ZESZYTY NAUKOWE POLITECHNIKI ŚLĄSKIEJ 1987
Seria: MATEMATYKA-FIZYKA z. 56 GEOCHRONOMETRIA Nr 4 Nr kol. 912

Mieczysław F. PAZDUR

Radiocarbon Laboratory
Silesian Technical University, Gliwice

Marian PORWOŁ
Polish Academy of Sciences, Gliwice

MICROCOMPUTER DATABASE SYSTEM FOR RADIOCARBON DATES: A PILOT PROJECT

Summary. The article presents short description of basic parameters of relational database for radiocarbon dates on IBM PC compatible microcomputer. Database consists of six relations indexed according to laboratory number of radiocarbon dates. Database entries are classified according to the study field, problem or process dated, type of dated material, and location of dated site.

1. INTRODUCTION

Since the advent of computer-assisted systems of information storage and retrieval an increasing interest is observed in adaptation of existing commercial systems of data handling to various fields of natural sciences, in particular, to storage and handling of large and steadily increasing sets of ^{14}C dates. Global production of ^{14}C dates, estimated actually to more than 10^4 dates per year, was a stimulating agent for a number of individual research projects, undertaken in several Radiocarbon Laboratories (Gulliksen, 1983; Otlet, Walker, 1983; Moffett, Webb, 1983a,b; Moffett et al, 1985; Johnson, 1985; Mook et al, 1985; Wilcock et al, 1985). The systems already developed and described are based on different database packages, mostly commercially available standard packages. However, most of described ^{14}C database systems are dedicated to archaeological samples.

Total number of samples analyzed for ^{14}C concentration in Gliwice Radiocarbon Laboratory actually exceeds 3,000, with yearly output equal to about 300. Taking into account significant number of ^{14}C dates which have been obtained in recent years in collaboration with radiocarbon laboratories in Groningen, Hannover, Lyon, Louvain, Berlin, Tallin and Tartu, and perhaps some other, which may be estimated to at least 500, and an increasing activity of Radiocarbon Laboratories in Łódź (Trzeciak, 1986) and Kraków (Grabczak, Kuc, 1986), total number of ^{14}C determina

tions in environmental natural samples in Poland may be estimated to about 4,000, with accumulation rate close to 400 per year. Such great data sets need appropriate computer-assisted systems for storage and retrieval of ^{14}C dates and associated information concerning dated samples, sites, profiles, research projects, etc. This necessity has been pointed out during the I Conference "Methods of Absolute Chronology", coorganized in Gliwice in 1983 by the Committee of Quaternary Research of Polish Academy of Sciences and Gliwice Radiocarbon Laboratory (Pazdur, 1986).

With the beginning of the new Central Research Project CPBP 03.13 in 1986 it was possible to start research on establishing of an appropriate database system, designed specially to ^{14}C dates connected with all different fields of Quaternary studies in Poland. The aim of this article is to present preliminary information concerning the hardware and software environments, logical structure of the database, and general characteristics of various possibilities of this system.

2. HARDWARE AND SOFTWARE ENVIRONMENTS

Hardware environment for the database system is provided by 16 bit microcomputer EMIX 86 XT, compatible with IBM PC of international standard. It consists of central unit with processor 8088 and coprocessor 8087 with memory 512 KB RAM, two floppy disk drives 5,25" (360 KB each) and additional hard disk memory WINCHESTER 20 MB. This hardware was chosen because of several reasons: 1) it is minimum but sufficient hardware environment enabling storage and handling of appropriate amount of information (20 MB memory is equivalent to approximately 2000 printed pages); 2) it is controlled by operating systems PC DOS and/or CP/M of international standard; 3) it is easy available in Poland at relatively low prices (ca 11 mln zł in 1986).

Table 1

General characterization of database structure

Relation	No. of fields	Content
BAS	21	all fundamental data
REG	20	regional info
LAB	18	relevant lab info
ADM	13	managing info
SITE	17	relevant field data
REF	19	references

Software environment for the database system is provided by commercially available package dBASE III, working under operating system PC DOS programming. It consists of relational database with his own specific language enabling fast data sorting and retrieval. The commands of dBASE III may be compiled using dBASE III COMPILER. The package enables simultaneous working with 15 files (including maximum ten data files). Database may contain up to 10^9 records, consisting of 128 fields of different type (with total length 4000 characters, and, moreover, may include also the MEMO type fields for storage of large texts of up to 5000 characters). Final reports are prepared in form of text files and may be either edited under appropriate text editor (such as, for example, WORDSTAR or LETTRIX) or stored as separate new files.

Table 2

Structure of relation BAS

	Field Name	Type	Length	Description
1	LABCODE	C	3	lab. code (eg Gd or Hv)
2	LABNO	N	5	lab. number
3	SNAME	C	25	sample name (site name)
4	SNAMEEXT	C	20	sample name extension
5	DTYPE	C	2	type of date
6	SAGEBP	N	5	sample age (conv BP)
7	SAGEDT1	N	4	age error (+)
8	SAGEDT2	N	4	age error (-)
9	DC13	N	6(2)	$\delta^{13}C$
10	STYPE	C	12	type of sample material
11	STYPEEXT1	C	20	material (extension 1)
12	STYPEEXT2	C	25	material (extension 2)
13	SFRACT	C	10	fraction dated
14	SOCCUR	C	25	occurence of sample
15	SCONTM	C	30	contaminants
16	SAGEEXP	C	25	expected age
17	SCOMMENT	C	50	comments
18	PROBL	C	25	form, process or sediment
19	PROBLEXT	C	25	and its extension
20	STUDFLD	C	25	study field
21	DATELIST	C	8	date list

The system may be characterized as user-oriented menu-drived interactive system which may be operated by users not familiar with computer programming languages. Finally, the system may cowork with the high-level

integrated package FRAMEWORK.

Database system is open, i.e. both its structure and content may be actualized at any time, and enables systematical appending of new records. It enables retrieval of any set of data selected from all relations and some or all records according to certain specified criteria. Protection of database files is actually performed simply by copying of selected parts of files on diskette.

3. STRUCTURE

Database structure is represented in form of six relations, shown in Table 1. First relation BAS containing all fundamental data on dated sample and dating results is shown in Table 2. An example of single record in BAS relation is given in Table 3.

Table 3

Example of single record from relation BAS

Field	Contents	Remarks
Record No.	114	
LABCODE	Gd	
LABNO	2318	
SNAME	Bukowno	
SNAMEEXT	2	
DTYPE	fs	a
SAGEBP	6070	
SAGEDT1	130	
SAGEDT2	0	b
DC13	-25.00	c
STYPE	charcoal	
STYPEEXT1		d
STYPEEXT2		d
SFRACT		e
SOCCUR	soil level	
SCONTM	rootlets	
SAGEEXP	Late Holocene	
SCOMMENT		f
PROBL	eolian	
PROBLEXT	dune	
STUDFLD	paleogeography	
DATELIST		g

Remarks: a)date is "Finite" and error is "Symmetrical" (fs);
b) 0 is introduced instead of repeating the value from previous field;
c) -25.00 denotes assumed value of $\delta^{13}C$;
d) more detailed information concerning sample type is not available;
e) no fractionation of sample material, bulk sample dated;
f) no comments available;
g) date not included in published date list

4. INDEXING AND CLASSIFICATION OF DATABASE ENTRIES

All relations in database are indexed by using two index fields which univocally determine all other information related to a given sample (or date). Theses fields are: laboratory code (LABCODE) and laboratory number (LABNO) of each record, for example, for sample listed in Table 3 (Gd-0000) the two indexes are Gd and 0000. The indexes appear as first two fields in each relation and provide identifiers for all operations performed on records from database files.

Logical structure of database is designed to maximize the possibility of data retrieval and handling according to certain specified criteria which are expected to be of primary importance for Quaternary studies. This was achieved by setting all the most important key words in the BAS relation. These keywords refer to:

1. study field
2. problem/process dated (with possible extension)
3. sample material (with possible two extensions).

All entries are classified with respect to the study field into one of 3 groups: paleogeography, paleoecology, and geology. First two groups seem to be self-explained; the last one contains ^{14}C dates of samples dated mainly because of their stratigraphic significance, or samples which cannot be classified strictly because of lack of suitable information or because of other reasons. Next keyword refers to type of studied process or form. Paleogeographical samples are subdivided into 7 groups related to main types of geomorphologic forms, distinguished by Klimaszewski (1981): FLUVIAL, EOLIAN, GLACIAL, PERIGLACIAL, SLOPE, MARINE and KARST. Each group may be further subdivided; for example keyword EOLIAN may have one of the following extensions: DUNE, COVER, LOESS, DEFLATION.

Third important keyword refers to sample material. Following basic types of sample material are distinguished as keywords: WOOD, CHARCOAL, PEAT, MOSS, GYTTJA, DY, SAPROPEL, SOIL, HUMUS, PEATY MUD, BONE, SHELL, ANTLER, SPELEOTHEM, TUFA, CONCRETION, DETRITUS. Each of these keywords may have one or two extensions, depending on information provided by submitter of dated sample. The most suitable extensions for WOOD, CHARCOAL, BONE, SHELL, ANTLER are provided by corresponding Linnean names (if identified) or other more general terms (for example, BONE - extension may read as ANIMAL or HUMAN, for SHELL extension may be given according to environmental criterion, in generalized form as MARINE, FRESHWATER, LAND. Extensions for keyword SPELEOTHEM are: STALAGMITE, STALACTITE, STALAGNATE, FLOWSTONE, CAVE PEARLS; for tufa: SINTER, MOSS TRAVERTINE, ONCOIDS, STROMATOLITE, CALCAREOUS MUD; for PEAT: REED, SEDGE, FEN, WOODEN, SEDGE-MOSS, BIRCH-WOOD, or HYPHNUM, CLADIUM, SPHAGNUM, ERIOPHORA, CALLUNA, MOLINIA, CLADOPHORA.

Regional classification is obtained in four different ways, according to:

1. geographical coordinates
2. basins of main rivers and then of their tributaries
3. general subdivision of the territory of Poland into physiographic units
4. administrative subdivision.

For paleoecological samples regional classification follows subdivision of Poland into paleoecological regions and subregions, given by Ralska-Jasiewiczowa (1982).

ACKNOWLEDGEMENTS

This study is supported through grant from the Central Research Project CPBP 03.13. Many valuable comments to this project were made by Dr B. Nowaczyk and Dr T. Szczypek, as well as some other colleagues, to whom the authors are gratefully indebted.

REFERENCES

Grabczak J., Kuc T., 1986, Datowanie C-14 metodą scyntylacyjną: Zesz. Nauk. Politechniki Śląskiej, Ser. Mat.-Fiz., Geochronometria Nr 2, 35-41

Gulliksen S., 1983, Radiocarbon database: a pilot project: Radiocarbon, v. 25, 661-666

Johnson I., 1985, MINARK microcomputer database system: [In:] The 12th Internatl Radiocarbon Conf., Abstracts (S. Gulliksen and R. Nydal, eds.), Tapir, Trondheim, p. 40.

Klimaszewski M., 1981, Geomorfologia: PWN, Warszawa

Moffett J. C., Orton C. R., Webb R. E., 1985, Application of a radiocarbon database to the study of culture-historic problems: [In:] The 12th Internatl Radiocarbon Conf., Abstracts (S. Gulliksen and R. Nydal, eds.), Tapir, Trondheim, p. 56

Moffett J. C., Webb R. E., 1983a, Database management systems, radiocarbon and archaeology: Radiocarbon, v. 25, p. 667-668

Moffett J. C., Webb R. E., 1983b, Database management of radiocarbon dates: [In:] Proc. 22nd Internatl Symposium on Archaeometry and archaeological prospection, Bradford, England (A. Aspinall and S. E. Warren, eds.). Univ. of Bradford Press, p. 67-72

Mook W. G., Engelsman F. M. R., Taayke E., 1985, The Groningen radiocarbon data bank: [In:] The 12th Internatl Radiocarbon Conf., Abstracts (S. Gulliksen and R. Nydal, eds.), Tapir, Trondheim, p. 41

Otlet R. L., Walker A. J., 1983, The computer writing of radiocarbon reports and further developments in the storage and retrieval of archaeological data: PACT, v. 8, p. 91-105

Pazdur M. F., 1986, Współczesny stan i kierunki rozwoju chronometrii radiowęglowej: Zesz. Nauk. Politechniki Śląskiej, Ser. Mat.-Fiz., Geochronometria Nr 2, 15-27

Ralska-Jasiewiczowa M., 1982, Introductory remarks: Acta Palaeobotanica, v. 22, p. 3-7

Trzeciak P., 1986, Oznaczanie chronologii bezwzględnej metodą radiowęglową stosowaną w Radiochemicznej Pracowni Muzeum Archeologicznego i Etnograficznego w Łodzi: Zesz. Nauk. Politechniki Śląskiej, Ser. Mat.-Fiz., Geochronometria Nr 2, 45-54

Wilcock J. D., Otlet R. L., Walker A. J., Drodge J., 1985, Establishment of a working database for international exchange of carbon-14 using the Harwell Status package and universal transfer formats: [In:] The 12th Internatl Radiocarbon Conf., Abstracts (S. Gulliksen and R. Nydal, eds.), Tapir, Trondheim, p. 55

Wpłynęło do Redakcji 15 stycznia 1987 r

MIKROKOMPUTEROWA BAZA DANYCH DLA DATOWAŃ RADIOWĘGLOWYCH: ZAŁOZENIA WSTĘPNE SYSTEMU

Streszczenie

W artykule przedstawiono krótki opis podstawowych parametrów relacyjnej bazy danych dla datowań radiowęglowych utworzonej na mikrokomputerze kompatybilnym z IBM PC. Baza danych składa się z sześciu relacji o wspólnym indeksie określającym numer laboratoryjny datowanej próbki. Poszczególne rekordy w bazie danych są sklasyfikowane na podstawie dziedziny badań, datowanej formy geomorfologicznej lub procesu geologicznego, rodzaju datowanego materiału, oraz położenia datowanego stanowiska.

МИКРОКОМПЬЮТЕРНАЯ СИСТЕМА БАНКА ДАННЫХ ДЛЯ РАДИОУГЛЕРОДНЫХ ДАТЕРОВОК: ПРЕДВАРИТЕЛЬНЫЕ ПОЛОЖЕНИЯ

Резюме

В докладе представлено краткий перечень основных параметров реляционной системы хоронения информации о результатах радиоуглеродных датировок с использобанием микрокомпютера IBM PC. Система включает шесть реляций имеющих совместный указател которым является лабораторный номер датированного образца. Классификацию данных проводят по области исследований, по датированным геоморфологическим форман или геологическим процессам, по типе датированного органического вещества и по местоположению.