NOTES AND COMMENTS

REPORT ON THE INTERNATIONAL RADIOCARBON DATA BASE (IRDB) WORKSHOP, ARCHAEOLOGY AND ^{14}C CONFERENCE, GRONINGEN, THE NETHERLANDS

A J WALKER and RENEE KRA

The Internationa Radiocarbon Data Base (IRDB) Workshop was held at the Archaeology and ^{14}C Conference, Groningen, The Netherlands, on September 11, 1987.

The participants were:

Renee Kra, Chair	USA
W G Mook, Organizer	The Netherlands
F M R Engelsman	The Netherlands
E Taayke	The Netherlands
J Lakling	The Netherlands
A Van der Wijk	The Netherlands
J Van der Plicht	The Netherlands
Jacques Evin	France
Gordon Cook	United Kingdom
E M Scott	United Kingdom
R L Otlet	United Kingdom
A J Walker	United Kingdom
Roy Switsur	United Kingdom
J C Vogel	South Africa (RSA)
D P Agrawal	India
Steinar Gulliksen	Norway
Rainer Beger	USA
M F Pazdur	Poland

GOALS

The objects of our discussion were:

1) to consider a suitable system for the data base, in particular, the value of a catalogue-type system;

2) to re-evaluate, in the light of discussions in (1), the proposed 29-field format that resulted from preliminary discussions on the data base;

3) to agree on a minimum data entry format.

DISCUSSION

Renee Kra, in the chair, reviewed the situation to date. The need for some form of international data base had been widely recognized, but specific details on system and format for the data still had to be agreed. For the type of data base, two systems were possible:

1) a catalogue-type system which would merely point to where the full information on a date was held (eg, laboratory files, laboratory data base, site publication);

255

2) a full data base system containing all information on each date and acting as an adjunct to or possibly, in due course, a replacement for the journal, *Radiocarbon*.

Advantages and disadvantages of the two systems were discussed with opinion being divided between those who felt that the catalogue-type system was incomplete and would lead to less misuse of data and those who felt that the power of the mainframe computer data base was its ability to store vast amounts of data without restrictions and that this should be exploited. The deciding factor proved, however, to be cost. Storage of data on a large mainframe computer is expensive and calculations suggested that the amount of storage required for the international data base would thus be prohibitive. The first type of data base, the catalogue, was therefore, accepted and it was suggested that Ms Kra should also consider the possibility of keeping such a data base on a PC rather than the mainframe since considerable storage space is available now on hard disks. The PC had the further advantage of being portable and not, therefore, relying on the continued support of any particular institution for the storage of the data, *ie*, for the very existence of the data base.

After this, discussion moved to data format and in the first place, the original, expanded entry format proposed by Renee Kra. This consisted of 29 fields as follows:

FILE NAME: Lab code + no. of file . year (eg, HAR1.87)

FIELDS	CHARACTERS
1. Lab code (eg, HAR)	10
2. Lab no. (eg, 2314)	10
3. Country	20
4. Region (eg, state, province, district)	20
5. Town (expanded character field)	25
6. Site name	20
7. Latitude	15
8. Longitude	15
9. Discipline	10
10. Site type	25
11. Series	20
12. Sample material (expanded field for identification; use semi-colon (;) as delimiter)	25
13. Sample code (expanded field; use ; as delimiter)	25
14. Measurement— BP age	10
15. Other measured value (eg, $\delta^{13}C$)	10
16. Other measured value	10
17. Other measured value	10
18. Culture/period	20
19. Context	
20. Submitter comment	
21. Laboratory comment	Text
22. Calibration comment	
23. References (separate file, eg, HAR1.87R)	

24. Collector (last name; first name; date)   25
25. Collector institution (address; telephone)   25
26. Submitter (last name; first name; date)   25
27. Submitter institution (address; telephone)   25
28. Sample publication (eg, date list)   25
29. Free field (for miscellaneous information)   25

The primary advantage of this format lay in its comprehensiveness and unlimited data storage capabilities but its chief drawbacks were 1) an unwieldy structure and potentially ensuing lack of cooperation on the part of data suppliers and 2) the prohibitive cost of mainframe computer storage mentioned above.

For the catalogue-type system, a much shorter format, with fewer fields, was required and after much discussion, the following format, consisting of 11 fields, was accepted by the meeting:

1. Lab code and number (eg, HAR-1829)
2. Sample name
3. ^{14}C determination with error (expressed in whatever form, eg, age BP, δ^{14}C per mil, etc)
4. δ^{13}C value
5. Sample material
6. Country
7. Geographic coordinates
8. Site name
9. Discipline (eg, archaeology, geology)
10. Association (eg, culture or period)
11. Reference (eg, publications, manuscripts, laboratory)

The main assets of this revised format would be 1) ease of data entry and retrieval, 2) less misuse of data by virtue of its brevity and ensuing necessity for research into primary sources (eg, laboratories), and 3) costs would be kept at a minimum especially if data could be stored on a PC.

Discussion continued on the precise definition of some of the categories, which seemed problematic and, in particular, on the subject of minimum entry. The question arose—could there be gaps in the 11 fields or must each one be complete in order to make an entry on the data base?

It was agreed that it was not necessary for every field to be filled but that where no entry existed, the word "none" should be used (eg, for δ^{13}C where it has not been measured). Clearly, there must be certain categories which must be defined to make it possible for an entry to be accepted but although this was discussed, no final decision was taken.

<div align="center">CONCLUSIONS</div>

The participants agreed that The International Radiocarbon Data Base (IRDB) should be established as a catalogue-type system to serve as an index for ^{14}C determinations. It should provide basic information only, *ie*,

the 11 categories contained in the accepted format, above, but it should also supply full references to more detailed information. Renee Kra will pursue avenues for general adoption of this format, for organization and management of the project, and for funding the setup and continuing costs of the Data Base.