全球沉积岩数据库(SEDBA)的现状及展望

南君亚

(中科院地球化学研究所,贵阳,IGCP269项目中国工作组组长)

全球沉积岩数据库(简称SEDBA)在UNESCO-IGCP和IUGS支持下,由IGCP 269项目设计建成。建立SEDBA的目的是实现对全球沉积岩数据的存储、管理、检索和利用,以促进信息的国际交流和全球地质对比研究的开展。

SEDBA可存储各类沉积岩样品的采样地点、构造位置、地质年代、地质特征、结构构造、沉积作用及环境、古地理等宏观信息,以及化学分析、组分及模态分析、碳质及烃类产出特征、重矿物分析、粒度分析、工程及物理参数测定、化石及孢粉分析、流体包裹体测温及各种流体特征、X射线衍射分析等分析测试数据,并附文献库、图件登记文件和采样登记文件。

SEDBA系统由26个文件按三个层次构筑而成。目前研制成功的FOXSEDBA和PARADOXSEDBA两套管理软件可在PC系列微机上实现对SEDBA的管理和应用。

SEDBA今后的发展方向,将通过国际合作,实现SEDBA与其它有关数据库的联网及通讯,形成一个多功能、综合性的数据库系统,在沉积模型及模拟、盆地分析、岩相古地理、层序地层学、沉积地球化学、能源及矿产资源勘探和开发等课题中开展应用研究,更新和完善SEDBA的结构及功能。

A GLOBAL DATABASE IN SEDIMENTARY PETROLOGY (SEDBA)

Nan Junya

Chairman of Chinese Group, IGCP Project 269

Institute of Geochemistry, CAS, Guiyang, Guizhou, China

ABSTRACT

A global database in sedimentary petrology (SEDBA) was developed by IGCP Project 269 under the support of UNESCO-IGCP and IUGS. The author analized the significance of establishing SEDBA, introduced the main achievement of IGCP Project 269 and described the charac-

teristics and functions of SEDBA. According to the suggestion of sedimentologists at home and abroad, the author also proposed an imagine on developing SEDBA in the future, that is to develop the application of SEDBA in geosciences by realizing the connection and communication within different databases and by using the connected databases to solve actual geological problems.

SIGNIFICANCE OF ESTABLISHING SEDBA

With the development of geosciences and the application of new techniques in geosciences, geological information rapidly increased. The traditional methods of data management can not meet the need. The development of geosciences will be hampered by the discordance between abundant information and backward data management.

Since the electronic computer was invented in the world, a great deal of information including letters, numbers, symbols, curves, geometric graphs, photos and voice can be stored in computer and can be transferred, caculated or processed easily. The function of computer is further enhanced by the database technique. Therefore, the emergence of computer and database technique opens a new epoch in the information science, which creates a favorable condition for automatically managing and processing geological information. The establishment of geoscience databases will make a great contribution in widening the application scope of data, raising the utilization ratio of data, improving the function of processing data and promoting the quantisation and modernization of geosciences.

According to the reports, so far there have been about 500 databases on geosciences in the world, which involved various fields of geosciences and gave the rein to the positive role in geological research, retrieval of geological information and bibliography as well as productive and economic administration.

In developming the geoscientific databases, however, one of important problem is that the utilization scope of the databases is often restricted because they were designed for a specific objective or in a narrow scope, which couldn't be adapted for international exchange of geological information. Thus, it is necessary to create some geological databases in terms of international cooperation.

A Global Database in Sedimentary Petrology (SEDBA) is a system for storing, retrieving and utilizing petrological data pertaining to sedimentary rocks, which was developed by IGCP Project 269 and supported by UNECSO-IGCP and IUGS. The SEDBA project began in 1988 and chaired by Prof. Niichi Nishiwaki-Nakajima of Nara University of Japan. About 60 researchers from 23 countries joint the project. As SEDBA is an international database system, in which above problem has been well solved. The geologists of every countries can retrieve the data pertaining to sedimentary rocks in the world from SEDBA. Therefore, it is useful to promote the international exchange of sedimentological information and the international geological correlation researches. SEDBA as a useful tool

can be used in basin analysis, sedimentological modeling and simulation, paleogeography, sedimentary geochemistry, energy and water resources as well as mineral deposit exploration, etc. The construction of SEDBA provided a successful experience for creating an international geological database.

MAIN ACHIEVEMENT OF IGCP PROJECT 269

To create a global database in sedimentary petrology, a series of complicated problems were faced by us. Firstly, as sedimentary rocks exist the characteristics of diversification in type, origin and information as well as the complication of spatial distribution, and it must be considered that all of information in the database should be useful in geosciences, we had to design a prototype database which should be of an excellent data structure and could store and retrieve most of useful information relevant to sedimentary petrology and sedimentology. Secondly, since SEDBA should be an international database, it must be considered how to disperse the system to all countries, in which there are different hardwares and softwares of computer, and how to organize and manage the system in the world. Thirdly, as the SEDBA was produced by the colleagues from different countries, it was not easy to get the uniform design scheme and create the prototype database during each annual meeting of IGCP Project 269.

Though there were so much difficulty mentioned above, IGCP Project 269 successfully managed to accomplish the plan in time under the effort of the members.

Establishment of standard on SEDBA

The ENCODING TEXTBOOK FOR A GLOBAL DATABASE IN SEDIMENTARY PETROLOGY was worked out and published, which is referred to as the standard for creating SEDBA. According to the standard, SEDBA contains the data on the geographical location, tectonic setting, geological properties, depositional process and environment, paleogeography, sedimentary structure and texture, and various analysis values of following eight types rocks: clastics, pypoclastics, carbonates, evaporites, cherts, phosphates, metalliferous deposits, coal and lignite of any geological ages, except for the unconsolidated sediments. The collective sheet for bibliography and graphic documents are also included. Although the actual graphic data haven't been managed in SEDBA so far, it will be realized in the near future.

Within SEDBA, the rock SAMPLE is treated as the basic unit of information access. A number of samples are referred to as a SET OF SAMPLES as long as they are related to each other in geological point of view. The definition of a SET OF SAMPLES requires the judgement of the data submitter.

The data are arranged in a series of data files on the basis of three levels.

The first level embraces following five files: COVER1, COVER2, HEADER SHEET, COLECTIVE SHEET FOR BIBLIOGRAPHY, COLECTIVE SHEET FOR GRAPHIC

DOCUMENTS, which are used to record the information about a set of samples and the general information about a sample, respectively. The second level consists of seven SPECIALIZED SHEETS for recording the general geological properties of clastics, pyroclastics, carbonates/evaporites, cherts, phosphates, metalliferous deposits, and coal/lignite, respectively. The third level contains 14 kinds of DATA SHEETS which are used to store the data about chemical analysis, carbonaceous material, hydrocarbon occurence, microprobe point analysis, isotopic analysis, component and modal analysis, fossils or pollens, granulometry, fluids characters and analysis, heavy minerals, X-ray diffraction, engineering and physical properties, intrinsic graphic documents and sampling procedure of sample, respectively. According to such data structure, SEDBA can be constructed as a relational database system easily.

The List of Recommended Terms for SEDBA has been compiled and published, in which more than 2400 terms about 10 respects of sedimentology are embraced. The FIPS Geopolitical code is also attached in the list. Some specific fields in SEDBA are coded, the terms of which are listed in data sheets or textbook. The abbreviation terms are also established and listed in the textbook. The establishment of these terms and codes realized the data standardization and normalization.

According to the standard, SEDBA was designed as such a kind of database that is independent of the type of computer and the data management system by choosing delimited ASCII file format as the transformation format, which can be recognized by every type of computer and translated to other format of data file by some database softwares. Consequentely, the choice of environment on running SEDBA is rather free. In this way, the most difficult problem has been solved and the SEDBA can be dispersed to every countries.

Construction of SEDBA management softwares

On the basis of the standard on SEDBA, two kinds of SEDBA database management softwares, PARADOX SEDBA and FOXSEDBA, have been produced by French Group and Chinese Group, which provide the actual tools for creating, managing and retrieving data files of SEDBA. According to the decision of IGCP Project 269, each database management system was composed of three independent softwares which have different usage. They are used for data submmiters, SEDBA local officers and users, respectively. Most of functions of DBMS, such as APPEND, EDIT, DELETE, BROWSE, RETRIEVE, LINK, OUTPUT, etc., were included in different softwares. Also, each software was recompiled by runtime version. In this case, these softwares are not only easy to disperse to every countries without copywrite problem, but also easy to be used by different users.

The minimum requirement of hardware and software for supporting PARADOX SEDBA and FOXSEDBA is as follows:

- IBM-PC/XT or its compatible machine;
- RAM equal or more than 640Kb;

• Hard disk with at least 5Mb free space;

• MS-DOS version 3. 0 or upper.

About 400 example data have been collected from different countries and stored in the related data files by using SEDBA database management softwares, which were used to examine the function of SEDBA prototype database and its softwares. Through the examination, it proved that the structure of SEDBA was reasonable and feasible, and the functions of its softwares were powerful.

Organization on SEDBA

To organize and manage SEDBA in the world, a SEDBA Central Office and eight SEDBA Regional Offices including China, East Asia, South Asia, Middle East, Africa, West Europe, North America, Australia and New Zealand, have been established.

The SEDBA Central Office is situated in Japan and responsible for the database as whole, especially for the maintenace of data structure and the control of ID numbers. It must also be the center of data communication, that is, it will accept data from regional offices and distribute data to users and regional offices. The central office will be also in charge of the communication with WDC or other related database systems.

The Regional Offices will be responsible for the management of the data itself, that is, to collect and update the data within its region, describe the data on the data sheets and enter the data into files. If necessary, it will also construct a database which is based on the standard of SEDBA but modified according to the requirement of the users in that region, for example, to translate the charecters into specific language. Therefore, the regional offices will be the service center for the specific region.

So far, some regional offices, such as Russia, East Europe and South America, haven't been established.

Distribution of SEDBA in the world

Besides the member contries of IGCP Projec 269, SEDBA has been distributed in many countries through a series of international or national training courses or short courses as well as the introduction in conferences or publications. According to the incomplete statistics, up to now, SEDBA has been dispersed to Albania, Algeria, Australia, Bangladesh, Bulgaria, China, Cuba, Czech, Egypt, Ethiopia, France, Germany, India, Indonesia, Iran, Italy, Japan, Poland, Romania, Russia, Thailand, Turkey, UK, USA, Yemen, Yugoslavia, etc. The sedimentologists of these countries have known the SEDBA and mastered the techniques of using SEDBA very well. They will become the core members to develop the application of SEDBA. More and more geological organizations, institutions and scholars are interested in SEDBA and pay attention to its development. Some of international and domestic organizations on information administration would like to connect with SEDBA. Hence, SEDBA is of a good prospect in the world.

THE FUTURE OF SEDBA

As the main target of IGCP Project 269 was to construct a prototype database for SEDBA, a lot of problems, such as how to collect a large number of data entering in SEDBA, how to access the graphic data in SEDBA, how to improve and upgrade SEDBA, how to realize the connection and communication with other geological databases, how to develop the application of SEDBA in sedimentology and related geosciences, etc., have not been well solved. In other words, the future task should be to transfer SEDBA from prototype database to an actual database and to develop its application in geosciences.

1. On data collection

The key of transfering SEDBA from prototype database to an actual database is how to collect data from geologists continuously. In general, geologists would like to submit their data, when they can obtain data from the database. In other situation, geologists would like to submit their data and to use SEDBA, when they understand how to prepare data for SEDBA and how to use SEDBA in their studies. The latter could be solved by organizing a series of training courses or short courses. But the former may be difficult to solve, because there are only a few data in SEDBA up to now. Consenquetly, at the preliminary stage, the active supporters, including all of members of IGCP Project 269 and other sedimentologists who are interested in SEDBA, should be requested to submit their own data to SEDBA. In this way, the data can be accumulated uninterruptedly and the SEDBA will become an actual database gradually.

As the SEDBA management softwares, for example FOXSEDBA, were divided into three independent parts: FOXSED1, FOXSED2 and FOXSED3 which are mainly used for data entry, ID control and data management, data retrieval and output, separately, the geologists may create their own database and develop the application by FOXSED1 and FOXSED3. In this case, many smaller SEDBA databases can be established and various application will be developed. Through a series of coordinated meetings or symposia, these databases could be merged to become a large actual SEDBA database.

Following three laws could be considered for encouraging geologists to submit their data to SEDBA: (1)determining the extents of their sharing SEDBA in accordance with their contribution for SEDBA; (2)having a cooperation with some special serials or magazines and ruling that no paper can be published unless the author conscientiously and carefully completes the SEDBA data sheets; (3)the paid service of SEDBA should be used for the users who need data but don't submit data to SEDBA.

2. On graphic data access in SEDBA

In sedimentological studies, a large number of graphic materials, e.g. maps, figures, drawings, photos, etc., have been accumulated, which are quite important informatin and must be included in SEDBA. In SEDBA, there are two files to have been created, which are

used to store some information about graphic documents, such as type, author, object, size, color, etc. of a graphic document. In fact, those are only registriston files, not the actual graphic data files. As well know, graphic data files differ from general data files in format and storage condition. To realize the access of graphic data in SEDBA, we need to used some capacious equipments, e. g. CD-ROM, and develop specific programs to manage and process graphic data. The interface between general data files and graphic data files should also be produced.

3. On updating and upgrading of SEDBA

At present, SEDBA is only a prototype database in which the structure may be modified in accordence with the practice in the future. For example, some fields in data files may be useless which should be canceled from the files; some new kind of data should be added in SEDBA. Morever, according to the requirement of users, some softwares should be improved and some specific functions of SEDBA softwares can be implemented. Therefore, it is necessary to update and upgrade SEDBA graduately.

4. On application of SEDBA

Developing the application of SEDBA in sedimentology and related geosciences is the most important respect in the near future. The application of SEDBA could involve many respects, such as sedimentological modeling and simulation, sequence stratigraphy, basin analysis, paleogeogarphy, sedimentary geochemistry, exploration of mineral deposits and energy resources, and so on. The problem is how to develop the application of SEDBA, which includes two meanings: what kind of theories and methodologies should be used for developing the application and who will do it?

According to the experience of IGCP project 269, without international cooperation under the support of UNESCO-IGCP and IUGS, it were difficult to construct such a kind of complicated database during a short term. Similarly, without international cooperation, the development of SEDBA would be slow and hard in the future. Therefore, it is a good channel to propose a new project to IGCP Board. So far, there have been some similar databases to be developed in different organizations without any communication. This is not good condition for information exchange. On the basis of the suggestion of many sedimentologists, the emphysis of new project should be to develop the database application in geoscieces by realizing the connection and communication within different databases and by using the connected databases to solve actual geological problems with examples. If the connection and communication within different databases can be realized, the scope of information will be enlarged and a more efficient and integrated database system will be produced in the world, which will lead more integrated researches in geosciences and promote the development of SEDBA.

Firstly, we should consider what databases will be connected and communicated by SEDBA. It is suggested that in the first stage, we can choose GEORECORD developed by CASP of UK and the National Stratigraphical Database produced by MGMR of China as

the connected databases. With the development of the project, other databases will be graduately added. Through analizing and comparing the characteristics of the databases, the content and data structure of the databases, DBMSs and computers used to establish the database, and the application softwares of these databases can be clarified, on the basis of which the concrete plan and procedure for connecting and communicating them will be determined.

The method and technology for the connection and communication should be selected. Currently, the most advanced method for the communication within different databases is by using computer network system, which should be used in the project. However, there is no the condition in many countries, especially in developig countries. Consequently, it is also considered that the transformation of data by using such standard file format as ASCII file format may be used in the preliminary stage in the project. The softwares used to transform data files should be developed. Also, the transformation of graphic data in different databases should be considered.

To promote the application of the connected databases in geosciences, following application programs should be developed: (1)basic application programs with the functions of retrieval, input and output, which will be suitable for all of databases to be connected; (2) graphic data processing programs; (3)application programs used in sedimentology and related geosciences.

REFERENCES

Nan Junya, 1989. Development plan of a global database in sedimentary petrology, News Letter of Chinese Society of Mineralogy, Petrology and Geochemistry, No. 2, pp. 152—153., (in Chinese).

Nan Junya, 1990. New progresses on a global database in sedimentary petrology, News Letter of Chinese Society of Mineralogy, Petrolgy and Geochemistry, No. 1, pp. 37—38, (in Chinese).

Niichi Nishiwaki-Nakajima, 1990. Construction of a global data base in sedimentary petrology, Abstract of COGEODATA Symposium on Computerized Basin Analysis for Geognosis of Energy and Mineral Resources, Gustrow, DDR, p. 25.

Niichi Nishiwaki-Nakajima, 1990. A global data base in sedimentary petrology, Abstract of Geoinfo Ⅳ, Ottawa, Cancda pp. 55—56.

Ploquin A. (ed.), 1991. Encoding textbook for a global data base in sedimentary petrology, SEDBA-IGCP 269, doc. geodiffusion 1, 100pp., 6fig,. Paris.

Nan Junya, Ye Jianliu and Yang Weidong, 1991. A global database in sedimentary petrology and its prospect of application, Abstract of Chinese Fourth Mineralogy, Petrology and Geochemistry Symposium, pp. 465—466 Hefei, (in Chinese).

Nan Junya, Ye Jianliu and Yang Weidong, 1992. A global database in sedimentary petrology and its development in China, 《CODATA Bulletin》 Vol. 24, No. 2, pp. 163, 13th International CODATA Conference, Beijing, China.

Yang Weidong, Nan Junya and Ye Jianliu, 1991. List of recommended terms for SEDBA, Chinese Group of Chinese Group, 60pp.

The project supported by National Natural Science Foundation of China.