THE INTERNET DATABASES OF THE WORLD DATA CENTER FOR GEOPHYSICS, BEIJING

Fenglin Peng[1,2*], Dan Wang[2,3], Xinbao Zheng[5],Lijun Xing[2], Keyun Tang[1,2], Bangyan Yue[1,2], Xiaoyang Shen[1,2], Jian Zhang[2,3], Le Peng[4], and Qinghua Huang[4]

[1]*Institute of Geology and Geophysics, CAS;*
[2] *WDC for Geophysics, P.O. Box 9825, Beijing, 100029, China*
[3]*Gradute Univ. of CAS;*
[4]*Peking University;*
[5]*Thinghua University;.*
**Email:* peng@geophys.cn

ABSTRACT

Since the 1990s, our center's geophysics database has been available on the Internet (http://gp.wdc.cn and http://wdc.geophys.cn). Based on HTML language, the website offers simple data service and is being constantly upgraded and improved. We have adopted ORACLE as the database and use JSP (Java Server Pages) technology to create dynamic pages. The basic function of the network is to store, check, inquire, and renew the data. Users can query and download the data in txt form and view pictures generated by Matlab. At present the system is running well in both versions: Chinese and English, each having a unified style. The pages are simple and convenient to use. Its users come from most provinces of China, including Taiwan, and from developed or developing countries, such as the U.S.A., Germany, Japan, Singapore, Bulgaria, Canada, etc.

Keywords: Geophysics data**,** World Data Center, Internet database

1 INTRODUCTION

The study on geophysics is rich in intellectual challenges and has tremendous potential to dramatically alter and improve our society and our lives. The geophysical research in China has a very long history. For example, the earliest geomagnetic survey in China was made before 1800, and the modern geomagnetic observatory in China was built in the 1870's. The accumulation of various kinds of data and information resources has a broad range and abundance records. With the development of the Internet, it is necessary to have databases online for scientific researchers to use these data conveniently. In recent years, we have gradually built up a web database in our center.

2 COLLECTION AND ARRANGEMENT OF GEOPHYSICAL DATA

The existing data resources derive mainly from long-term geophysical experiment and monitoring. We have geomagnetic main field charts in papers with long histories, microfilms of magnetograms rescued from Sheshan observatory, CDs, hard disks, etc. The disciplines include solid earth geophysics, marine physics, space physics, atmosphere physics, and several second class disciplines, such as geothermics, geomagnetism, seismology, ionosphere physics, geoelectricity, earth electromagnetism, etc. Classified by region, we have data from the Southeast, Yunnan-Guizhou, South China Sea Islands, Qinghai-Tibet Plateau, and the Antarctica. The applications of these data resources refer scientific research, national defense, communication, aviation, navigation, mining, etc.

According to the current demand, we have tentatively designed the following databases:
(1). Basic parameters of geophysics;
(2). Geomagnetism;
(3). Gravity;
(4). Geothermics;
(5). Space physics;
(6). Structure of the deep Earth;
(7). Scholars' information of Chinese geophysics

Supported by The Ministry of Science and Technology of China, and Chinese Academy of Science

3 THE BRIEF DESIGN PROCESS OF WDC FOR GEOPHYSICS, BEIJING

In the late 1990's, with societal development, the demand of sharing scientific data on the Internet increased. Meanwhile, the rapid development of computers, networks, and database technology made this idea practical. According to the task of the pilot project of the Ministry of Science and Technology, our center (Gao, 1992) began to utilize computer networks as carriers of data services, successively setting up the website and database on the server of the research room and research institute.

At first, the website was based on HTML language and offered simple data service. In 2000-2002, we set up a database based on ASP language and SyBase environment, which were on the system having a special network server and magnetic disc array. By adopting IBM X220 as the network server and RaidTech NAS as the network magnetic disc system, our total storage capacity reached nearly 300GB.

In 2003, we used PHP language and MySQL environment for a minute mean value observed network. In the process of querying magnetograms in the database, we used the same method as the WDC magnetism data center in Japan, where magnetograms are produced and stored in the database in advance. However, this method not only took up a large amount of magnetic disc space but also involved a large amount of data transmission, which slow down the search speed.

In 2004, we upgraded the system again and adopted Java and SQL Server 2000 to set up the minute mean values observation network. In this system, we do not generate magnetograms in advance but adopt Java to produce the desired picture and produce the relevant webpage dynamically. This approach improved the search speed significantly and also avoided taking up a large amount of magnetic disc space.

4 TECHNOLOGY ROUTE OF WDC FOR GEOPHYSICS, BEIJING

In WDC for Geophysics, Beijing, we have adopted SQLServer2000 as our database. It works well for the client and server environment. It is organically integrated with Windows XP and offers a scheme of information management systems of enterprise layer. In July 2006, we built a new Chinese website on another server. We used Tomcat and Apache as the web server engine: Tomcat for dynamic pages and Apache for static pages. The program language is Java Server Pages (JSP) and the database is ORACLE.

4.1 Web server engine

The operating system is the Windows 2000 server. We adopted the Apache HTTP server as the web server engine and TOMCAT as the service_container: Apache for static pages and Tomcat for dynamic pages. At present, dynamic pages are mainly used for showing magnetograms for different time periods according to users' needs.

4.2 Databases

We use ORACLE to design our main database on the Internet. The main characteristics of ORACLE are as follows:
(1) Support on all known platforms, not only Windows-based platforms.
(2) PL/SQL is a more powerful language than T-SQL.
(3) More fine-tuning to the configuration can be done via start-up parameters.
(4) It has strong transaction processing functions and adopts various kinds of methods to guarantee the integrity of the data.

4.3 Programming language

Java Server Pages (JSP) technology is based upon Java and is used to create dynamic pages (Yan, 2004). It provides a simple and convenient way for creating and displaying the web pages that generate content dynamically. JSP technology enables Web developers and designers to rapidly develop and easily maintain information-rich, dynamic Web pages that leverage existing business systems. As a part of the Java technology family, JSP technology enables rapid development of Web-based applications that are platform independent. JSP technology separates the user interface from content generation and enables designers to change the overall page layout without altering the underlying dynamic content.

4.4 Java database connectivity

We use Java as the database connectivity in our web database. Java, being robust, secure, easy to use, easy to understand, and automatically downloadable onto a network (Wang, 2005), is an excellent language for database applications. What is needed is a way for Java applications to talk to a variety of different databases. JDBC (java database connectivity) is the mechanism for doing this.

JDBC extends what can be done in Java. For example, with Java and the JDBC API, it is possible to publish a web page containing an applet that uses information obtained from a remote database. An enterprise can use JDBC to connect all of its employees (even if they are using a conglomeration of Windows, Macintosh, and UNIX machines) to one or more internal databases via an intranet. With more and more programmers using the Java programming language, the need of easy database access from Java is continuing to grow.

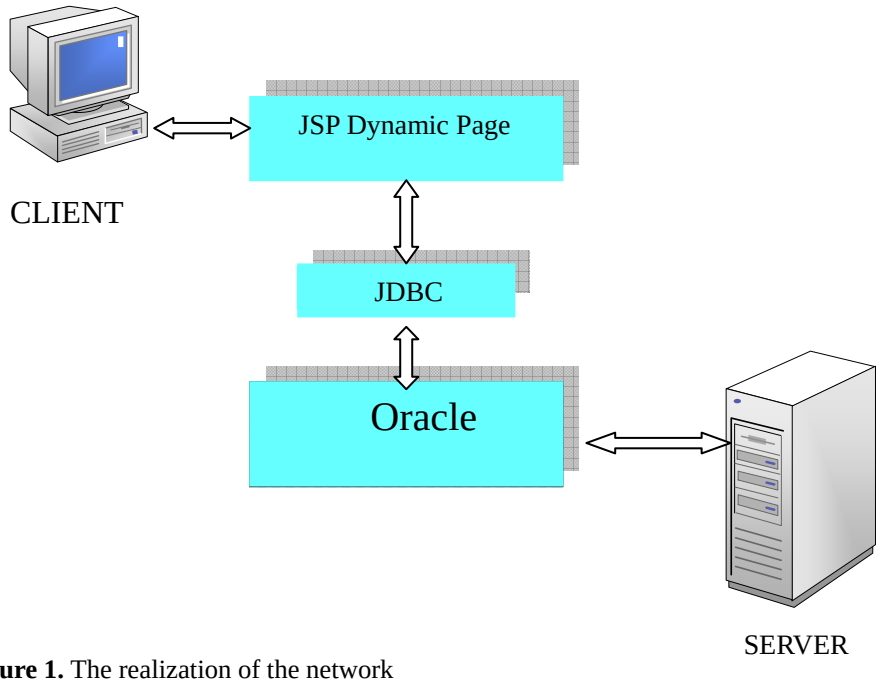

Figure 1. The realization of the network

The system inherits all the advantage of the former systems, while at the same time, it can support central registration in one data center, using data sources from other data centers. From the index page, one can search and download any files for free, and upload data from the portal.

5 A BRIEF INTRODUCTION TO THE DATABASE ON THE INTERNET

The website is available at two domain names: http://gp.wdc.cn and http://wdc.geophys.cn. Both point to the main server of the World Data Center for Geophysics, Beijing. The database is also a part of the geosystem science data-sharing platform in China.

5.1 Introduction to the website

Figure 2 gives a look at the core part of our website: on the left is a list frame: Geophysics Data, the World Data Center System, Research Organizations in China, Geophysics Observation in China, Related Links, etc. On the right are notices of meetings, news, etc. All links in the left frame lead to the corresponding pages.

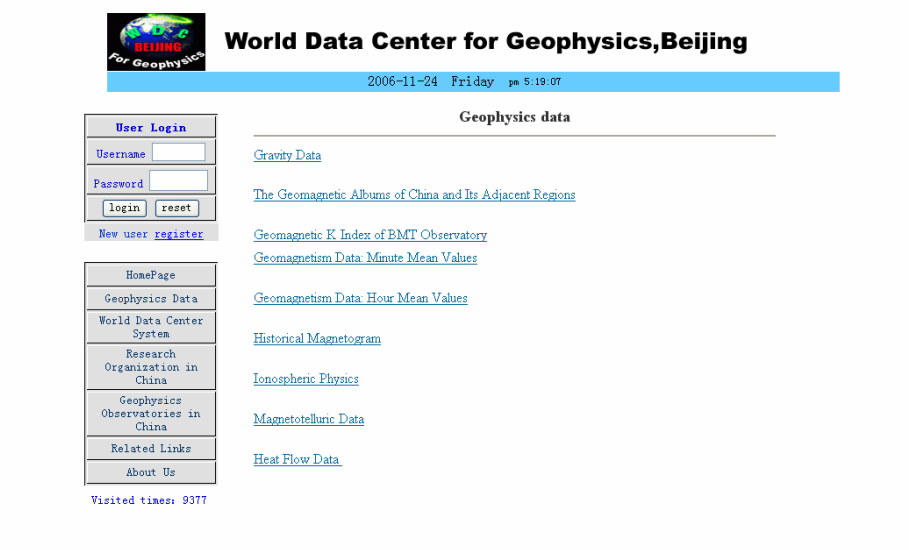

Figure 2. Database groups

For example, if one clicks on Geophysics Data, the system will list the available datasets, which include Gravity Data; The Geomagnetic Albums of China and its Adjacent Regions; Geomagnetic K Index of BMT Observatory; Geomagnetism Date: Minute Means Values; Geomagnetism Data; Hour Means Values; Historical Magnetogram; Ionospheric Physics; Magnetotelluric Data; Heat Flow Data; and Sub databases. Each of them has different search methods, including searching by site, provinces, city, county, and latitude or longitude or on both.

Figure 3 illustrates the selection of Geomagnetism Data: Minute Means Value. One can choose start time, continue time, and display format including data and charts. Also, one can click "see the data list" to get all the dates, and print or save them. Figure 4 illustrates another search result, with four different color curves representing different geomagnetic elements.

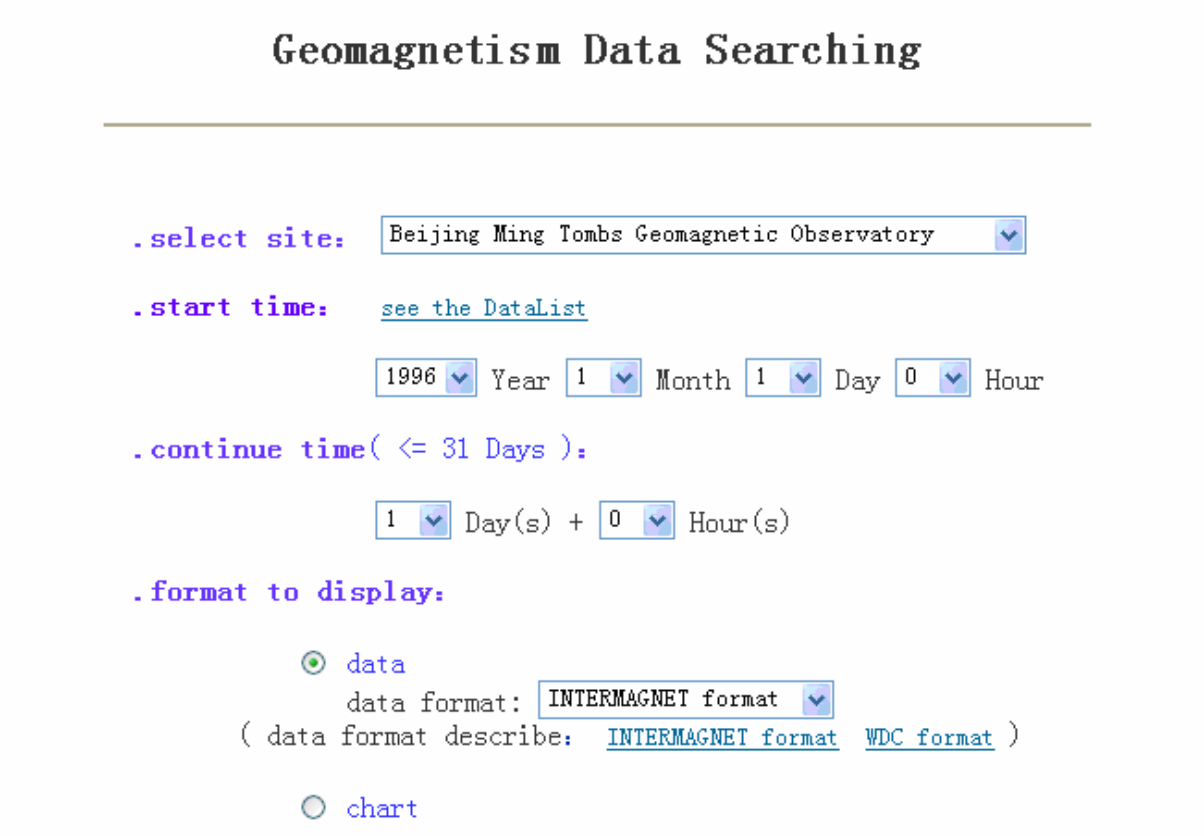

Figure 3. Geomagnetic minute data searching page

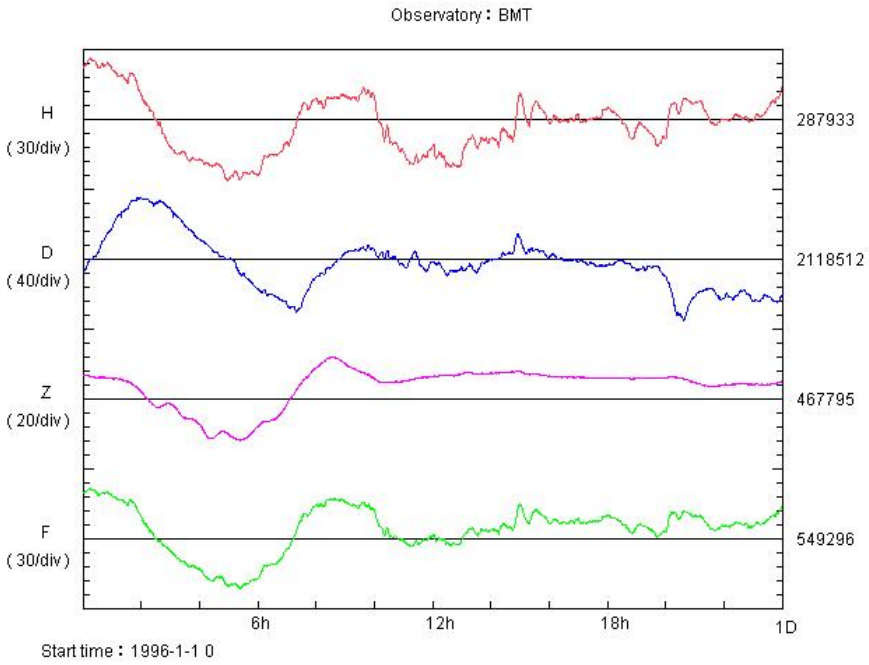

Figure 4. Dynamic magnetogram of the BMT geomagnetic observatory

5.2 Introduction to the website

The homepage of website http://wdc.geophys.cn includes news, data search, data upload, resources navigation, etc. The tree menu on the left contains databases catalogued by disciplines such as gravity, geomagnetism, geodynamics, and geothermics. Each searches for data; also keywords can be typed in the middle of the page.

6 CONCLUSIONS

Geophysics contains a multiplicity of data sources and large volumes of data. Through wide collection, international exchange, and mirroring, a large amount of valuable geophysics data has been collected in our data center. The cost for obtaining the data held in WDC for Geophysics, Beijing will not more than the cost of copying and sending the requested information. Online access is always free.

In the past years, hundreds of personal users and organizations from the fields of sciences, education, government, and industrial enterprises have obtained and utilized this data. In the future, we will edit more data to put on the web and provide more convenient service to the users.

7 REFERENCES

Gao, M., Lu, W., Wen, X., et al. (1992) China Geophysics Database in WDC-D for Geophysics. *CODATA Bulletin 24(2).*

Sun, J., et al. (2006) *China Geo-data Sharing Union* (Chinese-English Bilingual version).

Yan, Y. (2004) *An introduction to Java server pages.* Jilin Electronic Publishing.

Wang, W. (2005) *Java programming language.* Jilin Electronic Publishing.