initiative. CODATA Data Science Journal, p79-83, 2006