INTERNATIONAL GEOLOGICAL CORRELATION PROGRAM

Circular 80-5b
IGCP-163-IGBA
29 December 1980

Report of IGCP Project 163 for 1980

Contents

Introduction

This circular is an updated and somewhat abbreviated version of the report submitted to the IGCP board early in November.The project has been very active, and to keep this account within reasonable limits only major current activities of obvious interest to the Board and its Scientific Committee are noted under most of the main headings listed above. For reasons discussed in the next section the question of the life of the project is now a matter of prime concern. Some form of continuing international organization of the work is essential. IGCP recognition has been and continues to be of enormous assistance in this respect.

National Groups

New national groups have been formally established in Bulgaria, the United Kingdom and France. A Japanese group will soon be activated. During the report year official project correspondents were appointed by Bulgaria (R. Ivanov), France (P. Grandclaude), Israel (M. Eyal), and the United Kingdom (M. Le Bas).

Like the existing groups, the new ones are designed to maintain and improve communication between individual scientists -- currently contributors but ultimately, it is hoped, users -- and the central office. If there were no international project there would be no occasion for these

national groups. To facilitate completion of the retrospective literature scans they have undertaken for it, however, a number of national groups have proposed programs that extend their activities well beyond the currently approved termination date of the international project. The latest example is the Turkish group, for which a five year plan has recently been approved. (The situation is not without its lighter aspects; our Bulgarian delegate urges us to apply promptly for inclusion in the official program of the next meeting of the International Mineralogical Association, which will convene in Sofia some months after we have ceased to exist. We shall nevertheless make application.)

Central Office Activities

The most difficult step in the development of an electronic base from non-electronic data is the movement of data from the last hard copy to the first machine readable form. It became apparent in 1979 that the project coding form, designed primarily for the convenience of the contributor and so not of the conventional card-image type, was not suitable copy for commercial or sub-professional key punching. Backlog accumulated at an alarming rate and to meet the emergency a conversational program that accepts free form input in response to frequent prompts was developed and tested during the last part of 1979. In the winter and spring of 1980, with funds made available by the US-IGCP Committee, student assistants were trained in the operation of the program and used in a successful attempt to eliminate the backlog of over 2000 completed coding forms.

Unfortunately, however, the long delay in processing this material had led to a considerable loss of momentum; after an intial spurt in response to our Toronto and San Diego meetings and attendant publicity, receipt of completed forms tapered off to such an extent that in the first two months of 1980 none at all were received. Interest began to revive with the first distribution of galleys for proofing, however, and this fall more than 600 completed forms have been received. We must now try to maintain reasonably prompt processing of data contributions.

This may be difficult in the immediate future because of an ongoing review and appraisal of project file structure and processing, growing out of discussion at the recent Madrid meeting (see Circular 80-5a). Opinions have been solicited from computation centers associated with national groups in England, France and Bulgaria, and from well qualified external reviewers in France, Japan, Norway and the United States. We anticipate that this review will be concluded early in 1981.

Our plan is for national or regional centers and well situated individual contributors gradually to take over the initial data transfer operation, so that the central office receives -- on tape or perhaps by network -- material that is ready for machine editing. To facilitate development along these lines, operating versions of the data transfer program are now available at three national centers and tape copies will be made available to any individual contributor on request. Further,

through the good offices of T. Wright and L. Hoover, the project now has access to the United States Geological Survey's computer network; it is hoped that this will soon materially facilitate the interchange of data files, at least within the US group and possibly also for contributors situated near overseas nodes of the network. For the present, however, responsibility of the central office for transfer of data from coding sheets to machine readable form must continue.

Summary of Current Holdings

Data contributed through June of 1980 are now maintained in electronic card-image files in four installations; the central office, the western office of the US group in Cheney, Washington, and the offices of the Spanish and Turkish groups, in Madrid and Ankara. These files are not replicates, but much of the Turkish and Spanish material is also included in the central office files. The Western US office reports that its electronic holdings now include 1000 specimen descriptions, all of rocks collected by the Deep Sea Drilling Program. Major blocks of information on file in the central office include data for:

(a) lavas of the Canary Islands (1350 specimen descriptions), submitted by J.Brandle of the Spanish group.

(b) igneous rocks of Turkey (434 specimen descriptions), submitted by C.Unan of the Turkish group.

(c) lavas of Etna (399 specimen descriptions), submitted by R.Cristofolini of the Italian group.

(d) pre-Cambrian igneous rocks of northern Brazil (116 specimen descriptions), submitted by J.O.Santos.

(e) rocks of the Idaho batholith (45 specimen descriptions), submitted by W.Greenwood of the US group.

(f) lavas of Fiji and the adjacent ocean floor (136 specimen descriptions), submitted by J.Gill of the US group.

(g) rocks of the McMurdo volcanic complex, Antarctica (76 specimen descriptions), submitted by P.Kyle of the US group.

(h) systematic scan of major US publications (1218 specimen descriptions), contributed by D.Barker, K.Bladh, R.Boutilier, J.C.Butler, J.R.Butler, F.Chayes, J.Creasey, D.Fiesinger, L.Fukui, M.Gander, W.Greenwood, W.Kleck, B.Marshal, W.Romey, D.Rubel and D.Sundeen of the US group.

All but items (a) and (b) of this list were transferred from coding sheets to machine readable form in the central office. To facilitate proofing and machine editing, most of this material is being retained in

short disc files identified by contributors' names. When editorial work is completed these and the central office copies of the Turkish and Canaries material will be pooled into a single file of which tape copies will be made available, on request, to those who have assisted in compiling it.

No firm inventory of coded data awaiting transfer to machine readable form is possible because the work of preparing the forms is widely distributed among volunteers who are under no obligation to report regularly; only when forms reach the central office can they be inventoried. About 250 DSDP specimen sheets await processing in the Cheney office of the US group. As already noted, over 600 specimen sheets have been received at the central office this fall; most of these are part of the US group's systematic scan of major US literature. The Indian group reported in May that it expected to complete its scan of the Indian literature in January of 1981. In a recent letter Dr. Saha reports that by late spring the group plans to begin transfer of information from approximately 2500 coding forms to machine readable form.

Machine translation of large numbers of specimen descriptions from local to project format is contemplated by the French and Bulgarian groups, and may also be undertaken by the nascent Japanese group. Considerable effort may be involved in building translators for each such operation, so it may be desirable to defer work of this type until the current review of project format has been completed.

Meetings

The project met on 15-17 September in Madrid; circular 80-5a is an account of that meeting. The chairman was an invited speaker at the biennial CODATA Conference in Kyoto on 8-11 October, and took advantage of the opportunity to review certain rather general aspects of the work of the project. The 1981 meeting of the project is tentatively scheduled for Hawaii, and in 1982 we hope to meet at the International Mineralogical Association Congress in Sofia.

Publications, in preparation or in press

Preliminary findings of a numerical study, by F.Mutschler and W.Greenwood, of the specificity (unspecificity?) of petrographic description, requested by a resolution at the San Diego meeting, were presented at Madrid by Mutschler. The report was enthusiastically received and the authors were urged to prepare the material for publication in some widely circulated geological journal.

The Kyoto address of the chairman, "Attitudes toward data in the hard and medium hard sciences", will appear in the Proceedings volume of the CODATA conference, to be published by Pergamon Press in 1981.

An experimental study of the relative costs of storing derived variables in a base and computing them at operation time, stimulated by discussion at the Madrid meeting, is being prepared for release by F.Chayes. It has not yet been decided whether publication will be in the form of a journal article or a project circular.

INTERNATIONAL GEOLOGICAL CORRELATION PROGRAM

September 30, 1980

Circular 80-5a
IGCP-163-IGBA

Report of the 1980 Meeting of IGCP Project 163

IGCP Project 163 met in Madrid on 15-17 September in quarters graciously provided by the Department of Petrology and Geochemistry and the Computation Facility of the Universidad Complutense. The following attended:

J. Brandle, Spain

F. Chayes, U. S. A.

J. Gill, U. S. A.

Ph. Grandclaude, France

M. Horder, UK

R. Ivanov, Bulgaria

G. Kakatsakis, U. S. A.

M. LeBas, UK

F. Mutschler, U. S. A.

Th. Pantazis, Cyprus

C. Unan, Turkey

From reports delivered at the meeting it appears that nearly 5000 specimen descriptions have now been brought to card-image file status in the project format. Most of these must still be proofed by contributors, and for ease of access all are at present being held in separate files indexed by contributor's names. A detailed analysis of current holdings will be included in our annual report. Discussion followed the outline set forth in the previously distributed agenda. This notice lists only the major decisions taken and actions agreed upon.

(1) Discussions must be inaugurated with national and international agencies or institutions that might undertake to complete, house and manage the new base. It does not seem likely that the necessary arrangements can be effected within the currently authorized life of the project; the chairman was several times requested from the floor to bring this matter to the attention of the IGCP board and, at the appropriate time, to request renewal or extension of the project.

(2) The coding form now in use is somewhat unconventional in that it does not provide column-by-column structure of the coded information, and so probably cannot be used efficiently as copy by subprofessional labor. Since the project currently has neither the volume of work that would require the use of commercial transfer facilities nor the funding that would permit it, the problem is not immediate. But we should be prepared, for if the workload builds up sufficiently we may have to go commercial. F. Mutschler agreed to undertake prompt development of a more conventionally structured coding form. M. Horder and M. Le Bas will collaborate on a form suitable for use by either IGBA or GEXEC.

(3) The grammar used in moving data from coding forms to initial machine readable form has been developed without benefit of professional guidance or consultation. Fears were expressed that it was unnecessarily complex and that the complexities might lead to inefficient passage from "card image" format to packed base. This view was far from unaminous but it was agreed that external review was in order. Such review(s) will be obtained.

(4) At the present time, modal analyses are included only as "additional information" that can be retrieved and listed but not used internally for sorting and computation. This stratagem was adopted because such analyses are still rather rare and the language in which they are reported is non-standard, perhaps even unstandardizable. There was much feeling that modal data should be recorded in such fashion as to permit their use in reduction and sorting but it soon became apparent that the necessary standardization of symbols could not be developed on site. R. Ivanov agreed to prepare a note outlining practices used by the current Bulgarian geological information system for this purpose. His note will be issued as a project circular, and used to stimulate discussion of the problem.

(5) In extended discussion about proposed modifications of the literal scaling factors currently attached to trace element amounts it was decided to abandon them outright in favor of a numerical scale preceded by a '-' sign, thus eliminating one more external symbol table now required by contributors and data-transfer operators. The new format will be like the so called "scientific notation" normally used in computation except that the exponentiation symbol is a '-' instead of an 'E' and the exponent must be a digit in the range 0-9 inclusive. Trace element data in current holdings will be edited to conform to the new format, and appropriate modification of the Contributors' Instruction Manual (IGBA Circular 79-2) will be prepared.

(6) Current procedures for framing and tagging specific kinds of information in the "Additional Information" Block of the coding form were objected to by one participant on the ground that their use in search procedures may be rather costly. The current procedure will be continued pending demonstration of an effective alternative.

(7) There was heated discussion of the advisability of including 'derived' variables--those calculated from the raw data--in the base, and it was clear that thorough tests of the relative costs of computing them on demand vs storing them ab initio must be undertaken. Such a test is now being designed at the central office, but tests should also be undertaken elsewhere. What is at issue is a balancing of the cost of central processor time against the cost of active storage space, and this balance may vary both over time and from installation to installation. The question is perhaps particularly important to us because of the large number of derived variables used in igneous petrology.

(8) Preliminary findings of a numerical study by F. Mutschler and W. Greenwood of the unspecificity of petrographic description, requested by a resolution at the last meeting, were presented by Mutschler. The results to date seemed so striking that the authors were strongly urged to prepare the material for publication as soon as possible.

* * *

The meeting was useful to the project in many ways that cannot be conveyed by a list of this kind. We are all grateful to IGCP for the basic grant that made it possible, to various national IGCP committees that provided auxiliary travel assistance, to the Spanish IGCP committee for its generous support of local activities, to Professor Alfredo Pacheco and the members of the Department of Petrology and Geochemistry for their friendly hospitality during the meeting and in the field excursion that followed, and to our Spanish representative, J. L. Brandle, who handled local arrangements with patience and forbearance above and beyond the call of duty.

Felix Chayes

Felix Chayes, Chairman
IGCP-163-IGBA

FC:mm