2. SEDAN—THE GERMAN PALEOCLIMATE DATA CENTER FOR MARINE AND LACUSTRINE SEDIMENTS

M. Diepenbroek

At the Alfred Wegener Institute for Marine and Polar Research (AWI), a paleoclimate data center for marine and lacustrine sediments (SEDAN) will be set up within the next few years at the request of those scientists in Germany who work in the marine aspects of PAGES. The project is funded by the Federal Ministry for Research and Technology (BMFT). In Germany about 25 institutions sample and analyse marine and lacustrine sediments. With the foundation of new institutes, the organization of special research programmes, new or reconstructed research vessels, the amount of existing samples and data has increased at a great rate. At present about 50 km of sediment cores exist. The datacenter will guarantee the consistent and long-term storage of data derived from these samples and will make them available for the scientific community, especially for modeling groups working on paleoclimate.

Characteristics

Figure 1 shows essential attributes of the datacenter that were recognized during the conception phase. The first aspect in the data domain is the availability of data. Both historically important data sets and recent data sets should be available. Data quality, a crucial aspect in the data domain, can be estimated. Quality information is parameter dependant. In general, it might be sufficient to specify the method of measurement and calibration, the precision of the data, and—if possible—to give citations on data sets. In many cases however, additional and more specific information is needed. Problems in the data domain are caused by the variety of possible analyses, interpretations, calibrations and alterations in these fields.

The first crucial aspect in the structural domain is the usability of the system. Easy entry requirements, simple handling and good responsiveness will contribute to the effectiveness of the system. To cope with the growing requirements of users and with future developments in information technology, the system must be adaptable to improvement and enlargement. Problems in this field arise from the spatial distance between institutes and working groups and the different hard- and software environments. Legal aspects have also to be recognized. Copyright and similar questions are always controversial. The copyright question has to be solved in accordance with the users. Finally, there must be some benefit for the data producers to be motivated to store their data in the datacenter.

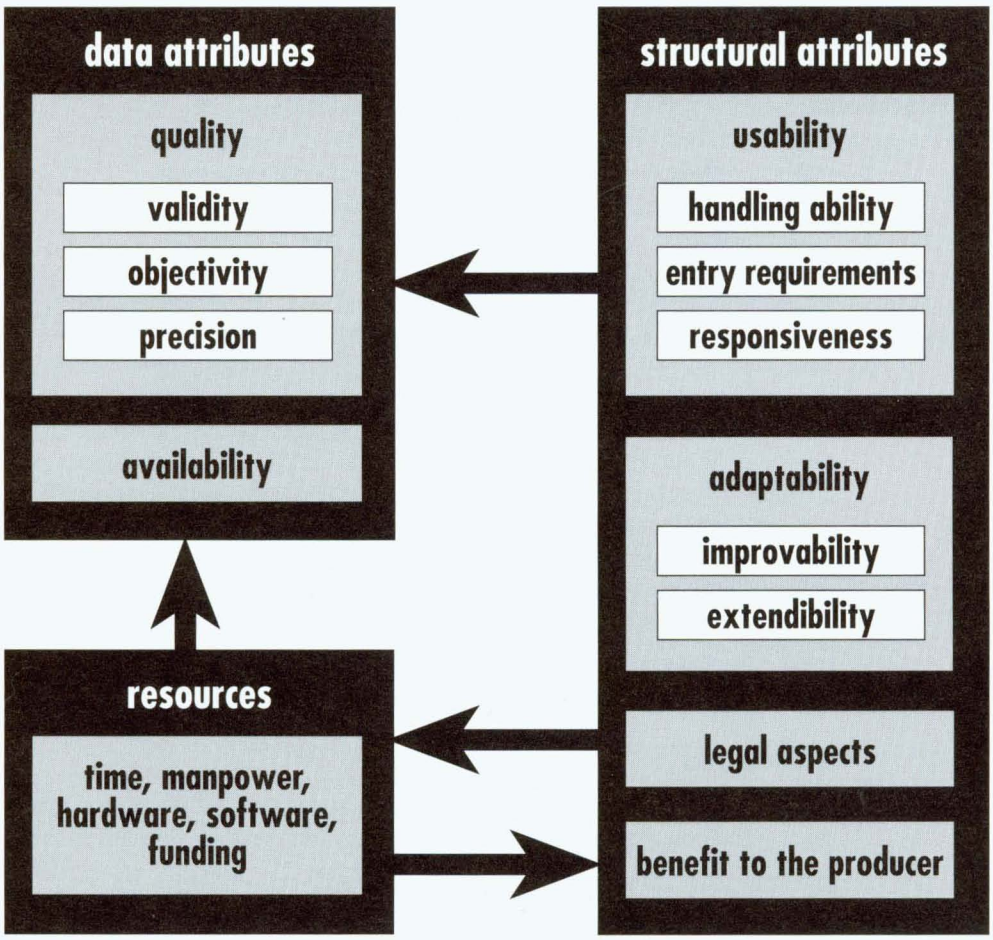

Figure 1
Critical aspects of the projected system

Specifications

From the previously mentioned aspects and problems it is apparent that a flexible data model is needed that meets the requirements of a heterogeneous and dynamic scientific environment. Further requirements are the realization of a data quality check and detailed data security. The users should be connected on-line to the database. Access should be through high level front ends (geographical browser—retrieval, import, and export). At the same time the system will be tied into the international data exchange (e.g. with the NGDC or the WDC A via ftp or mail). The motivation of producers to deliver their data to the data center can be viewed under different aspects. On one hand the producer receives a powerful tool for storing and managing his data. On the other hand he is responsible for "his" data pool and should be able—in a progressive state—to manage his pool independently. Accordingly, the data sets are related to the name of the producer. Moreover, the data will be made citable through publishing an abstract of the data description in the "Reports for Polar Research". This will solve the copyright question and has a positive influence on the availability and quality of data.

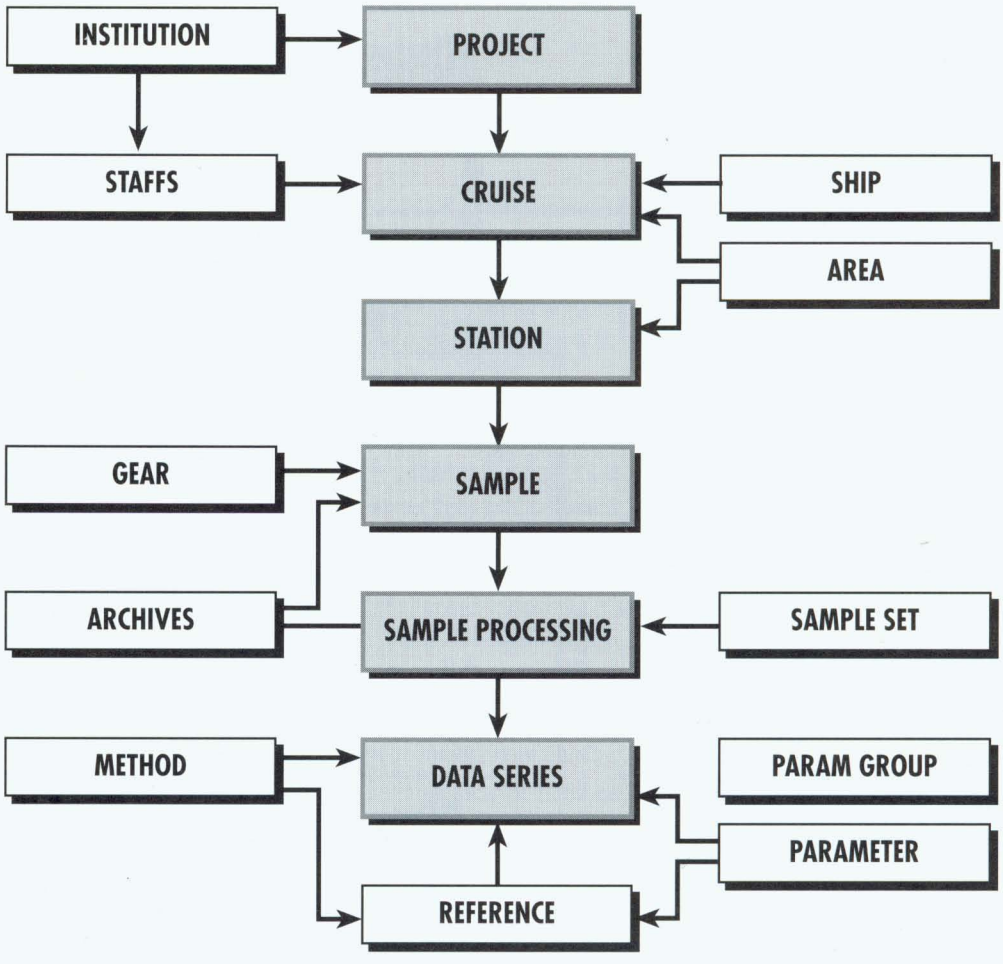

Figure 2
Data model

Proposals coming from several quarters have also suggested that all paleoclimate projects financed by the national funding organizations are obligated to deposit their results in the data center after publication of the data at the latest (similar to the practice of the NSF).

The data model and the network concept are currently in use in the sediment core database of the AWI. The data model (Figure 2) shows a hierarchical order of the meta information, beginning with the projects, which are often bound to an institution, then the cruises, the stations, and the samples. The sample processing and the data series description form the lower levels. Other information supplement these tables, as the gear list or a list of all citations connected with the data. The essential part of this model is formed by the combination of data series, parameter, and method, which guarantees that the system can also be easily adapted to individual and changing requirements. Novice users simply add their specific parameters and methods (if not yet present) to be able to store all desired data types.

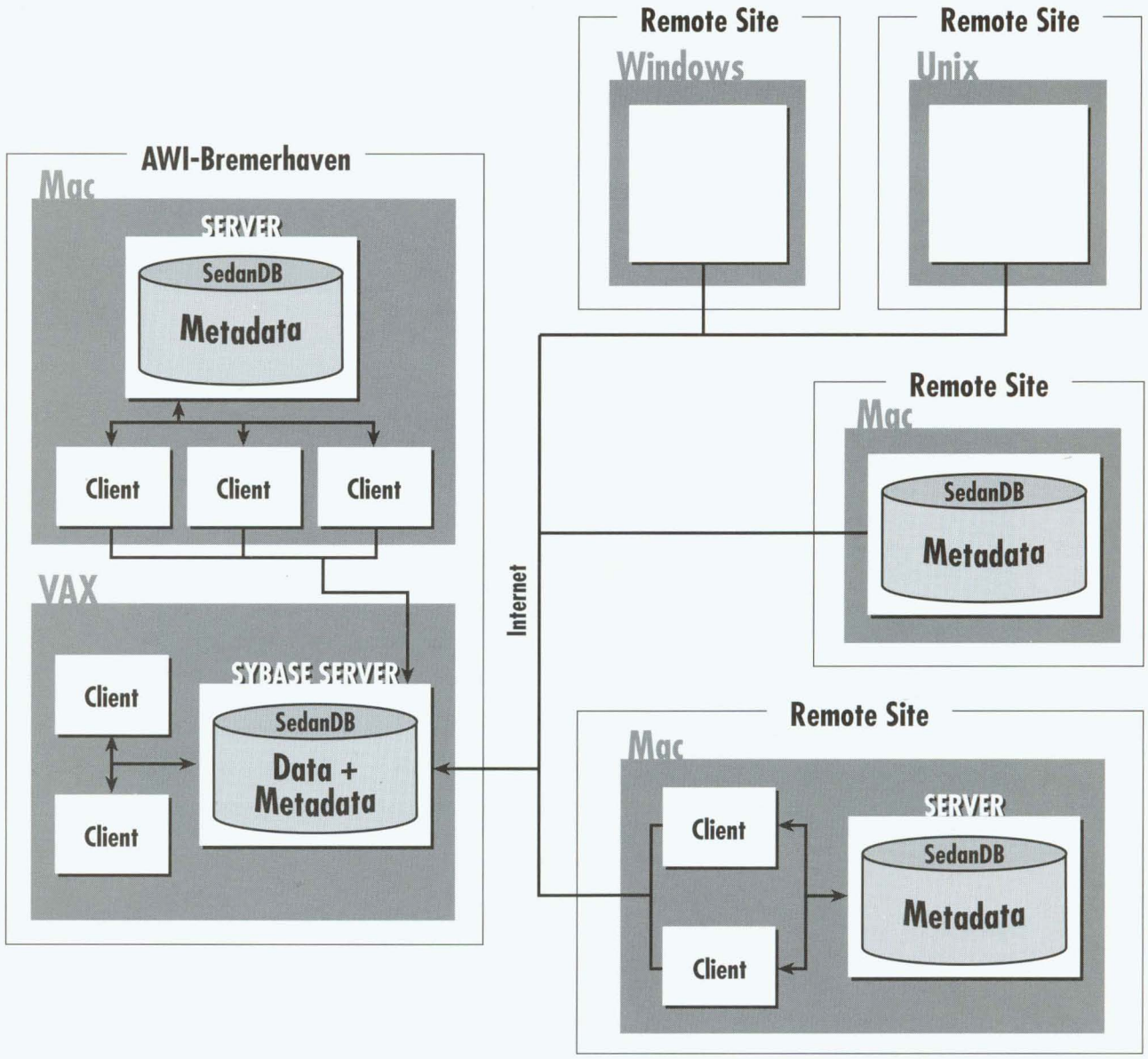

Figure 3
Network concept

The network concept (Figure 3) is based on the client/server-principle. The database server runs on a VAX 6610. Remote users outside Bremerhaven will have access to the server via Internet. To improve the access speed, the meta information is mirrored on local servers. The mutual update of meta information runs in the background. The initial phase of the project (3 years) will be financed by the Federal Ministry for Research and Technology (BMFT). Staff consists of a scientist and a programmer. The necessary hard- and software, as well as basic services (network, system management, communication, backup), are available through the AWI. In 1995 a first status report and an evaluation of the project will be prepared.