PERGAMON

Computers & Geosciences 28 (2002) 1201–1210

COMPUTERS & GEOSCIENCES

www.elsevier.com/locate/cageo

PANGAEA—an information system for environmental sciences

Michael Diepenbroek[a,*], Hannes Grobe[b], Manfred Reinke[b], Uwe Schindler[c], Reiner Schlitzer[b], Rainer Sieger[b], Gerold Wefer[a]

[a] *Center for Marine Environmental Sciences (MARUM), University Bremen, Bremen 28334, Germany*
[b] *Alfred Wegener Institute for Polar and Marine Research, Bremerhaven 27515, Germany*
[c] *Physics Department, University of Erlangen-Nuremberg, Erlangen 91058, Germany*

Received 23 March 2001; received in revised form 20 April 2001; accepted 5 May 2001

Abstract

PANGAEA is an information system for processing, long-term storage, and publication of georeferenced data related to earth science fields. Essential services supplied by PANGAEA are project data management and the distribution of visualization and analysis software. Organization of data management includes quality control and publication of data and the dissemination of metadata according to international standards. Data managers are responsible for acquisition and maintenance of data. The data model used reflect the information processing steps in the earth science fields and can handle any related analytical data. The basic technical structure corresponds to a three tiered client/server architecture with a number of comprehensive clients and middleware components controlling the information flow and quality. On the server side a relational database management system (RDBMS) is used for information storage. The web-based clients include a simple search engine (PangaVista) and a data mining tool (ART). The client used for maintenance of information contents is optimized for data management purposes. Analysis and visualization of metainformation and analytical data is supported by a number of software tools, which can either be used as 'plug-ins' of the PANGAEA clients or as standalone applications, distributed as freeware from the PANGAEA website. Established and well-documented software tools are the mini-GIS PanMap, the plotting tool PanPlot, and Ocean Data View (ODV) for the exploration of oceanographic data. PANGAEA operates on a long-term basis. The available resources are sufficient not only for the acquisition of new data and the maintenance of the system but also for further technical and organizational developments.

Keywords: Data management, Data mining, WDC-MARE, PangaVista, GIS

1. Introduction

Information handling and knowledge management have become important backbones in Global Change research. One of the milestones for the implementation of a global information network was the foundation of the system of World Data Centers (WDCs) by the International Council of Scientific Unions (ICSU) in the late 1950s (Panel World Data Centers, 1996). In recent decades public interest in environmental problems has increased enormously and research focusing on these

*Corresponding author.
E-mail addresses: mdiepenbroek@pangaea.de (M. Diepenbroek), hgrobe@awi-bremerhaven.de (H. Grobe), mreinke@awi-bremerhaven.de (M. Reinke), uschindler@pangaea.de (U. Schindler), rschlitzer@awi-bremerhaven.de (R. Schlitzer), rsieger@pangaea.de (R. Sieger), gwefer@uni-bremen.de (G. Wefer).

0098-3004/02/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0098-3004(02)00039-0

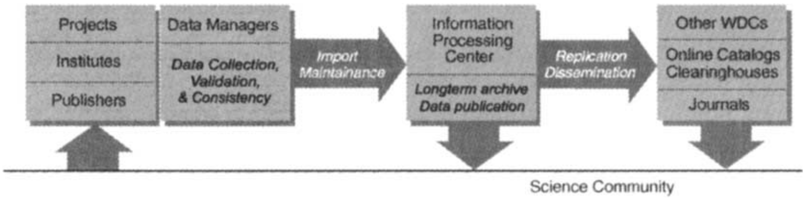

Fig. 1. Data management components, their interaction and resulting data flow.

subjects has been intensified. At the same time developments in computer and network techniques allowed the realization of more sophisticated information systems with increased storage capacities and a general accessibility through the Internet (Mounsey and Tomlinson, 1988).

PANGAEA—the network for geological and environmental data—was initiated in 1993.[1] From the beginning, PANGAEA was supposed to be more than a data archive—it was conceived as a scientific tool (Diepenbroek et al., 1999). This included a flexible and comprehensive data model to compensate for the heterogenous and dynamic scientific environment, flexible geocoding of data, the possibility to extract any part of data from the system, and a comprehensive and quick view on the retrieved information through user friendly visualization tools.

Funded by the German Ministry of Research and Technology during the years 1994–1997, PANGAEA initially provided data management services on the national level, and later expanded its role to include EU and international projects.

In recent years the data management "business" has evolved into a general service open to various scientific groups and programmes. Activities are concentrated within the fields of marine geoscience and bioscience. In 2000 the PANGAEA group initiated the foundation of the World Data Center for Marine Environmental Sciences (WDC-MARE).[2] The WDC system offers an acknowledged framework to embed data management services into the science community, which means improved communication networks and more opportunities for project participation, data acquisition, and funding. WDC-MARE uses the information system PANGAEA as an operating platform.

This paper describes the structures, functionality, and contents of a modern scientific information system used as a basic infrastructural element for Global Change research.

[1] http://www.pangaea.de.

[2] WDC-MARE—http://www.pangaea.de/WDC-MARE.

2. PANGAEA data management

The essential service supplied by PANGAEA is scientific information management, in particular project data management. To be successful, it requires a high level of organization both related to the technical as well as scientific and human aspects of information handling. To guarantee a significant availability of data, data managers should play an active role in science, that is, data management should be a necessary and creative component for scientific work. Technical aspects should be hidden to users as far as possible and interfaces be supplied that enable logical views on the data.

2.1. General data management structures

The central component of PANGAEA is an information-processing center (IPC) whose primary objective is the long-term storage of data (Fig. 1). Data sources include research projects, institutions, and individual scientists having an interest to publish their data. For each data set, a data manager is responsible for the collection, validation, and the consistency of the data. The data manager plays a critical role in the processing of data, and must have scientific expertise, good communication skills, and a clear overview on the project they are working on. The managers are usually funded directly from the projects or institutions associated with the data, and adhere to a formal data management plan.

During the import process, the data documentation (metadata) is checked for completeness and validity. This manual procedure is supplemented by an system of generic and parameter-specific validation routines, which is frequently updated. These routines are based on employed analytical methods and parameters, which requires a given standard unit, range of acceptable values and the possible precision of the data. This information will be used by validation routines during the import to identify suspect values, e.g., outliers.

Generally, data quality can be described in terms of the validity of methods and the precision and objectivity of measurements. It is not essential to have only excellent quality data sets; however, it is important that the quality can be estimated. The completeness of the meta-information, including, in particular, the descriptions of the analytical methods and the reference where the data have been published first is crucial in the understanding of the analytical data.

The ultimate goal is the publication of data. In PANGAEA there is no fixed rule for this. Publication depends on the wishes of the data producers. Unpublished data can be password protected, a useful feature in particular during the runtime of a project to enable the efficient exchange of preliminary and or sensitive data between working groups.

If data sets are referenced by a publication, links to the publication are added to the data descriptions and—where possible—an information in the publication refers vice versa to the related data sets stored in PANGAEA. A closer link between the publication of data and the traditional publication process improves the availability and quality of data. In PANGAEA this was exercised already for a number of publications (e.g., Fischer and Wefer, 1999)

Certain data sets might also be replicated into other data centers, in particular in those cases where PANGAEA has cooperations with other WDCs. The replication of data after publication is declared policy of the WDCs.

Each data set in PANGAEA is represented by a uniform resource locator (URL). In 2001, data set descriptions including such links will be further disseminated for usage in metadata systems, e.g., the Global Change Master Directory (GCMD).[3] For this purpose a metadata server is set up, which fullfills recent metadata standards and supplies a Z39.50 interface (Lynch, 1997). The interface can be used to download metainformation or can be incorporated into clearinghouses, thus avoiding the explicite replication of metainformation.

2.2. The data model

The great variety of parameters, methods, calibrations and interpretations used in environmental reconstructions, as well as the modification of established methods, are major obstacles for the integrative use of data sets in a common system. The challenge of managing these heterogenic and dynamic data was met in PANGAEA through a highly flexible data model. This was achieved by a combination of a simple and fully normalized relational database structure in combination with the functionality of middleware components.

The model reflects the standard processing steps for data in the natural sciences (Fig. 2). Different institutes/projects (PROJECT) working in the earth science fields carry out expeditions (CAMPAIGN) for sampling.

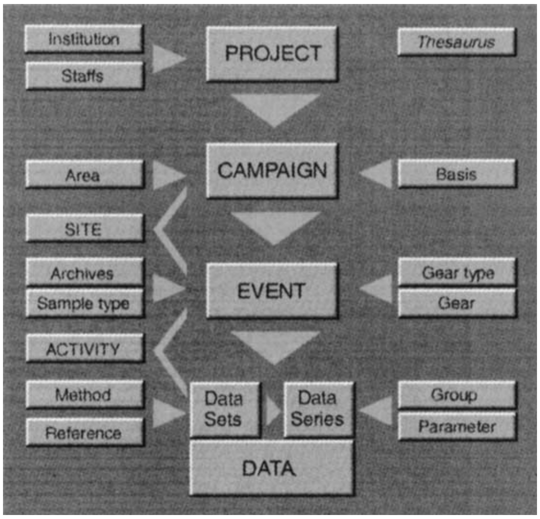

Fig. 2. Data model used in PANGAEA. Model is used as entrance level for ART and 4D clients. For this purpose model was slightly simplified. Arrows show principal relations in model.

During an expedition at a number of locations (SITE) different samples may be taken or measurements made (EVENT). At distinct points/intervals the medium to be investigated (e.g., sediment, water or ice) is sub-sampled or measured as required (ACTIVITY). Finally, from each sample or measurement one or several analytical data values may be generated, which can be found on the ‘DATA’ level. These main levels are supplemented by related tables of ancillary information about staffs and institutions, gear, area, and used vehicles (Basis), literature (reference), methods, and parameters. The latter are gathered into parameter groups for a better overview.

Data descriptions are provided at different levels: data series and data sets. A data series consists of one or several data points for one parameter. Data sets are typically composed of a number of data series, e.g., a CTD profile, an age model, or a pollen diagram. Descriptions of data sets follow the standard of the GCMD (Fischer and Wefer, 1999) and include information on parameters and methods used, principal investigators, references, size, and temporal and spatial coverage of the data set. The essential part of the model is the combination of the ‘DATA’, ‘Parameter’ and ‘Method’ tables, which allows the definition and storage of new, unique parameters by the user at any time. The technical data types can be numerical, text or pictures.

The model enables a flexible geocoding of data. Up to five geocodes can be used simultaneously for the description of data points in space and time: latitude,

[3] NASA’s Global Change Master Directory (GCMD), http://gcmd.nasa.gov.

longitude, datetime, absolute age, depth in water, sediment, ice, or altitude. The data model thus allows for the combination of all data types and the extraction of any part of data. Such a model is therefore a useful requisite for the analysis of complex data inventories or data mining (Han and Kamber, 2001).

2.3. Technical structures

PANGAEA is based on a three-tiered client/server architecture (Fig. 3) (Eckerson, 1995). On the client side (frontend—first tier) a number of interfaces are offered for access to the system. The middleware (second tier) is an application server with several components for the import, retrieval and maintenance of the data. A relational database finally is the central storage facility for all types of data (blackend—third tier). Each of the client programs and the SQL server communicate with the intermediate layer, the application server, whereby the communication between application server and web clients requires an additional web server.

All components are encapsuled as far as possible and use standard interfaces to communicate with each other. Middleware and frontend components are generic to ensure a flexible functionality. Because open system design components can be altered or new ones added without affecting existing components. This concept extends the lifetime of components and reduces overall maintenance costs.

2.3.1. The database server

The SQL server used is a Sun E10K Solaris system with four Processors, 8 GB RAM, and sufficient disk space running SYBASE (RDBMS) as database management software. All the information served by PANGAEA is completely stored in a relational database on this server.

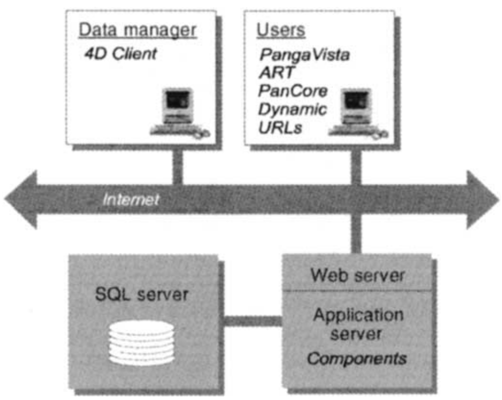

Fig. 3. Technical structure of information system.

2.3.2. The middleware

An application server keeps the semantics necessary for the control of the information flow, information quality and consistency. The combination of a relational data structure with the logic of a middleware enables quasi-object-oriented views on the stored information. The components for the retrieval transform the logical requests from clients into SQL requests and convert the result sets into client readable formats. Thus, complex data matrices (e.g., time slices) can be retrieved without the user knowing about the underlying storage structures. Other components ensure that new data are imported properly into the relational database and that suspect values are flagged and unvalid information is rejected. Many of these tasks are queued in the middleware, thus clients only need to deliver tasks without having to wait for results. Most of the computing is done by the middleware.

2.3.3. The clients

A number of interfaces enable users to access the system on different levels, metainformation and analytical data. The interfaces vary in flexibility and complexity of usage, from simple search engines to complex retrieval and data processing tools. Except the 4D client, which is the central data management interface, they are all web-based.[4]

2.3.3.1. PangaVista. This simple web-based search engine enables the retrieval of whole data sets, referenced by dynamic URLs (Fig. 4). The PangaVista is linked with a Thesaurus comprising all of the metainformation related to the analytical data, thus can be used with a variety of keywords like principal investigator, author, title, parameter, method, project, sampling location, etc. Keywords can be combined to create boolean expressions, with syntax identical to that used by AltaVista.[5] Results contain a short description of the retrieved data sets with the URLs to the related data included. Data can be downloaded as HTML or tab-delimited text. All data sets contain a metadata header according to the DIF standard used by the GCMD.

2.3.3.2. Advanced retrieval tool (ART). ART is a web-based tool to retrieve analytical data and all types of metainformation. The flexibility of this tool allows for the definition of complex retrievals as well as for the configuration and formatting of result sets, which are shown in listings and visualized as plots and/or maps (Fig. 5). The simplified data structure of PANGAEA is the opening user interface, which allows users to enter all levels by selecting the desired field. Results can be

[4] http://www.pangaea.de/Retrieval.

[5] http://www.altavista.com.

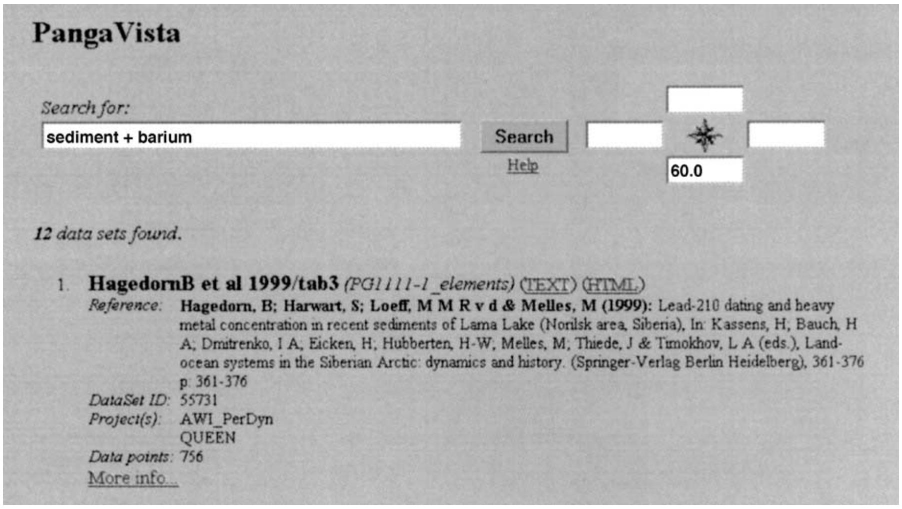

Fig. 4. PangaVista is simple search engine, which allow users to retrieve desired data sets by seeking matches to user-specified keywords and geographical constraints. Here, number of data sets were found for query 'sediment and barium for latitudes higher 60° North'.

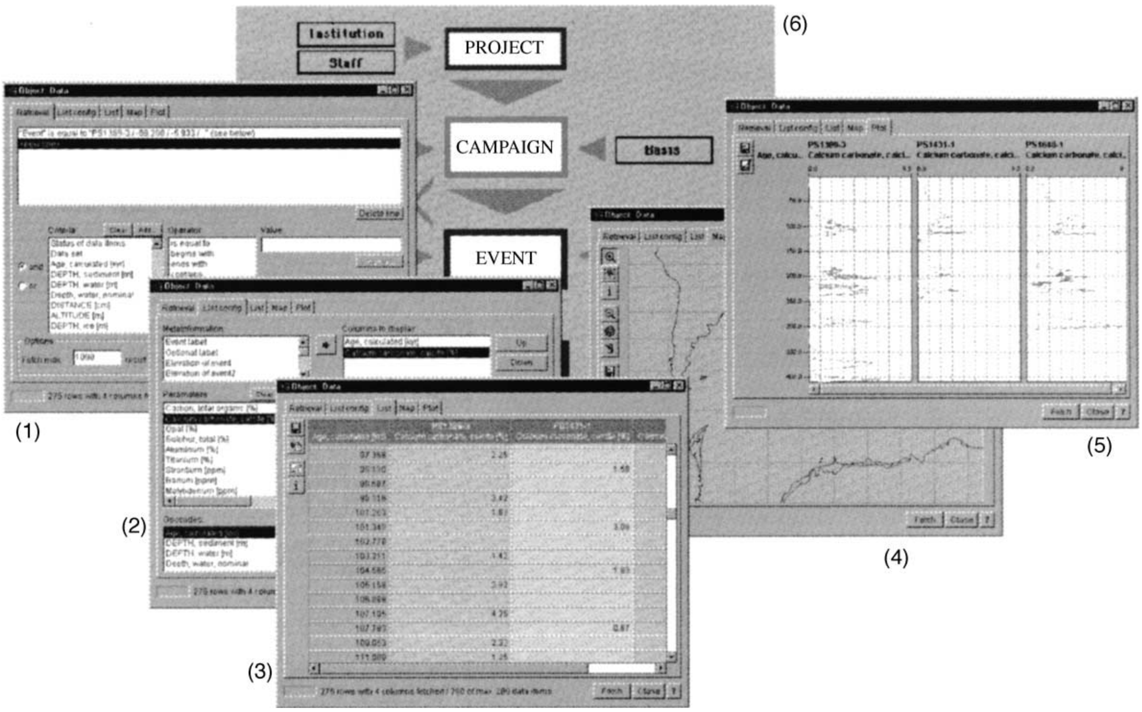

Fig. 5. ART is data mining tool, that allows for complex retrievals on all levels of information. Example shows retrieval for analytical data. After entering data level, user interface opens window containing five tabs: (1) 'Retrieval'—definition of query conditions; (2) 'Config'—output configuration; (3) 'List'—listing of result set; (4) 'Map'—visualization of geographical coverage; (5) 'Plot'—plotting of time series. Default screen for entering different levels is graphic of simplified data model (6).

downloaded as tab-delimited text, plots and maps as 'ebf' or 'ps' files.

Through its functionality ART can be classified as a data mining tool. It enables the user to retrieve analytical data for any section of the globe and to combine any types of data, provided that the geocoding of the data is compatible. Thus, time series from different sites or even different environments can be

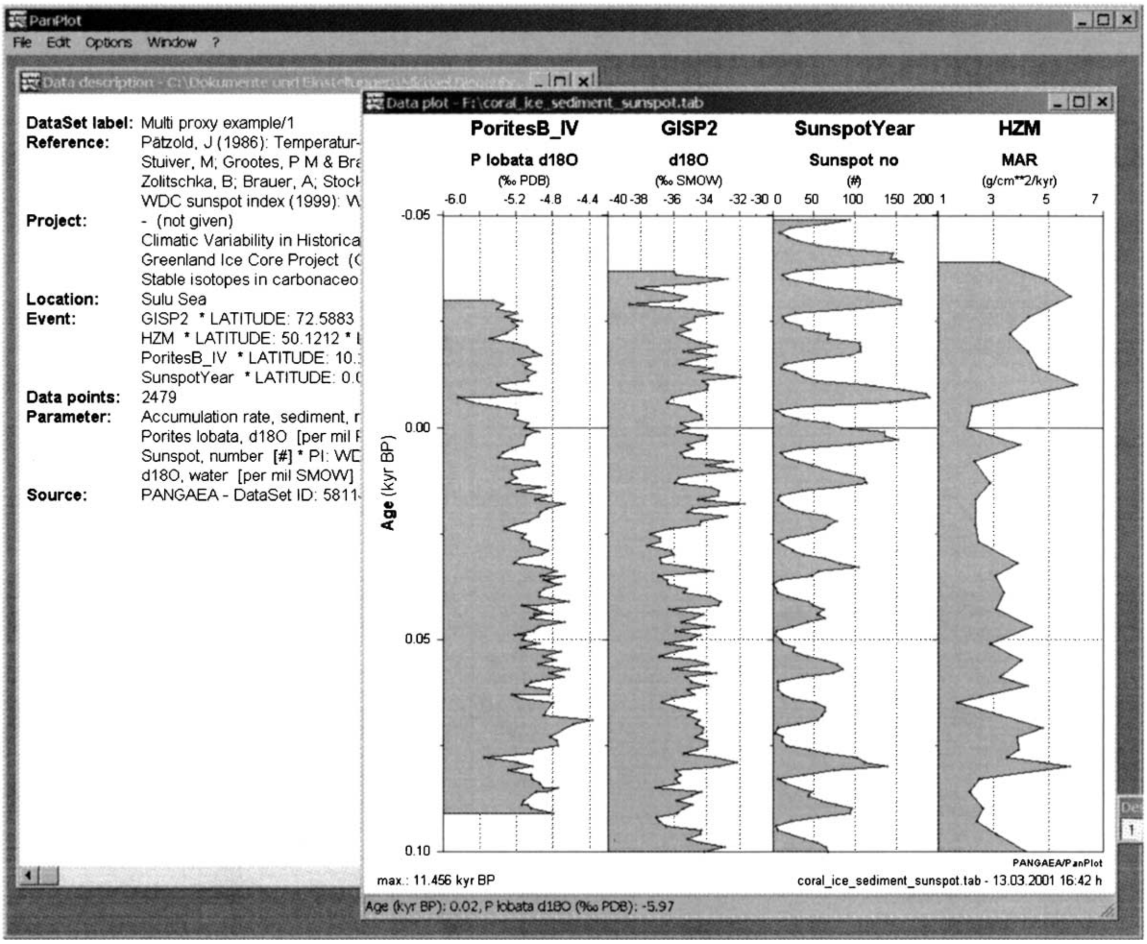

Fig. 6. Data set retrieved with ART showing time series from different environments: oxygen isotope data from corals and ice, number of sunspots, and accumulation rate of lake sediments. Data set was plotted with PanPlot.

plotted on uniform scales (Fig. 6)[6] or timeslices can be extracted for a specific type of data.

For the GUI of ART a Java applet was used; the middleware transforms the logical requests and prepares the result sets for the client. A context sensitive help system is supplied for all levels.

2.3.3.3. PanCore. PanCore is a web-based interface for the retrieval of sampling and site locations. Definition of retrievals is done through forms. Geographical constraints can be set within the included map, which is also used for the visualization of the results. The list of retrieved locations can be downloaded as text. PanCore is a subset of components from the ART interface.

[6] http://www.pangaea.de/PangaVista?query=58114.

2.3.3.4. 4D client. The 4D client (4th dimension) is the general administrative tool for the processing and maintenance of the information stored in PANGAEA. It supplies special routines and graphical user interfaces (GUIs) for the import of analytical data and all types of metainformation, for the update of existing information, and for the definition of data sets. The latter enables administrators to define new data sets by using existing data sets of metainformation, thus allowing for different and dynamic views on the data without causing redundancies.

2.3.3.5. Static retrievals. PANGAEA enables the usage of predefined retrievals on any sub-set of analytical data and all types and quantities of metainformation. Retrievals and output configurations can either be defined with ART or with the 4D client. Data set descriptions are supplemented by default with retrieval

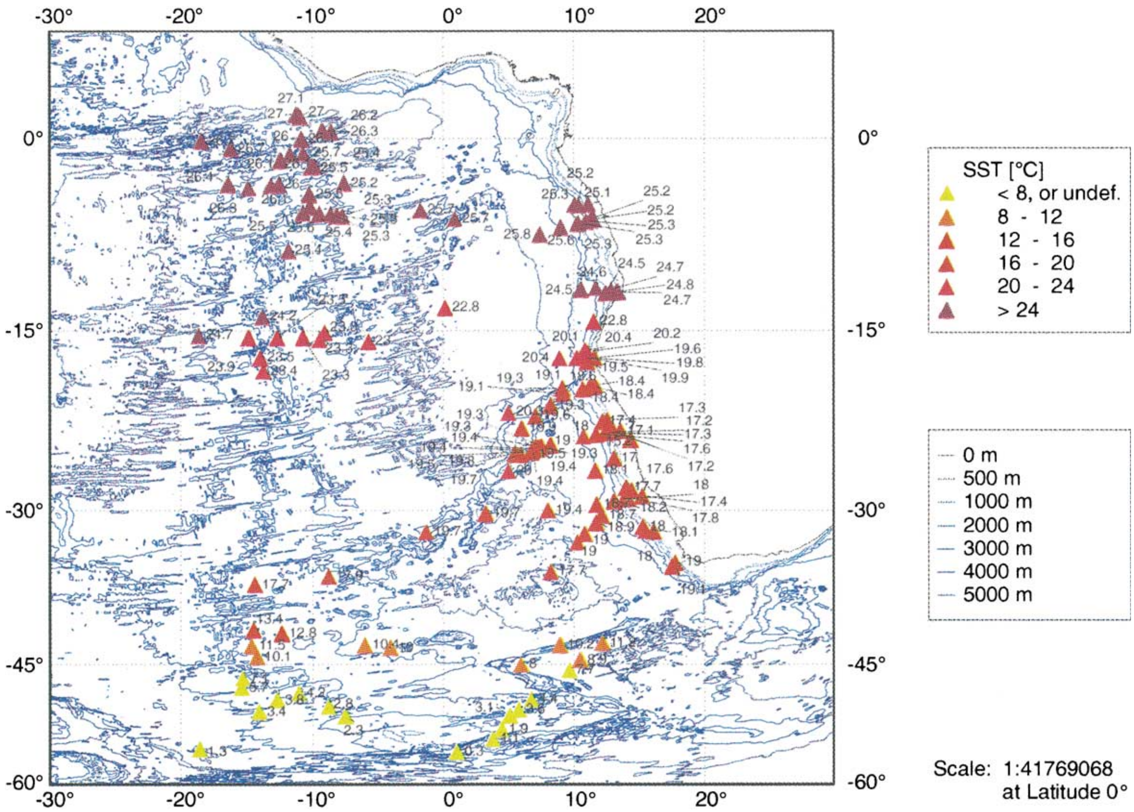

Fig. 7. Map generated with PanMap showing recent sea surface temperatures (SST) calculated from alkenones (Müller et al., 1998).

and configuration information. Static retrievals are refered to by URLs and can be used in HTML pages and pdf files. This feature in particular allows for the dynamic referencing of data sets from within an online-publication.

3. PANGAEA visualization and analysis software

PANGAEA serves a series of software products to be used either as standalone applications or in conjunction with the information system. Four of these products are described below. Each of these are well-established scientific tools, of excellent quality, and used worldwide.[7]

3.1. PanMap

For the geographical presentation of data, the PANGAEA tool PanMap was developed. PanMap is a simple to use geographical information system (GIS) running on Macintosh and Windows platforms. It can handle stationary and vector data, which are organized in layers. The GIS can either be connected directly to the database front end or can be used as a standalone application. After calling PanMap from PANGAEA, the software is automatically started, data are exported from the database, imported by the mapping software and plotted in a default map. The software allows for the configuration of maps with different projections, the import of additional vector and stationary information and the labelling of stationary information. Styles of map elements can be changed for stationary data, in particular color scales can be defined according to the numerical values of related data (Fig. 7). Maps can be exported for further processing in other applications.

An extensive collection of map resources is available on the PANGAEA web site, among which are the General Bathymetric Chart of the Oceans (GEBCO)[8] in four different resolutions, the World Vector Shoreline (WVS)[9] and the Global 30 Arc-Second Elevation Data

[7] http://www.pangaea.de/Software.

[8] General Bathymetric Chart of the Oceans (GEBCO), http://www.ngdc.noaa.gov/mgg/gebco).

[9] World Vector Shoreline (WVS), http://www.ngdc.noaa.gov/mgg/fliers/93mgg01.html.

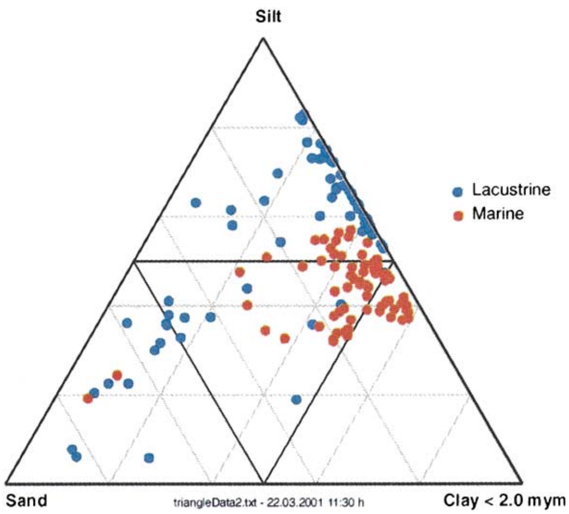

Fig. 8. Ternary diagram plotted with PanPlot showing textures of marine and lacustrine sediments.

Set (GTOPO30),[10] which was converted into vector data. Users can also prepare individual maps for usage with PanMap. Tools are provided to digitize printed maps, the resulting output data as well as any individual contour lines can be converted to the proprietary PanMap format.

PanMap is available on the internet for several years. It has become one of the most used freeware GIS worldwide.

3.2. PanPlot

The PANGAEA tool PanPlot allows the visualization of multiple data series as plots (Fig. 6) or ternary diagrams (Fig. 8). Up to 255 parameter can be plotted at uniform scales. PanPlot can handle numerical, datetime, and text data. Values on the ordinate can be of type numerical or datetime. The input format is a plain ASCII spreadsheet. Scales and graphic features can be modified by the user and distinct parameter can be selected from the data matrix. Labelling as well as whole plots can be rotated by 90°. The export format of graphs and plots is platform specific (EMF, PICT). Like PanMap the plotting tool might be used as a stand-alone application or as a plug-in of the information system. Data sets exported from PANGAEA are read by PanPlot inclusive the metaheader, which is shown in a separate window. The software is available for Macintosh and Windows platforms.

[10] Global 30 Arc-Second Elevation Data Set (GTOPO30), http://edcdaac.usgs.gov/gtopo30/gtopo30.html.

3.3. Ocean Data View (ODV)

ODV (Schlitzer, 2001a) is a software package for the visualization of oceanographic data on Windows and Unix platforms. The package can be used to create and manage large sets of proxy and (pale)oceanographic data and provides tools for easy exploration and the graphical display of these data as property/property plots, color sections and color distributions on isosurfaces (Schlitzer, 2001b). The data collection format of ODV is optimized for dense storage and direct data access allowing the storage of large station collections on desktop computers.

ODV is also distributed with the Electronic Atlas of WOCE Data (eWOCE—Schlitzer, 2000).

3.4. PaleoToolbox

The software used to calculate the parameters for paleoenvironmental reconstructions (e.g., paleotemperature) from proxy data (microfossil assemblages) with transfer functions (Imbrie and Kipp, 1971) was converted for use on a Macintosh (MacTransfer) (Sieger et al., 1999). The file format used by MacTransfer to calculate the raw data is proprietary and thus not supported by the database. The preprocessor included in the PaleoToolbox routines converts the proxy data from a given microfossil assemblage into the import file format for MacTransfer. It is controlled via a ‘project file’ which can be assembled with the PaleoToolbox software. Export files can be directly processed with a spreadsheet software or PanPlot. The PaleoToolbox software also allows microfossils to be counted with the aid of a computer keyboard. Another way to calculate paleoclimatic parameters from proxy data is the Modern Analog Technique (Imbrie and Purdy, 1962; Houston, 1980). The software MAcMAT, as a further part of the PANGAEA tools, reads the proxy data and calculates, for example, paleotemperatures.

4. Operation resources and significance of PANGAEA

The institutional framework for PANGAEA, respectively the World Data Center, is supplied by the Alfred Wegener Institute (AWI), Bremerhaven and the Center for Marine Environmental Sciences (MARUM), Bremen. Several persons are responsible for the technical and scientific organization and development. The data management group sums up to about six fulltime scientists. Their number varies depending on the number of projects PANGAEA is involved in. The computer center of the AWI is maintaining all the basic IT services to run the information system, that is hard- and software, the local networks, and backup services. Fast

Table 1
Contents of the PANGAEA information inventory (2/2001).

Projects	97
Campaigns, expeditions	2642
Events, sampling sites	121,066
Parameter (data types)	14,063
Methods	613
References to publications	3812
Data sets	7529
Data series	151,517
Data items	29,298,452

access to the PANGAEA servers is enabled by the Bremen highspeed network (152 Mbit).

PANGAEA has a long-term perspective. Both the MARUM as well as the AWI have stated their commitment for at least 10 years.

Data management services on an international level have been supplied since 1996. Until the beginning of 2001, PANGAEA was a partner in 23 projects covering many fields of environmental sciences.[11] In each project, at least one half-time or full-time scientist had the role of data manager. The project data management contributed substantially to the amount and quality of archived data. It is important to acknowledge that the acceptance of an information system is—besides user-friendliness—essentially determined by its contents (Table 1).

The access frequencies on the PANGAEA servers double about every year. In the second half of 2000 the average number of visitor sessions per day on the web site was 200 and the mean number of daily data base queries was 250 of which about 60 were retrievals for analytical data.

Access logs as well as the response of users clearly show the growing acceptance and significance of web-based information systems for science.

5. Conclusion

PANGAEA is an information system for the long-term archiving and publication of any georeferenced data related to environmental sciences. The flexibility of the data model compensates for the heterogeneity in the different scientific fields. Analytical data are stored consistently together with the related metainformation necessary for their understanding and evaluation. Data set desciptions follow international metadata standards. Online publications can refer to related data sets via dynamic links. With its comprehensive GUIs and the built in functionality for import, export, and maintenance of information, PANGAEA is a highly efficient system for scientific data management and data publication. For the visualization of data sets, high-quality software tools are distributed as freeware from the PANGAEA web site. Users have the choice of several web-based interfaces to retrieve information from the system, ranging from a simple search engine to a sophisticated data mining tool. The mining tool allows for the retrieval and combination of any sub-set of analytical data and metainformation, thus producing added values to the data inventory. Data mining tools are particularly useful for projects that have to synthesize a variety of different data resources, e.g. multi-proxy approaches or modelling experiments with a need for observational data for the verification of results.

[11] http://www.pangaea.de/Projects.

Acknowledgements

Thanks are expressed to Christopher Cogan with assistance in editing.

References

Diepenbroek, M., Grobe, H., Reinke, M., Schlitzer, R., Sieger, R., 1999. Data management of proxy parameters with PANGAEA. In: Fischer, G., Wefer, G. (Eds.), Use of Proxies in Paleoceanography—Examples from the South Atlantic. Springer, Berlin, Heidelberg, pp. 715–727.

Eckerson, W.W., 1995. Three tier client/server architecture: achieving scalability, performance, and efficiency in client/server applications. Open Information Systems 10, 1, 3(20), 1–12.

Fischer, G., Wefer, G. (Eds.), 1999. Use of proxies in paleoceanography: examples from the South Atlantic. Springer, Berlin, Heidelberg, http://www.pangaea.de/Projects/SFB261/Use_of_Proxies.

Han, J., Kamber, M., 2001. Data Mining. Concepts and Techniques Morgan Kaufmann Publishers, 500pp.

Houston, W.H., 1980. The Agulhas current during the late Pleistocene: analysis of modern faunal analogs. Science 207, 64–66.

Imbrie, J., Kipp, N.G., 1971. A new micropaleontological method for quantitive paleoclimatology: application to a late Pleistocene Caribbean core. In: Turekian, K. (Ed.), The Late Cenozoic Glacial Ages. Yale University Press, New Haven, CT, pp. 71–181.

Imbrie, J., Purdy, E.G., 1962. Classification of modern Bahamian carbonate sediments. Memoirs, Association of Petroleum Geologists 1, 253–272.

Lynch, A.C., 1997. The Z39.50 Information Retrieval Standard, ISSN 1082-9873, http://www.dlib.org/dlib/april97/04lynch.html.

Mounsey, H., Tomlinson, R., 1988. Building Databases for Global Science. Taylor & Francis, London, 419pp.

Müller, P.J., Kirst, G., Ruhland, G., von Storch, I., Rosell-Melé, A., 1998. Calibration of the alkenone paleotemperature index UK'37 based on core-tops from the eastern South

Atlantic and the global ocean (60°N–60°S). Geochimica et Cosmochimica Acta 62, 1757–1772.

Panel World Data Centers, 1996. Guide to the World Data Center System. Issued by the Secretariat of the ICSU Panel on World Data Centers, http://www.ngdc.noaa.gov/wdc.

Schlitzer, R., 2000. Electronic atlas of WOCE: hydrographic and tracer data now available. EOS Transactions, American Geophysical Union 81 (5), 45.

Schlitzer, R., 2001a. Ocean-Data-View, http://www.awi-bremerhaven.de/GEO/ODV.

Schlitzer, R., 2001b. Interactive analysis and visualisation of geoscience data with ODV. Computers and Geosciences 28 (10), 1211–1218.

Sieger, R., Gersonde, R., Zielinski, U.A., 1999. New extended software package for quantitative paleoenvironmental reconstruction—EOS electronic supplement, http://www.agu.org/eos_elec/98131e.html.