

PANGAEA - Network for Geological and Environmental Data

In 1993, scientists from various German research institutes working in the field of paleoclimatology initiated a project to implement a common information system which guarantees longtime storage of data in consistent formats and provides access for the scientific community. Based on the discussion and recommendations of this group, the information system PANGAEA was developed which is now operated by the Alfred Wegener Institute for Polar and Marine Research (AWI) and the Center for Marine Environmental Sciences (MARUM).

The system is able to store any type of data with related descriptions. Due to its importance for marine geological work, it is also able to give information about the availability of sampling material. The system has standardized import/export routines, access with uniform retrieval functions and tools for the visualization of data. The system is designed as a network using client/server technology with access and data exchange through the Internet.

The emergence of an integrated earth systems science also calls for a full knowledge of past conditions, and for data sets that are drawn as composites from different methods, techniques and archives of paleoclimate. The major challenge of PANGAEA in this effort is not only to have the data archived and accessible but to allow retrieval and extraction of individual parts from a comprehensive data collection.

Keeping the complexity of paleoclimate data in mind, the resulting system still needs flexible access tools with high functionality. Thus users of PANGAEA need to go through a short tutorial to learn the principles of the data model and the retrieval process. Due to the simplified data model and the uniform tools on all tables, the system is easy to use once the principles are understood.

In the PANGAEA standard retrieval (<http://www.pangaea.de/retrieval>) the full functionality of the system is realized. To fulfill the requirements of a more specific data access, individual retrieval tools can easily be designed due to the generic structure of the software. As an example the user may try the interface PanCore which was created to obtain

information about available sediment samples/cores in international repositories. A similar interface for a search to published data sets using an author's name or a keyword is in preparation.

The further possibility for retrieving data from PANGAEA is the direct download interface. Using the standard retrieval a data curator can save retrieval information on his local web server. A click on a link to this information included in a web page starts the retrieval and provides the requested data set as a tab-delimited text file (e.g. <http://www.pangaea.de/projects/SFB261/>, click on 'data' in the reference of Müller P. et al, 1994).

The data model

The simplified data structure of PANGAEA is shown as a graphic on the opening screen of the client software (Fig. 1). The graphic allows the user directly to access all tables. The structure mirrors the processing steps for scientific data in general. In projects (PROJECT) cruises are carried out for sampling or measurements (CAMPAIGN). During a cruise at a number of locations (SITE) different samples may be taken or measurements are made (EVENT). At distinct points/intervals the medium to be investigated (e.g. sediment, ice or coral) is sub sampled or measured (SAMPLE). From the investigation of each sample parameter specific values will result (DATA). Data can be retrieved and exported in any combination together with meta-data in an individually configured table.

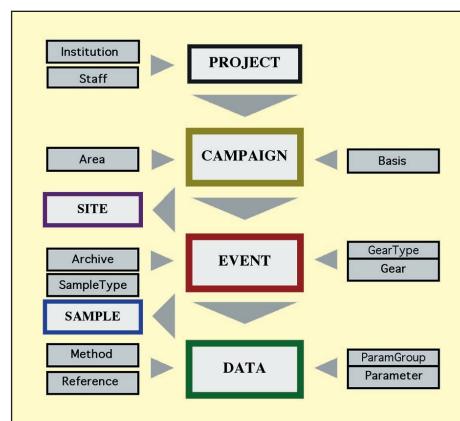


Figure 1: The hierarchy in the PANGAEA data model mirrors the path from the sampling in projects to the published data.



Data handling

The meta-data comprises related information about cruises, sampling sites/sets and storage facilities. The most important meta-data are the location (latitude/longitude) and the elevation of an event. Scientists and institutes related to data sets are stored with their full addresses. Other related items, such as the names of ships, gear or sample types, are defined in lists which can be updated at any time. The parameter list is organized in groups and consists of about 5000 definitions used in marine and paleoclimate research. References for cruise reports or published data can be typed in or imported from professional bibliography software.

The import of meta-data is organized through predefined forms which are available for the import of references, campaigns and sites/events. Analytical data are imported via tab-delimited text with the parameter names or the PANGAEA-ID in the header of the input matrix. Meta-information related to the data (method, owner, comments) has to be defined prior to the import, the relations are made during the import process.

To improve the data consistency, data sets can be stored at different levels of processing. The primary data (e.g. counts of a microfossil assemblage or weights of granulometric investigations) are the raw data for calculations and interpretations. Archiving the raw data allows future recalibration or interpretations. The secondary data are those values calculated from the raw data, and in many cases are given as percentages or other units of concentration. Parameters evaluated from the secondary data (e.g. by statistical methods) are defined as tertiary data (e.g. paleotemperature).

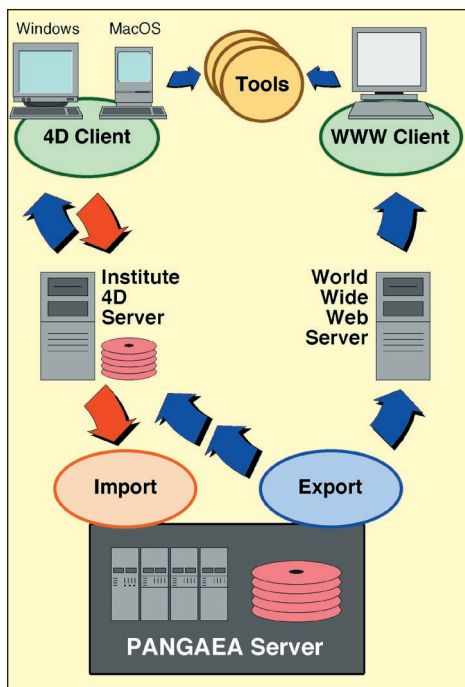


Figure 2: The PANGAEA network uses client/server technology through the Intranet/Internet to communicate and distribute data between institutes.

The retrieval tool for finding and extracting data from the system is uniformly designed for all levels and allows the use of complex combinable search criteria. Data can be exported as tables or plotted with one of the visualization tools. Tables can be sorted and configured individually. Multiple data sets can either be displayed with identical parameters and locations in one column, or the data can be split by locations and/or different versions of data sets into separate columns, thus allowing the direct comparison of data sets from different investigations or locations.

When dealing with the archiving and publication of data, copyright has to be considered. If an information system also stores unpublished data, it is crucial for the acceptance and the trust of the database that the data can be protected. In PANGAEA the owner of a specific data set is able to define access rights for individual users or groups.

Technical setup

PANGAEA uses client/server technology through the Internet with a main server (SUN E10000, 8 processors, 8GB internal memory, 100GB hard disk capacity) running SYBASE as the database management software. The client soft-

ware for import/export, written in 4th Dimension (4D), runs on the operating systems MacOS and Windows. The 4D clients are used by the data curators for project data management (<http://www.pangaea.de/projects/>) and for the import of data from different institutions participating in the network. The client software to be used on the World Wide Web is written as a Java applet and allows read only access on published data for anyone. Registered users can also share unpublished data. The applet includes a draft map and plot functionality for quick overviews. Due to the use of Java applets, users of the web client should use recent browser software versions. Non-Java interfaces are in preparation.

External tools

For the geographical presentation of data the PANGAEA tool PanMap was developed as a stand-alone mapping application for vector data (<http://www.pangaea.de/software/>). PanMap can be used to draw sampling sites with meta-data or analytical data in a geographical context. Maps can be configured with different projections, the styles of map elements can be changed, additional vector data or site information can be imported and managed in different layers, and graphics can be exported. The General Bathymetric Chart of the Oceans (GEBCO) is provided as a bathymetric data source.

The PANGAEA tool PanPlot enables the user to plot data versus depth/altitude or time. Scales and graphic features can be modified by the user and distinct parameters can be selected from a data matrix. PanPlot and PanMap can be directly accessed by the 4D clients. When using the Java web client, data have to be downloaded to the PC first and the visualization software then has to be started by the user. PanMap and PanPlot are freeware.

MICHAEL DIEPENBROEK

Center for Marine Environmental Sciences (MARUM), Bremen University, Germany
mdiepenbroek@pangaea.de

HANNES GROBE

Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany
hgrobe@pangaea.de
<http://www.pangaea.de>

In Memory of Hans Oeschger

Hans Oeschger, one of the founding fathers of PAGES died on December 25, 1998. His role in the establishment and promotion of PAGES drew on both his distinction as a scientist and his commitment as a person. His passing has evoked many tributes in both national and international media. His distinguished career, culminating in the award of both the Tyler and the Revelle prizes, is described more fully and with great insight and affection in these accounts and they reinforce the immensely high standing he had among his peers. Here, we have tried to bring a specifically PAGES dimension to our commemoration by inviting two of his close associates during the development of the project to give their personal recollections. The first of these is Jack Eddy, who worked alongside Hans in the early planning stages:

Hans was one of several proponents for a paleostudies component in the initial architecture of ICSU's International Geosphere-Biosphere Program, and with that goal in mind he accepted membership on the first Scientific Committee for the IGBP in 1985/86. Based on his own early work, and that of his colleagues at the Physical Institute that he directed in Bern, he recognized the valuable role paleoclimate information could play in clarifying some of the main issues that then clouded conjectures regarding the enhanced greenhouse effect. He was also concerned that projects of the WCRP had put so little emphasis, or trust, in paleoclimate data. Hans helped identify candidates who might serve on the initial steering committee for what came to be known as the IGBP Past Global Changes (PAGES) project. He hosted the early meetings in Bern and served as its co-chairman.

Members of the IGBP Working Group on Global Changes of the Past, and later of the PAGES Steering Committee, favored expanding the concept beyond Hans' initial thoughts regarding greenhouse warming and isotope chemistry. With these other insights the PAGES concept was broadened to embrace many different disciplines; to include Holocene as well the somewhat longer time scales of the last glaciations; to utilize PAGES as a vehicle to coordinate the efforts of disparate (and sometimes competing) factions in paleoscience; and to open up and centralize paleo data sources around the world as a way of achieving this kind of needed organization and synthesis.