

I am speaking here in the name of a research group sponsored by a French government agency called D.G.R.S.T. This group is active in the field of soil science data processing (informatique pédologique).

Two main tasks of this agency is the coordination of various work relating to ecology; and also, giving stimulus to young people to create information procedures in this field. This group embodies all sectors of activity concerned with soil science from university laboratories to private companies. Also, foreign scientists are welcome to join us.

Among one of the more recent activities of our group is the standardization of soil science data. In each case, unanimity has to be reached among the group members on each item standardized. You can realize how difficult this task can be.

Our group has published last year a booklet which is the first of a series on soil science data processing. Its title translated into English is: Glossary of Soil Science, description of the Horizons for Data Processing. This booklet is on exhibit on our Administrative Secretary's desk.

In our method no code names are used when the data are punched. Natural language is used for the description of the soil horizons. This is of vital importance to us because many of our people work under severe climatic conditions. Furthermore, we are planning to generalize the use of small wire recording equipment in the field.

Field data are punched on the IBM 29 or 26 as they reach us in Paris. Data follow each other, separated by a full stop, from column 1 to column 80 of the card. Now we have access to the computing facilities of S.T.A.D. and C.I.R.C.E. which include the UNIVAC 1108, the CONTROL DATA 3600, the IBM 360-75 computers.

Programs for storage, retrieval and treatment of descriptive data are already operational. These programs, called R, D and S, are written in Fortran IV by REGIE INFORMATIQUE for both the 1108 and the 3600, and are also on exhibit. This investigation was supported by O.R.S.T.O.M.

I would like to talk to you about the R program.

R is a program to store a repertory on tape and to modify it when necessary. This repertory is an exhaustive version of the above-mentioned glossary. Several equivalent repertories written in other languages will be in the future stored on the same tape.

In the field of laboratory data, the following is an information system adopted by O.R.S.T.O.M. It remains unchangeable for at least one more year. A form is used to transcribe in format F4.0 the data of an imposed list of variates. All data undergo a $\log_{10}+4$ transgeneration. No missing data are allowed after Boolean selection of horizons and variates. Generalized D^2 measures of distance are computed (Cah. ORSTOM, sér. Pédol., 3, 79-88, 1965). A given clustering process is used (Biom.-Prax., 36-47, 1965). A tentative approach is followed to determine the most potent variates (Cah. ORSTOM, sér. Pédol., 6, 21-34, 1968; Statist. Theory and Method Abstr., 10, 510, 1969). Scattergrams and detransformed confidence limits for the means is also a standard output. All programs are written in Fortran IV and are executed rapidly on our CDC 3600 and UNIVAC 1108 computers. No attempt has been made to transpose them on other computers.

In the same field, many valuable methods are at our disposal. They were programmed by S.O.G.R.E.A.H., S.O.G.E.T.H.A. on the IBM 360-65 and by C.E.P.E. on the CII 10010 and the IBM 360-40.

Several proposals will come under discussion at D.G.R.S.T. next Spring. One of these proposals, which concerns the storage of laboratory data, is as follows: Results of all dosages should be stored. Each dosage is followed by the unit of measurement (as a negative power of ten), by a code name for the element or compound, by a code name for the method of dosage, and by a code name for the laboratory. The same profile and horizon identification is needed for the field data and laboratory data.

Now a few words about a second research group which is working on plant sociology data processing (informatique phytosociologique). This group is seeking agreement on a list of types of vegetation that should be valid on a world scale. The I.B.P. classification is to be examined, even in the absence of a French version. On the other hand, more detailed enumerations concerning forest canopies, herbage formations, etc, have been laid down. Most interesting of all, however, is a list of the "alliances" encountered in Western Europe, which is on exhibit.

Key punching of all former field descriptions seems to be the most needed task. Agreement