COMPUTER PROCESSING OF GEOLOGICAL DATA

D.T. Gittins

Department of Museum Studies
Leicester University

1. Introduction

This paper describes two pilot projects into computer information processing of geological data, one concerned with collection management, and the other with regional recording of geological sites. Both projects used the generalised information processing package, INFOL 2, for the handling and processing of the information. Some similar applications in archaeology are considered, namely artifact registration, site recording and collection management. Finally, some ideas on the integration of information systems on a regional and national basis are put forward.

2. INFOL 2

INFOL 2 is a generalised information processing package, written in standard FORTRAN IV, and is designed especially for use by non-programmers. It operates using a comprehensive vocabulary of English instructions through three main phases of operation :

ESTABLISHMENT - Where the user defines the list of Items in the Records to be processed, states any validation criteria, and inputs the data (using numeric tags) to create the computer file.

UPDATE - Where the user retrieves single or groups of records, and can eliminate, add, modify or replace records, or items within records.

INTERROGATION - Where records are retrieved and are then outputted to the user. Retrieval facilities include free - text searching, as well as multiple retrieval criteria. The interactive mode is especially designed for use in this phase of processing. Output can either be automatic - relying on the package to format the output, or controled - where the user defines the output format. In both of these, there are instructions that allow the user to multiple sort, embedd output in text, define page headings and similarly control the output format.

In both of the phases, the user can gave instructions to
embedd output in text, sort on more than one item, count
values in items, derive simple statictics, generate an
exchange file, specify titles and page headings and so on.

Other features of the package
include the facility to attach FORTRAN subroutines at any stage
during the program, and interrogate the file interactively
with either VDU's or teletypes. The file structure can be re-
defined subsequently if the user desires.

3. Collection Management and Documentation Project

This was undertaken
for the Fossil Mollusc Section of the British Museum (Natural
History) and used INFOL 2 on a CDC 7600 at the University of
London Computer Centre. The bulk of the processing was
undertaken from a remote batch terminal at the Bedford College
Computer Unit (Regent's Park, London), using only the most
basic facilities - punch cards, batch input and line printer.
In this case, we were attempting to see what could be achieved
even where facilities, manpower and experience were restricted.

The material
under study consisted of various species of fossils collected
from different locations in Natal, South Africa, and it formed
part of the collections of the Palaeontology Department. Since
no previous work had been done on it, the material was stored
merely in the order in which it had been uncrated in 1971.

After discussion, the
categories of information to be recorded were defined, and
recording began using standard forms.   The full contents of
the computer record are listed opposite.

Two recording forms were
developed: One for data about a single species of fossils, and
the other for recording data that applied to a number of records
(Normally stratigraphic age, sediment or store location).
However, an interesting development occured ; Once familiar
with the tags and method of recording, it was easier for us to
record the data onto simple lined paper, a format that was far
easier to keypunch also.

Contents of the Computer Record :

1. Record Number
2. Country Number
3. Country Name
4. Published Location
5. Field Location
6. Institution Housing the Material

7. Stratigraphic Age : Era
                         System
                         Stage
                         Section
                         Bed

8. Store Location : Room
                     Cabinet
                     Draw

9. Sediment : 12 Numerical Parameters
10. Sediment : Description

11. Taxonomy : Phylum
               Class
               Sub - Class
               Order
               Superfamily
               Family
               Genera
               Species

12. No. of Open Pairs of Valves
13. No. of Closed Pairs of Valves
14. No. of Single Valves or Individuals
15. No. of Bored Valves
16. % of the Fauna
17. Fossilisation Condition
18. Biotic Inferences
19. Comments
20. Catalogue Number
21. Ecomorphy
22. Collection Details
23. Cataloguer

After input of the basic records on punch cards, followed by file updating using the second recording form, the data file was processed by INFOL 2. With the necessary instructions we were able to produce catalogues, indexes and special reports that were aimed to satisfy as far as possibleany demand for information about the collection.

The project generated the following output from the data on file using the processing system and the necessary instructions from the users ;

A. Record Catalogue :

All the information for each record in sequential order of Record Number.

B. Indexes :

Indexes to the catalogue, ordered on different important items of information, namely :

Field Location
Published Location
% of the Fauna
Abundance
Record Number
Taxonomy (Genera & Species)
Stratigraphic Age
Store Location (Room, Cabinet, Draw)

C. Special Output :

Material in a special format or medium, or satisfying particular criteria ;

Catalog - Cards
Microfilm Index of taxonomy
Automatic Counting of Genera
Lists of 'in situ' fossils
List of fossils with bored shells

As you can clearly see, even with only basic computer facilities we were able to achieve a level of documentation of the material far superior to that possible using conventional, manually - oriented techniques. A wider range of information, as well as the material itself, was rendered more readily available to both the curator, and the researcher.

4. Regional Geological Site Recording Project

This is a pilot project into computer information processing of geological site records within the Earth Sciences Section of Leicestershire County Museum. It was initiated towards the end of 1975, and is at present (January 1976) concerned with the transcription of data from the existing manual files.

The aim of the project is to formulate and evaluate the proposed system, to serve as the basis for an analysis and evaluation of a possibly more extensive, integrated county site recording scheme.

INFOL 2 will be implemented on the County Council's Univac 1106 computer, and punching will be undertaken by their staff using a standard input form.

We estimate that up to 700 site records will finally be processed, and while that may not appear a large number, we will be attempting to document the sites in great detail.

The contents of the computer records has been established, and in all there are 48 categories of information, both textual and numeric. A separate list and details of the recording categories is kept, and recording of the data from the existing manual files uses a simple lined input form, which is used for keypunching. Only a few simple rules govern the recording of data, and no detailed documents are required for the purpose.

We aim to explore the capabilities of both batch and interactive processing, batch providing catalogues, indexes and special reports to serve requirements that can be foreseen, and interactive to answer enquiries that could not be planned for. However, we would also like to explore the possibilities of substituting interactive facilities where batch might primarily have been considered.

Output will include a comprehensive main catalogue, with indexes of Institutions, grid reference, location, owners, planning authorities, special scientific sites, educational uses, type localities, threatened sites, potentially valuable locations, and so on. With the facilities that computer processing provides, it should be possible to meet a variety of demands from within the museum, planning departments, conservation groups, educational bodies, researchers, field units, national bodies and the general public.

Contents of the Computer Record :

Identification :

    Computer Record No.
    Museum Record No.
    Institution
    Section
    Recorder
    Nat. Grid Reference
    Latitude and Londitude
    Other Coordinates

Description :

    Type of Site
    Condition
    Dimensions
    Locality
    Educational Use Grade
    Type Locality Details
    Site Description
    History Description

Personnel :

    Management Body
    Owner
    Tenant/Occupier
    Last Official Visits
    Actively Interested People
    Past Users

Status :

    Planning Authorities
    Planning Status
    Conservation Status

Geology Features :

    Stratigraphy
    Petrology
    Structures
    Mineralogy
    Palaeontology
    Relationships
    Geomorphology

General :

    Re-Appraisal Date
    Threats
    Potential
    Access Category
    Restrictions
    Approach Route

Cross - References :

    Maps
    Plans or Charts
    Transparencies
    Negatives or Prints
    Microfilm or Fiche
    Collection information
    Intra - Regional Files
    Inter - Regional Files
    Published Literature

5. Some Applications in Archaeology

The main applications that readily come to mind are in the fields of site excavation, collection management, and regional site recording, particularly artifact registration on excavation.

Excavation Directors should seriously contemplate the organisation of the excavation process towards subsequent computer processing and analysis. That is, artifact registration should be in a form that readily allows the data to be used as input to computer systems. Of course, this is more easily said than done, especially where no computer processing is available at the time of excavation.

However, there are at present a number of development programmes concerned with the automatic registration of artifacts, or the processing of artifact and site data, and once more widely implemented, it would be valuable to evaluate them in a variety of different field situations.

Information management systems would obviously benefit the archaeologist, whether research, collection or field - oriented. The greater availability of information is an obvious advantage. However, the documentation of existing collections is an extremely time - consuming process, and is merely a duplication of the time spent in excavation, where archaeological material is concerned. It would be obviated to a great extent if artifacts were computer - registered on excavation.

There are already existing information systems that handle regional archaeological site information, normally as a response to planning and development pressures. Even so, the examples are not numerous and this is another application that would repay further investigation.

6. Integrated Regional & National Information Systems

I would finally like to put forward the idea of an integration of both regional and national information systems. Is it possible, or desireable for future research and development to be oriented towards a sharing of the information gathered by field units, archaeological excavations, museums, researchers, and national bodies ? Could we come to the point where Regional Information Processing and Analysis Units centralise information collection and dissemination ? How might such a system be developed ? Who should be envolved ? How would they be organised ?

Hopefully, these questions will be considered and explored by workers not only within an archaeological context, but from a broad scientific viewpoint as well.

D.O.E. 1975	Principles of Publication in Rescue Archaeology
Chenhall, R.G. 1975	Museum Cataloguing in the Computer Age American Association for State and Local History
-------------- 1971	The Archaeological Data Bank : A Progress Report COMPUTERS AND THE HUMANITIES 5(3) : 159-169
Clack, P. 1975	The Information Retrieval Syatem of the Northern Archaeological Survey COMPUTER APPLICATIONS IN ARCHAEOLOGY 1975
Cooper, J.A. 1975	A National Scheme for the Documentation of Geological Sites BULLETIN OF THE GEOLOGICAL CURATORS GROUP 5:225-230
Doran, J.E. & Hodson, F.R. 1975	Computer - Based Archaeological Data Banks In : MATHEMATICS AND COMPUTERS IN ARCHAEOLOGY (Editors) PP. 317-330
Jones, M. 1975	The Role of the County Museum Service in Geological Site Recording NEWSLETTER OF THE GEOLOGICAL CURATORS GROUP 5:218-221
Kenworthy, J.B. 1975	The Fife Archaeological Index - A Computer Implementation COMPUTER APPLICATIONS IN ARCHAEOLOGY 1975
Leveque, F.H. 1975	INFOL 2 Reference Manual DEPT. OF COMPUTING AND CONTROL, IMPERIAL COLLEGE, LONDON
Whallon, R. 1972	The Computer In Archaeology - A Critical Survey COMPUTERS AND THE HUMANITIES 7(1):29-45