RoS – A NEW DATABASE SYSTEM IN THE GLIWICE RADIOCARBON LABORATORY

NATALIA PIOTROWSKA[1], MAŁGORZATA SZCZEPANEK[1], ANNA PAZDUR[1] and WIESŁAW ZAJADACZ[2]

[1]*Department of Radioisotopes, Institute of Physics, Silesian University of Technology, Krzywoustego 2, 44-100 Gliwice, Poland*
[2]*Zawadzkiego 42, 44-100 Gliwice, Poland*

Key words: RADIOCARBON, DATABASE, RoS

Abstract: The idea of creating radiocarbon database came up in late 80's. In this time the staff of the Gliwice Radiocarbon Laboratory started developing the local database GdRDB that should contain all dates of samples measured in this laboratory. This first database had a simple structure – only one file. Later modification of GdRDB database in 1994 introduced changes in its structure, and the modified database had seven files. The modification was made in order to adapt GdRDB to databases already existing (in Japan, Great Britain and Yugoslavia) and to make it match the standard established by the International Radiocarbon Database (IRDB).

The new database system was developed in the Gliwice Radiocarbon Laboratory in 2002. This system consists of five relational databases containing all information about submitters and institutions, samples, sites and the dating process. The access to these databases and management of data is provided through an application called *Register of Samples Editor* (RoSE). The RoS system is supplied with the *Data Analyse* application enabling analyses of database content and *Web Assistant* programme, which generates web pages containing information about dated archival samples.

1. INTRODUCTION

In the middle 80's there was about 3000 of radiocarbon dates already collected in the Gliwice Radiocarbon Laboratory. This situation gave rise to the idea of creating a local radiocarbon database (Pazdur and Prowoł, 1987). The first presentation of GdRDB database that contained 250 radiocarbon dates took place during the 3rd Conference "Methods of Absolute Chronology" in 1989 (Michczyński and Pazdur, 1989). The modification of the GdRDB database in 1994 brought changes in its structure. The novel structure consisted of seven files (**Fig. 1**, Michczyński and Pazdur, 1994). The modification was made in order to adapt the structure of the local database to already existing radiocarbon databases as well as to the proposed structure of the International Radiocarbon Database (IRDB). At that time several similar database structures had already been described by Moffet and Webb, 1983 (United Kingdom), Englesman *et al.*, 1986 (The Netherlands), Wilcock *et al.*, 1986 (United Kingdom), Omoto, 1989 (Japan), Obelič, 1989 (Yugoslavia), and van der Plicht, 1992 (The Netherlands). Most of these local databases had structures in very close to the proposed IRDB configuration. Unfortunately, due to many obstacles, the project of creating one international radiocarbon database has never been realised. Development of a new, much more efficient AMS dating technique resulted in rapid increase of number of radiocarbon dates. Many dates, which have been published, were spread in lots of articles and gathering them was a very laborious task. Problems arose also from the deficiency of funding for IRDB project (Walker and Kra, 1988; Kra, 1989a; Kra, 1989b).

Since that time some databases have been developed and maintained for the specific purposes, mostly for historical and archaeological applications including time and space interpretation of radiocarbon datasets (Michczyński *et al.*, 1995; Orlova *et al.*, 1998; Zaitseva *et al.*, 1998; Ashmore *et al.*, 2000; Burton and Levy, 2001). There were also several papers on international and local soil databases published in 1996 (Becker-Heidmann, 1996; Becker-Heidmann *et al.*, 1996; McNeely, 1996; Scharpenseel *et al.*, 1996).

Nowadays, when the availability and capacity of computer hardware and software are much more advanced than even ten years ago, it is difficult to imagine a running radiocarbon laboratory without any computer support - and it is supposed, that all of them apply more or less advanced storage and processing of data. The perfect example of advance in modern database formation was given by Suckow and Dumke (2001), who described the data model for laboratories applying many different methods of research. It is also worth to call the databases available through Internet, which can be found, e.g., via links on www.radiocarbon.org.

2. DATABASE GdRDB

The first radiocarbon database GdRDB, in the Gliwice Radiocarbon Laboratory was developed for dBASE III system platform and installed on IBM PC/XT computer running under PC - DOS v.3.1. The central element of this database was the file *RDB.DBF*, which contained all information about sample location and date. The database programme used also several index files (*RDBSITAL.NDX*, *RDBLABNO.NDX* and *RDBDATE.NDX*) to sort records: by the name of series and sample location, by the laboratory number, or by the date of measurement.

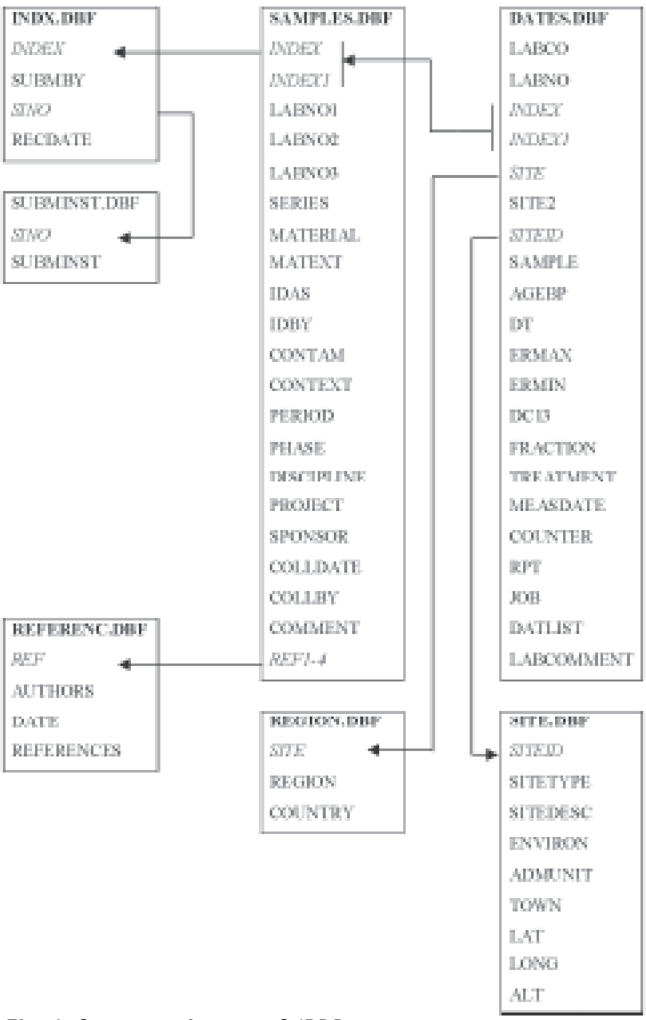

***Fig. 1.** Structure of system GdRDB (after Michczyński and Pazdur, 1994)*

Operation of this database was realised through the programme stored in the file *SERVRDB.PRG*. It could browse and sort the records, and print out information about dates selected by the user. This programme was using sub-programmes and procedures stored in the files: *RDBORDER.PRG*, *RDBSORT.PRG*, *SHORTREP.PRG*, *LONGREP.PRG* and *RDBPROC.PRG* (Michczyński and Pazdur, 1989).

The next step in development of the database GdRDB in the Gliwice Radiocarbon Laboratory was made in 1994. The database was modified and the structure contained seven files (Michczyński and Pazdur, 1994). Basic relations between these files are presented in **Fig. 1**. The files were related one to another with key fields *INDEX* and *INDEX1* (fields defined by the sample position in the register), *SITE* and *SITEID* (the sample location and its code), *SINO* (the code of submitting institution) and *REF* (references).

3. RoS SYSTEM

The most recent modification of the database maintained in the Gliwice Radiocarbon Laboratory started in 2002. This version of radiocarbon database, which has been called RoS system, consists of five relational databases. The system has a client/server structure using InterBase 6.0 The Open Source Database server and was developed in Delphi 5.0 Professional the environment. A very fast access to data is provided by the server based on InterBase Express components using native drivers (http://www.borland.com/interbase/index.html).

All dates of samples processed in the Gliwice Radiocarbon Laboratory have been already entered into the system. Unfortunately, detailed information about most of samples is unavailable. Lists of samples with complete descriptions were published on the Laboratory web site: www.carbon14.pl. New samples submitted to the Laboratory are entered into the databases system on a day-to-day basis.

Structure of RoS system

The RoS system consists of five related databases, which are as follows:

- *IBGDRDB3_VB.GDB* – a basic relational database, which contains information from the sample register;
- *IBRADIOCARBON.GDB* – contains supplementary information about samples which dating results were published as ten radiocarbon date lists in „Radiocarbon”; works also as a separate programme;
- *IBWORLD. GDB* – a library with information about geographical locations of sampling sites;
- *IBGDRDB1.GDB* – an auxiliary database (a copy of GdRDB_1 database) which contains information about samples collected between 1972 and 1995;
- *IBCAL.GDB* – a database with information necessary for calibration of radiocarbon dates.

The access to these databases and management of data is realised through the application called *Register of Samples Editor* (RoSE), which works in two language versions: Polish and English.

There are 29 tables connected by primary and foreign keys. The diagram of the tables and their relations is shown in **Fig. 2**. The tables are grouped into eight groups (A-H), each one including the most tightly connected tables. Hereafter a general description of the tables and their fields included in database system is given.

Group A consists of *Register*, *Submitter* and *Institution* tables containing information relevant to entering the sample into the laboratory register. The basic key of *Register* table is *Register number* in the field no. 0. The *Register* table is connected by external keys *id_submitter* and *id_institution* with *Submitter* and *Institution* tables, respectively. The field no. 9 in the table *Institution* (*id_country*) is an external key connecting this table with *Country* table in group B.

Details concerning the location of sampling site are placed in *Country* and *Sample location* tables in group B. *Country* table belongs to *IBWORLD.GDB* database, where all modifications of its records are made. All changes in this database are immediately visible in *IBGRDB3_VB.GDB* database. *Sample location* table stores all information about the sampling place and can be modified only in RoSE programme.

Group C contains detailed information about the sample derived from an information sheet, which is filled by the submitter, e.g. sample material, contamination, collector, expected age and period, and additional information. Nine fields in *Samples* table are connected with other tables, what enables selection of predefined data (e.g. series, material, field of study).

Tables containing information about job and project, within which the dating is performed, are grouped in group D. *Dates number* table has *Job*, *Project* and *Payer* tables attached and connects also *Samples* table with *Laboratory* or *Archive dates* table, where the results of dating are stored (Bookmark E). Here a boolean field no. 7 (*Archive*) has a special function. It stores information if the sample is still being processed in the laboratory (or the information is incomplete) or is already archived. The archival samples are required to have the fields containing basic data, such as laboratory number, measurement technique, age, uncertainty, preparation method, date of measurement, filled in and only these samples are available for advanced analyses or for publication in Internet.

Two tables: *Laboratory* and *Archive dates* form group E, where the details concerning dating process and results are stored. They are related to tables in F, G and H groups by a number of key fields. The tables in group F keep information about the sample with respect to the laboratory dating procedure, including different steps of preparation for different measurement techniques, and stores information about measurement progress. Group G contains three tables tightly connected with *Laboratory* and *Archive dates* tables. In *Type of date* table the type of obtained result is classified as "finite" with symmetrical or unsymmetrical uncertainty, "infinite", "less than" or given in pMC. *Laboratory number* table contains unique identification number of dating result, where the prefix depends on the measurement technique (Gd- for gas proportional counting, GdS- for liquid scintillation counting and GdA- for AMS). Group H comprises two tables with information relevant to final reports on dating. After the dating is completed and the sample is moved to Archive the report for Submitter may be automatically prepared. *Report numeration* and *Reports* tables belong to *IBCAL.GDB* database.

Handling RoS

The RoSE application enables managing data collected in *IBGDRDB3_VB.GDB* database, i.e. to fill in and edit fields, to modify information concerning samples (also archival samples). The fields are placed on several sheets (*Register*, *Samples*, *Laboratory*, *Archive Dates* and *Radiocarbon*) making the application user-friendly. The choice of sheet is made simply by clicking a proper icon on a toolbar. Adding information about samples to the database system consists in filling the fields in each record with information provided by Submitter (in the sample information sheet) or with information concerning laboratory dating procedure. The predefined fields are divided into two categories: libraries (*material*, *contamination*, *type of site*, *field of study*, *measurement device*) and card indexes (*submitter*, *institution*, *collector*, *site location*, *series*, *job*, *project*, *sponsor*, *pretreatment*). The card index allows easy addition and modification of items, as well as sorting or searching within records, while libraries have preliminarily defined content.

In case the new sample belongs to the register position, which has not been recorded, user should create one, using option *Insert Register Position*. In the first sheet *Register* basic information concerning register position, such as submitter, institution and other formal issues are included. If the position has already been created, a new sample is added using *Insert Sample* option. Information about sample is entered in *Samples* sheet. There are some text fields (*sample name*, *material identified as*, *carbon content*, *additional information about sample*, *date of collection*, *expected age-period*, *comment*) and predefined fields, where the choice of one of the values from libraries and card indexes is made: *material*, *contamination*, *sample collector*, *type of site*, *sample site* and *field of study*. Similar procedures are applied in *Laboratory* sheet, where user should enter information on *measurement method*, *measurement device*, *pre-treatment – fraction* extracted from the sample and *date of measurement*. After above fields are filled in and the laboratory procedure (consisting of several steps depending on measurement method) is 100% complete, the dating results are entered and an option *Finish process of dating* allows moving sample from *Laboratory* table to *Archive Dates* table. The view of *Archive Dates* sheet is similar to that of *Laboratory* sheet, although there are two more yes/no fields: *Internet* and *Calibration*. Samples with *Internet* field marked are available for publication of sample description in Internet, while, dates with *Calibration* field marked can be forwarded to the calibration programme linked with *IBCAL.GDB* database.

Radiocarbon sheet provides access to descriptions of samples, which dating results have been published in *Radiocarbon* journal as ten date lists. It may be also possible

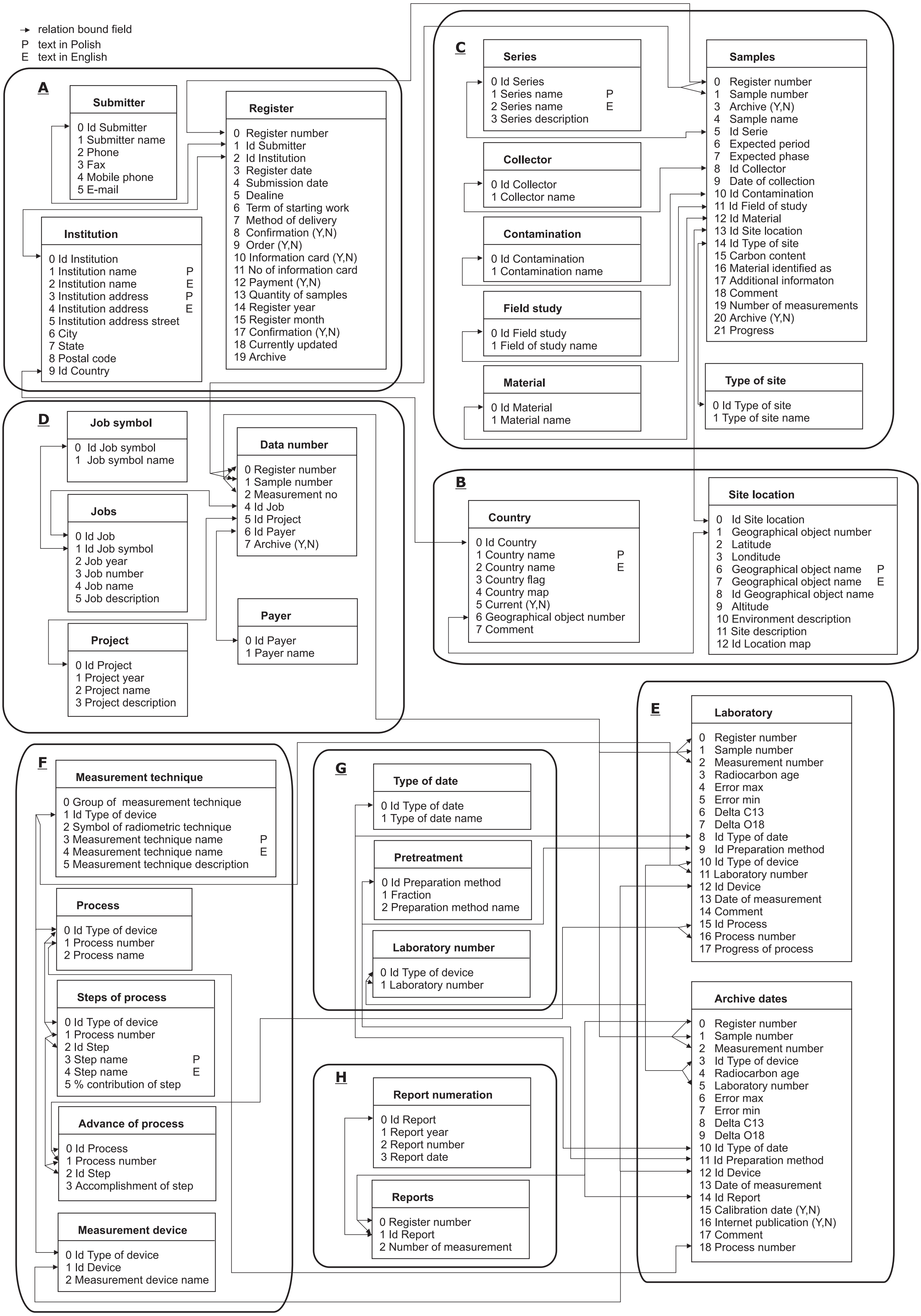

***Fig. 2.** Structure of tables and relations within RoS system*

to use similar option for future references to literature, where the dating result is published.

RoSE application enables also simple searching for a sample by name or laboratory number. More advanced analysis can be made by using *Data Analyse* application, which is briefly described hereafter.

Auxiliary applications

RoS system is supplied with *Data Analyse* and *Web Assistant* applications. The first programme allows analysing the database content by searching with use of advanced criteria, concerning almost all fields included in database and also the ranges of radiocarbon age. The results of analyse are exported to a file of *CSV* (coma separated values) or *HTML* format, which can be edited in other applications enabling further analyses of obtained data. The user defines fields included in exported file and their order, as well as the way of sorting of results. *Data Analyse* application can only process archival samples

The *Web Assistant* programme generates web pages containing information about samples, which are marked for Internet in *Internet* on *Archive Dates* sheet. The samples are sorted by site location and the application creates *HTML* files with sample location lists in alphabetical order. The number of these files depends on the user who specifies the number of sites per one page. The *Web Assistant* programme creates also the page listing the samples from a specified site and a file containing an information sheet for each sample. All this files are cross-linked enabling fast off-line access to data and easy transfer to the web page of the Gliwice Radiocarbon Laboratory (www.carbon14.pl).

4. EXAMPLE OF APPLICATION

Information collected in RoS database system allowed to carry out statistical analyse of group of gyttja (organic sediments) samples dated in the Gliwice Radiocarbon Laboratory. Deposition of organic material in some regions is closely related to the intensity of plant vegetation. This could also indicate on the correlation of number of dated samples with climate conditions in a given period of time. Analysed samples were collected from sites located at the area of Poland. The searching with use of programme *Data Analyse* returned 161 samples, and results were put into *CSV* file consisting of following columns: *Material of Sample*, *Register Number*, *Sample Name*, *Age* (age with its uncertainty). The dates carried out on carbonate fractions were excluded from the set (13 samples), as well as samples with ages exceeding 14200 BP (6 samples), what resulted in acquisition of 142 dates for further analysis

Afterwards, the histogram of conventional dates for gyttja samples was created (**Fig. 3**). Few groups of dates occurring as the maximums and minimums in radiocarbon dates distribution depending on sample age can be distinguished on **Fig. 3**. During time interval 0-13,500 BP (Alleröd and Holocene) a two not very clear maximums can be noticed. This situation could indicate on climate conditions favourable for gyttja formation in above periods. Unfortunately, the groups are not enough explicit to state when these periods of times were started and finished without analysing shape of histograms and having any clear mathematical criteria. But possibility of obtaining this kind of information proves that analysing of radiocarbon dates as a group of physical data can be used to isolate periods of different climate conditions (Goździk and Pazdur, 1987).

Conventional radiocarbon time scale is not equal to scale of calendar years. Therefore, in order to compare palaeoclimatic and palaeoenvironmental data, the calibration of the conventional radiocarbon age to calendar age is required. Predominant number of analysed dates has covered the last 20,000 BP (120 of 125 dates). They were analysed with use of OxCal programme, what enabled obtainment of total probability distribution of calendar ages (Bronk Ramsey 1995, Bronk Ramsey 2001). Therefore an attempt of statistical interpretation of calendar

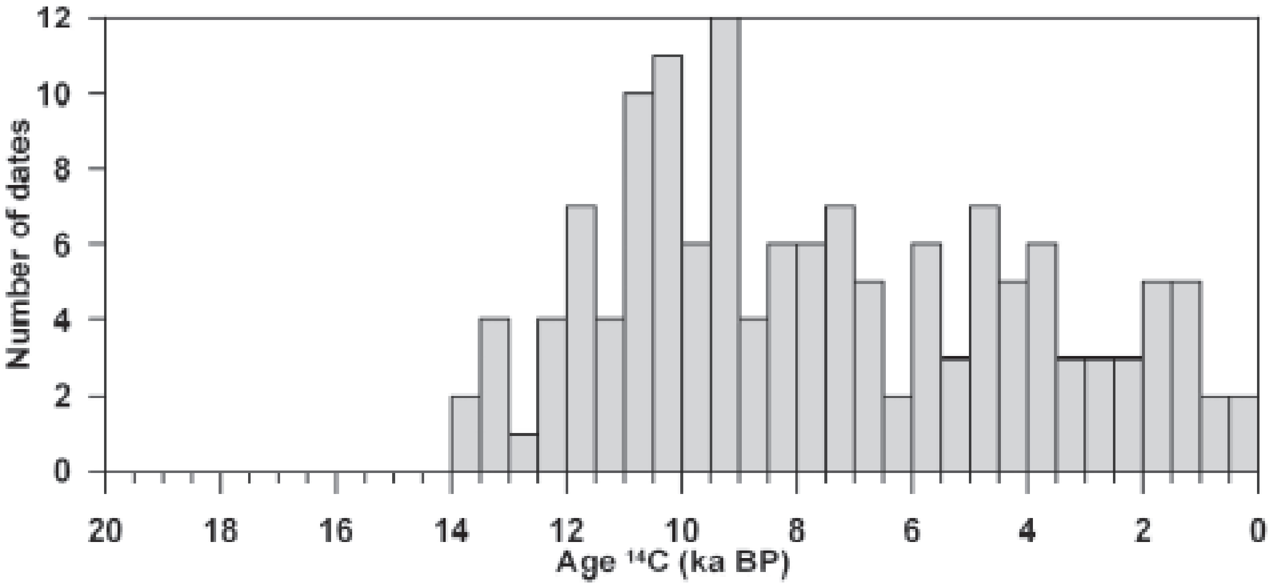

Fig. 3. *Histogram of conventional radiocarbon dates per 500-year intervals for analysed gyttja samples*

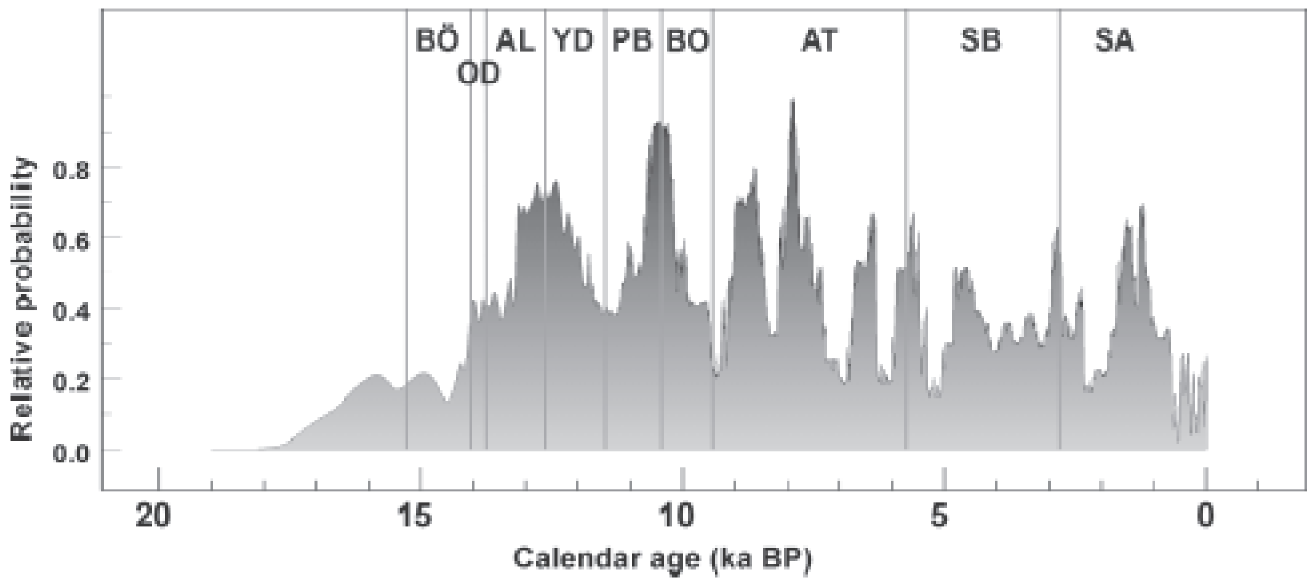

Fig. 4. Sum of probability distributions for 142 gyttja samples

dates was undertaken in frame of MSc thesis by Szczepanek (2002). Summing of probability distribution for each calendar ages (summing continuous functions) was performed and the total probability distribution was obtained, where individual maximums are better visible and separated, but also much more other maximums than in conventional dates distribution can be distinguished. The total distribution of probability of calendar age, which was carried out for the group of gyttja samples is presented in **Fig. 4**. The advanced analysis of shape of probability density distribution for peat samples is presented by Michczyńska and Pazdur (2004). In case of gyttja samples, the analysis seems to be more complicated, as the shape of distribution shows the characteristic features also presented by Michczyńska and Pazdur (2004), but they do not directly correlate with climatic zones of Late Glacial and Holocene, also marked in **Fig. 4**.

5. FUTURE DEVELOPMENT OF RoS SYSTEM

The main problem, which makes the possible analysis difficult, arises from large amount of samples with incomplete description. The current work with database system is mainly aimed at complementing the information stored at present only in paper sample description forms filled by submitters and in a variety of publications.

The RoS system has been created in a way allowing for future enlargement of its possibilities. In particular, the authors wish to develop a new calibration programme, which would be directly connected with the RoS system (*IBCAL.GDB* database). The *IBWORLD.GDB* database also needs much more work to become an archive of precisely described site locations containing also graphical information.

6. SUMMARY

The new database system was developed in the Gliwice Radiocarbon Laboratory and launched in 2002. This version of radiocarbon database, called RoS system, is based on five relational databases. The access to these databases is provided by and management of data is facilitated with using the application called *Register of Samples Editor* (RoSE).

RoS system is supplied with *Data Analyse* application enabling the analyses of database content by searching and sorting using advanced criteria. *Web Assistant* programme is a generator of web pages containing information about samples for publication in Internet.

Information collected in RoS database system allowed to carry out statistical analyse of group of gyttja samples dated in the Gliwice Radiocarbon Laboratory. Summing of probability distribution for calendar ages was performed and resulted in total probability distribution, where individual peaks can be clearly distinguished.

ACKNOWLEDGEMENTS

The development of RoS system was realised as a project financed by the State Committee for Scientific Research "Database of radiocarbon dating for environmental studies", research project No. 6 P04E 007 17, coordinated by Anna Pazdur. The authors acknowledge the participation of staff of the Gliwice Radiocarbon Laboratory in developing previous and present version of the database. We also thank Mr. Wojciech Pazdur for supplying the RoS system with graphics.

REFERENCES:

Ashmore, P. J., Cook G. T., Harkness D. D., 2000: A radiocarbon database for Scottish archaeological samples. *Radiocarbon* 42(1): 41-48.

Bronk Ramsey C., 1995: Radiocarbon Calibration and Analysis of Stratigraphy: The OxCal Program. *Radiocarbon* 37(2) 425-430.

Bronk Ramsey C., 2001: Development of the Radiocarbon Program OxCal. *Radiocarbon* 43 (2A) 355-363.

Becker-Heidmann, P., 1996: Requirements for an international radiocarbon soils database. *Radiocarbon* 38(2): 177-180.

Becker-Heidmann, P., Scharpenseel, H. W., Wiechmann, H., 1996: Hamburg radiocarbon thin layer soils database. *Radiocarbon* 38(2): 295-345.

Burton, M. and Levy, T. E., 2001: The Chalcolithic radiocarbon record and its use in southern Levantine archaeology. *Radiocarbon* 43(3): 1223-1246.

Englesman F. M. R., Taayke E. and Mook W. G., 1986: Groningen ^{14}C database. *Radiocarbon* 28(2A): 788–796.

Goździk J. and Pazdur M. F., 1987: Frequency distribution of ^{14}C dates from Poland in the time interval 12-45 kyr BP and its palaeogeographical implications. *Zeszyty Naukowe Politechniki Śląskiej, Seria Matematyka-Fizyka, z.56, Geochronometria* 4: 27-42.

Kra R. S., 1989a: The International Radiocarbon Data Base: a progress report. *Radiocarbon* 31(3): 1067–1076.

Kra R. S., 1989b: Report of the International Radiocarbon Database Workshop. *Radiocarbon* 31(3): 1080–1082.

McNeely, R., 1996: Geological survey of Canada soil database. *Radiocarbon* 38(2): 271-275.

Michczyńska D.J. and Pazdur A., 2004: Shape analysis of cumulative probability density function of radiocarbon dates set in the study of climate change in the Late Glacial and Holocene. *Radiocarbon* 46(2): 733-744.

Michczyński A. and Pazur M. F., 1989: Local microcomputer database for radiocarbon dates. *Zeszyty Naukowe Politechniki Śląskiej, Seria Matematyka-Fizyka, z.61, Geochronometria 6*: 27-35.

Michczyński A. and Pazur M. F., 1994: Gliwice radiocarbon data bank – present status. *Zeszyty Naukowe Politechniki Śląskiej, Seria Matematyka-Fizyka, z.71, Geochronometria* 10: 47-59.

Michczyński, A., Krzanowski, A., Pazdur, M. F. and Ziółkowski M. S., 1995: A computer-based database for radiocarbon dates of central Andean archaeology. *Radiocarbon* 37(2): 337-343.

Moffett, J. C. and Webb, R. E. 1983: Database-Management Systems, Radiocarbon and Archaeology. *Radiocarbon* 25(2): 667-668.

Obelič B., 1989: The Radiocarbon data base at Ruder Boškovič Institute Radiocarbon Laboratory. *Radiocarbon* 31(3): 1057-1062.

Omoto, K., 1989: The Radiocarbon Data-Base of Japan. *Radiocarbon* 31(3): 1063-1066.

Orlova, L. A., Kuzmin, Y. V. and Zolnikov, I. D., 1998: C-14 database and geographic information system for Western Siberia. *Radiocarbon* 40(1): 313-329.

Pazdur M. F. and Porwoł M., 1987: Microcomputer database system for radiocarbon dates: a pilot project. *Zeszyty Naukowe Politechniki Śląskiej, Seria: Matematyka - Fizyka, z.56, Geochronometria* 4: 19-25.

Scharpenseel, H. W., Pietig F., Schiffman, H. and Becker-Heidmann, P., 1996: „Radiocarbon dating of soils: Database contribution by Bonn and Hamburg." *Radiocarbon* 38(2): 277-293.

Suckow, A. and Dumke, I., 2001: A database system for geochemical, isotope hydrological, and geochronological laboratories. Radiocarbon 43(2A): 325–337.

Szczepanek, M., 2002: Baza danych dat radiowęglowych jako źródło informacji o zmianach środowiska w przeszłości (Database of radiocarbon dates as source of information about environmental changes in the past), MSc thesis, Silesian University of Technology, Faculty of Mathematics and Physics (in Polish).

van der Plicht, J., 1992: The New Groningen C-14 Data-Base. *Radiocarbon* 34(3): 493-499.

Walker A. J. and Kra R. S., 1988: Report on the International Radiocarbon Data Base (IRDB) Workshop. Archaeology and ^{14}C Conference, Groningen, the Netherlands. *Radiocarbon* 30(2): 255–258.

Wilcock, J. D., Otlet, R. L., Walker, A. J., Charlesworth, S. A. and Drodge J., 1986: Establishment of a Working Database for the International Exchange of C-14 Data Using Universal Transfer Formats. *Radiocarbon* 28(2A): 781-787.

Zaitseva, G. I., Dergachev, V. A., Timofeev, V. I. and Sementsov, A. A., 1998: C-14 chronology of archaeological sites in European Russia and changes in environmental processes: A database investigation. *Radiocarbon* 40(2): 759-766.