全球沉积岩数据库项目(IGCP-269)

中岛西协二一

(日本奈良市奈良大学社会研究学院)

在1988年设立了IGCP-269项目,试图借助设置的模型(prototype)数据库来检验用于沉积岩石学上的全球数据库可行性。数据库包容沉积岩的各种分析值和样本的详细资料。在个人计算机上可用对本项目来说是至关重要的,所以它也适用于发展中国家。由于研究者要获取其他地区数据来源的信息常常很困难,有必要设置地区组以汇集数据资料。一个逻辑部件相当于一个样本,并由标题程序块和内涵程序块构成。标题程序块的形式一般适用于所有类型的沉积岩,而内涵程序块对各种岩石类型是不同的。许多问题,不仅信息程序方面的技术问题,还有沉积学上的理论问题已经解决。包括检索和分析功能在内的数据库管理系统已采用几个商用软件研制出。世界上已开发了容有400件样本数据以上的模型数据库,并配置由沉积学家试验。包含巨大数量全球数据的实际数据库现在正在研制中,还要将其集成为一种认识基础,成为沉积学家研究沉积相、区、环境等方面的一种有效工具。

GLOBAL DATABASE PROJECT IN SEDIMENTARY PETROLOGY (IGCP-269)

Niichi Nishiwaki-Nakajima

Faculty of Social Research, Nara University,

1500 *Misasagicho, Nara City, Nara*, 631, *Japan*

ABSTRACT

In 1988 the IGCP Project No. 269 was established to examine the feasibility of global databases in sedimentary petrology by constructing a prototype database. The database is composed of variety of analytical values in sedimentary rocks and sample specification data. Availability on personal computers are essential for the project so that it can be used also in developing countries. Regional groups were established to collect data, because it is often difficult for researchers outside the region to get information on data sources. A logical unit corresponds to a specimen, and it is composed of a header block and an intrinsic block. The format of the header block is common for all types of sedimentary rocks, but that of the intrinsic block differs for each rock type. Many problems were resolved, not only technical problems in information processing, but also theoretical problems in sedimentology. The database management systems, in-

cluding retrieval and analytical functions were developed by using several commercial softwares. The prototype database with more than 400 sample data was developed and distributed for examination by sedimentologists in the world. The actual database, which contains a large amount of global data, is now under development, and it will be integrated into a kind of knowledge base for sedimentologists as a useful tool for studying sedimentary facies, province, environment, and so on.

INTRODUCTION

Many documentary and factorial databases are now under construction and/or in operation in various fields of science. and it is difficult for researchers to collect scientific information without using databases. The database is important for various geological researches as the principal source of data, especially in the preliminary stage of research.

There are so many projects on databases that it is necessary to discuss the basic problems of databases and to harmonize these projects. Several scientific committees on databases in geology have been established in the last decade.

In this paper the IGCP Project 269 is introduced, which has discussed many problems on global database and constructed a prototype database in sedimentary petrology.

PETROLOGICAL DATA FOR GEOLOGICAL ANALYSIS

It is important in the initial stage of geological research to examine data which have already been obtained in and around the target area by using some geostatistical methods. Data will be available from various kinds of projects such as oil exploration surveys, geological mapping, geochemical studies, geophysical surveys, etc. The accuracy and quantity of data may not be sufficient in many cases, as the members may not think that the data will be used for other purposes at the time of their collection, but they contain valuable informations for other researches.

Petrological analytical data are often included in many research report as one basic data on the surveyed area, and they are useful for many geostatistical analyses, It is helpful for the researchers if a complete set of data on the target area were obtained, which had been collected from various sources and standardized in the course of input to the database.

The final goal of our IGCP project is to construct a global database in sedimentary petrology which is composed of standardized data from the whole world. We should have a close contact with petroleum and mining geologists, who have a large amount of petrological data.

IGCP PROJECT 269

In 1985, the IGCP 163 on the global database for igneous petrology was successfully completed, and its successor project IGCP 239 has started. At the same time, the IUGS Subcommission on Data Bases for Petrology was established in 1985, and one of its working groups on sedimentary petrology continued the discussion on a global database in sedimentary petrology. According to the conclusion of the working group, a new IGCP project was proposed in October 1987, and the IGCP Project No. 269 was accepted by the IGCP Board in February 1988, in which feasibility and various related problems on sedimentary databases would be examined by construction a prototype database.

The project held six international meetings (1988 in Japan, 1989 in France, 1990 in China, 1991 in Egypt, 1992 in France, 1993 in Turkey) and many meetings of working groups for business and scientific discussions.

CONTACT WITH OTHER ORGANIZATION

Is it necessary to get a consensus among sedimentologists on the purpose and content of the database, because it is impossible to construct in without the support of sedimentologists in the world, and because it will not be used if it does not fit with requirement of sedimentologists. so we held international meetings and/or symposia on the database in sedimentary petrology at international conferences in geology and related sciences, such as the International Geological Congress, the International Sedimentological Congress, etc.

It is also necessary to contact with information scientists who are working in geology and related sciences, as we need the up-to-date knowledge and experiences in information sciences. We held meetings at conferences of that field, such as the International Conference on Geosciences Information,

Also it is necessary to contact with other international projects concerning sedimentology and/or databases in geology, not only for avoiding the overlap of contents of the database, but also for integration of data structure and database systems. IUGS Subcommission on Databases for Petrology, the IGCP 239 Project on Databases in Igneous Petrology, and the GSGP: Global Sedimentary Geology Project.

GENERAL PRINCIPLES

Two guidelines have been established for the project. One is compactness of the content of the database, which will be composed mainly of analytical values of sedimentary rocks with minimun information for sample specification. The other is availability on personal computers considering the use in developing countries, and it was suggested to use

"IBM—PC type"computers(not only IBM—PC itself).

The DBMS was selected after the examination of various softwares. It is generally accepted that several DBMS which are widely distributed and available in the most parts of the world should be used. That is,we decided to save time and labor by using commercial softwares,rather than developing original DBMS.

It was decided to establish the regional groups,not national groups,each of which will collect data in sedimentary petrology analyzed and/or published within the region. The regional group is important for database projects because it is often difficult for researchers outside the region to get information on data in the region. The following eleven regional groups were established:1. West Europe 2. South Asia 3. China 4. Middle East 5. East Europe 6. USSR 7. USA+Canada 8. Latin America 9. Japan+Korea 10. Australia+NZ and 11. Africa.

DATA CENTER

We established one central office and several regional offices for data collection and data distribution.

The central office is responsible for the database as a whole,especially for the maintenance 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 regional office is responsible for the data itself,that is,to collect and update the data from its region,describe the data on the data sheet,and enter the data into computer files. It will also construct a database which is based on the global database but modified,if necessary,from the original one for the convenience of users in that region,for example, translation into a specific language. It means that the regional office will be the service center for the specific region.

After the construction of the actual database,the database is better to be distributed though the World Data Center which is an authorized agent for scientific databases in the world.

BASIC CONFIGURATION

The basic opinion on the data structure is as follows. A logical unit corresponds to a specimens,and it is composed of a header block and an intrinsic block.

The format of the header block is common for all type of sedimentary rocks. The key variable 'rock type' is included in the common header,and it specifies which specific header sheet is used for the specimen.

The intrinsic block is divided into the specific header and actual data. The specific

header contains a specific description of the sample, and the format differs for each rock type. The actual data contains actual analytical values, and it is composed of various kinds of data sheets, each of which corresponds to a specific type of analysis.

There were many technical problems on the data structure to be solved in our project. For example, which kinds of variables are common to all rock types, how many rock types should be defined to design different data sheet formats, etc.

There is an opinion that the header block should be organized more sedimentologically. That is, more integrated variables, such as environments, sequences, facies, etc., should also be included. These variables are important for sedimentological research, and the user may require such integrated information together with analytical values. The final goal of our project would be a kind of knowledge base, which can treat such interpreted informations.

Another important suggestion on the database is to store image data, such as photographs, drawings, etc., since they contain different information other than numeric values. For this purpose we plan to use the CD-ROM, which is now becoming popular and more easy to use.

DATA STRUCTURE

The five working groups were established corresonding to the five major sedimentary rock types for the discussion on the data structure for each rock type. One more working group was added to discuss the header block, which will be commonly used for all types of sedimentary rocks.

1. Carbonate and evaporates
2. Siliciclastics
3. Pyroclastics
4. Coal and lignite
5. Iron and siliceous deposits and phosphate
6. Common header block

Each working group discussed the details of the data structure for each rock type, and their conclusions were examined and compiled at the Second International Meeting (France, 1989). Discussion within each working group was continued by mail communication. The draft of the data structure was distributed to many sedimentologists asking valuable comments and suggestions.

A special form was prepared to describe the data structure for each rock type at the Third International Meeting (China, 1990), because such standardized description of necessary for detailed discussion of data structure, and because such a form is very helpful when designing the data sheet format.

The following characteristics were described for all the fields of variables in our

database.

1. Field name
2. Type(Logical/Numeric/character)and Length
3. Number of maximum occurrence
4. Character of field (Indexed/Searched/Output only)
5. Importance (Mandatory/Important/Significant/Useful)
6. Coded table/Recommended term/Free term
7. Example of data

DATA SHEET

Each working group prepared a draft of data sheet format for each major rock type, which was distributed to all the members and related researchers. The draft was examined, and the first version of data sheet for all rock types was determined at the Second International Meeting(France,October 1989).

A new working group for designing the data sheet was established to design and print the data sheet. The data sheet was distributed not only to the members of the project but also many sedimentologists in the world at the end of 1989,requesting suggestions for the format.

Many corrections and several modifications were made to the first version of data sheet at the Third International Meeting(China,May 1990). New intrinsic data sheets were accepted and several new analyses were added. Finally the data sheet was published at the end of 1990,which was composed of the following sheets.

Header Block	Common Header	Header Sheet
		Bibliographic Sheet
		Sampling Sheet
Intrinsic Block	Specific Header	Silisiclastics Sheet
		Pyroclastics Sheet
		Carbonates Sheet
		Evaporates Sheet
		Iron Rocks Sheet
		Phosphates Sheet
		Chert Sheet
		Coal and Lignites Sheet
	Actual Data	Modal Analysis Sheet
		Chemical Analysis Sheet
		Microprobe Analysis Sheet
		Isotope Sheet
		Granulometry Sheet

Heavy Mineral Sheet
Fluid Analysis Sheet
X-ray Diffraction Sheet
Fossil and Palynology Sheet

Consequently, the following sheets will be used to describe one specimens in our database.

1 Common header sheet
1 Specific header sheet
n Bibliographic sheets
n Actual data sheets

EXPLANATORY TEXTBOOK

The data sheet should be used not only by members of the project, but also by other sedimentologists in various regions of the world, as it is necessary to include all data sources in the world. It is important for all researchers to understand the details of the data sheet.

As a result of the discussion at the Third International Meeting, it was decided to publish a manual on the data sheet, explaining how to describe on the data sheet. Each working group prepared a draft of the explanatory textbook on all the data sheets which are used for a specific rock type. The working group on the manual was established, which published the explanatory textbook for data submitter.

TERMINOLOGY

There are three choices for describing each field of data, that is, free terminology, recommended terminology, or code. The free terminology is flexible so that each researcher can use his own terminology, but it is difficult to retrieve specific data because terminology used is different for each researcher. The code system, on the other hand, is not so easy to adopt, because the code table is generally large and complicated, but it is very easy to retrieve specific data by using the code system, because only one code is allowed for a specific description.

To avoid both difficulties, we chose the recommended terminology for many fields in our database, using only the tables of recommended terms, and it is possible to and new terms when a researcher wishes to use other terminology. Of course the efficiency of retrieval is lower than the code system, but it is higher than the free terminology.

We prepared code tables only for several fields, on which the international code is established, of which code table is small and convenient for the submitter of source data.

DATA BASE MANAGEMENT SYSTEM

At the First International Meeting (Japan, 1988), it was decided that we should use some relational DBMSs which are widely distributed. We have examined several DBMSs, such as, d-Base III Plus, FoxBase, Paradox 3, Informix, etc. It was not easy to deterimine which one is better for our database at that time because the data structure of our database had not yet decided.

We have decided that each regional group may construct a prototype database by using a DBMS which the regional group has already bought, partly because of copyright problems, and partly because of the limitation of the hardware systems of each regional group.

Each regional group can choose any DBMS, though there is only one limitation for the selection of DBMS, that is , the source data should be supplied in the ASCII format for translation of data into other DBMS.

MEETINGS AND TRAINING COURSES

In the course of our project, the following meetings were held for scientific and/or business discussion, including several workshops and symposia. Also several training courses were organized for demonstration and education of our database system, and for calling new participants in our project.

International meetings

Date	Venue	* 1	* 2
1988 June 24—29	Nara Univ., Japan	19	6
1989 Oct. 8—16	Vosges, CRPG, France	20	9
1990 May 2—9	Beijing Film Acd., China	44	8
1991 Apr. 26—May 6	Egyptian Petr. Res. Inst. Egypt	28	10
1992 April 20—26	Univ. Rennes, France	18	13
1993 June 23—July 3	Middle East Tech. Univ., Turkey	30	8

* 1 Number of participants

* 2 Number of countries

International Symposia

International Symposium on SEDBA (1993 June 28—29, METU, Turkey)

International Symposium on the Applications of Computer and Database to Sedimentology (1994 July 1—8, Chengdu, China)

Other meetings (Excepting regional/national group meetings)

Meeting	Date	Venue	* 1	* 2
Business Meeting	1989 July 13	28th IGC, Washington	15	9

Business Meeting	1990 June 18	Geoinfo IV, Ottawa	12	9
Business Meeting	1990 Aug. 28	13th IAS, Nottingham	8	8
WG on Textbook	1990 Oct. 12—18	Univ. Nancy, France	10	4
Business Meeting	1992 Sept. 1	29th IGC, Kyoto	16	13

* 1　Number of participants

* 2　Number of countries

Training courses

Course	**Date**	**Venue**	*** 1**	*** 2**
Short Course WB-4	1992 Aug. 29	29th IGC, Kyoto	10	7
Short Course	1992 Oct. 5	3rd Sed. Conf. China	40	1
Training Course	1992 Dec. 8—12	Univ. Nancy, France	19	11
Training Course	1993 Fed. 22—24	Nara Univ., Japan	4	3
Training Course	1993 June 25—27	METU, Ankara, Turkey	30	8

* 1　Number of participants

* 2　Number of countries

SUMMARY OF THE PROJECT

The feasibility of the global database in sedimentary petrology was discussed for the last five years through the construction of the prototype database. In the course, the detailed data structure was determined in relation with the use in sedimentology, the data sheet format was designed for data description, the explanatory textbook for data submitter was published and its revised version will be coming soon, the runtime versions of database management system were developed for Paradox and Foxbase, and, more than 400 example data were entered into the prototype database SEDBA version 1.1, which has already been distributed to the world with runtime softwares, and training courses were also held. Actual database with large amount of data can be constructed by upgrading and extending our prototype database.

It is concluded that the global database is possible and necessary to be constructed in the successor project, as the purpose of our project to construct the prototype database has been completed.

MAJOR ACHIEVEMENT

The project was initiated by the two kinds of organizations. The one is the network for communication and data collection composed of one central office and several regional offices. The other is the working groups, each of which is responsible to discuss and work on a specific problems.

The data structure was carefully examined so that we can keep complete compatibility between Paradox and Foxbase versions. It should not be modified unless public examination was completed, as it is not fair nor acceptable if we modify frequently only by our own opinion. That is, we plan to modify the data structure after 3 years when we obtain a consensus through training courses and symposia on our database, where we may receive valuable comment and suggestions from many sedimentologists.

The prototype database has already been constructed by using several commercial softwares on IBM—PC type personal computers. It has been distributed to the world with free of charge, and it has been used and examined at many laboratories and institutions. Also we have reported on our prototype database at several scientific conferences and meetings, together with its demonstration by using the personal computers. Many sedimentologists are interested in the database, and express to join our project.

International training courses and short courses have already been held at several places, and we have got many many new participants to our database project. Now we have a global network of regional centers, each of which is responsible to data collection within the region.

It is concluded that the golobal database in sedimentary petrology is possible to be constructed by expanding our prototype database, though it is also necessary to upgrade our system in many points, such as, updating the retrieving procedures, development of application programs, acceptance of graphic data, use of new hardware for mass data storage, conversion to other operating systems and machine types, etc.

It is necessary for the development of our database system to demonstrate the examples of application to sedimentology, for example, quantitative stratigraphy, sequence stratigraphy, basin analysis, simulation of sedimentary processes, paleoenvironmental analysis, etc. It is also important to develop new ideas and techniques to apply our database for the more integrated research in sedimentology, such as, the development artificial intelligence and knowledge base.

We held the first symposium on our database in Turkey in 1993, and the second symposium in China in 1994. These symposia are important for discussing the future development of our database and its application.

CONCLUSION

According to the discussion through the construction of the prototype database, it is possible and necessary to construct the global databases in sedimentary petrology. Our project was completed as the prototype database has been constructed, which is the final goal of our project. The successor project should be organized to construct actual database by upgrading and expanding the prototype database.

DEVELOPED SOFTWARES AND DATABASES

S. Nion and A. Ploquin(1991) SEDBA:Runtime version of Paradox for SEDBA. 2 floppy disks(3. 5′ 2HD).

W. Yang, J. Nan, and J. Ye (1991) FOXSEDBA, Runtime version of Foxbase for SEDBA. 3 floppy disks(5′2DD).

IGCP 269 WG on Database (1992) Prototype SEDBA with runtime versions of Paradox/Foxbase and its brief manuals. Version 1. 1. 3 floppy disks(3. 5′2HD).

SPECIAL PUBLICATIONS

IGCP 269 WG on Data Sheet (1990) SEDBA data sheet, Version 1. CNRS-CRPG, 30 pp.

A. Ploquin ed. (1991) Encoding textbook for a global database in sedimentary petrology SEDBA IGCP 269. Doc. Geodiffusion, no. 1, 100 pp.

W. Yang, J. Nan and J. Ye(1990) List of recommended terms for SEDBA. Chinese Group of IGCP Project 269, 56+4 pp.

J. Ye, J. Nan, and W. Yang(1992) User's Manual for FOXSEDBA. Short Course Note, 3rd Nat. Sed. Conf. China, Chengdu, 20 pp.

J. Nan, J. Ye, and W. Yang(1992) Design and construction of the global database in sedimentary petrology. Short Course Note, 3rd Nat. Sed. Conf. China, Chengdu, 34 pp.

J. Ye, J. Nan, and W. Yang(1992) FOXSEDBA: a kind of SEDBA database management system. Course Note for the International Training Course on SEDBA, 23pp.

S. Nion and A. Ploquin(1992) A user's manual for sedimentary Database SEDBA with Paradox Runtime. Course Note for the International Training Course on SEDBA, 35 pp.

S. Nion and A. Ploquin(1992) Paradox SEDBA screen formats for data entry. Course Note for the International Training Course on SEDBA, 27 pp.

G. Eseller and J. Harbaugh, eds. (1993) Explanatory textbook on SEDBA. Middle East Technical Univ., Ankara(in press).

REPORT TO SCIENTIFIC JOURNALS

N. Nishiwaki(1988)On the start of the IGCP Project 269. Jap. Jour. Geol. Soc., 94(6)470.

N. Nishiwaki(1988)New IGCP Project on a data base in sedimentary petrology. Abst. Ann. Meet., Jap. Soc. Geol. Data Proc., 23.

N. Nishiwaki(1988)Proposal procedure of new IGCP project, with an example of the IGCP project 269. DELP News, (25)18—21.

N. Nishiwaki-Nakajima(1989)Construction of global database in sedimentary petrology. Abst. 28th IGC, Washington, (2)521.

N. Nishiwaki-Nakajima(1988)Introduction to the IGCP Project269. Jour. Sed. Assoc. Jap., (29)40—41.

N. Nishiwaki-Nakajima(1989)Report on the first international meeting of the IGCP 269. Jour. Geol. Soc. Jap., 94(9) 726.

J. Nan(1990)A global data base project on sedimentary petrology. Bull. Mineral. Petr. Geochem., 1990(1)37—38.

N. Nishiwaki-Nakajima(1990)Construction of a global data base in sedimentary petrology. Abst. Intern. Symp. on Computerized Basin Analysis, Gustrow, 25.

N. Nishiwaki-Nakajima(1990)Construction of a global data base in sedimentary petrology. Abst. Geoinfo IV(Intern. Conf. Geosci. Inform.), Ottawa, 22.

N. Nishiwaki-Nakajima(1992)News on SEDBA: Sedimentary database developed by IGCP 269; distribution of its prototype and its training course. IAS Newsletter, (121)4—5.

N. Nishiwaki-Nakajima(1992)A global database in sedimentary petrology(IGCP 269). Y. Kuroda and R. Tsuchi eds., Japan Contribution to the IGCP, IGCP Nat. Comm. Jap., 33—39.

N. Nishiwaki-Nakajima, J. Nan, A. Ploquin, and J. J. Royer(1992)A global database in sedimentary petrology. Abst. 13th CODATA Conf., Beijing, 134—135.

N. Nishiwaki-Nakajima(1994)Global data base in sedimentary petrology. Proc. 4th Intern. Conf. Geosci. Inform. Ottawa 1990, Geol. Surv. Canada Open file 2315, pp. 309—316.