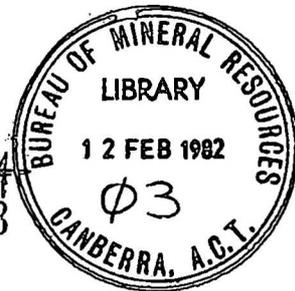


1981/56
ø3

BMR PUBLICATIONS COMPACTUS
(LENDING SECTION)



091788



BUREAU OF MINERAL RESOURCES, GEOLOGY AND GEOPHYSICS

RECORD

Record 1981/56

COMPUTERISED BIBLIOGRAPHIC DATABASES IN BMR

Compiled by K. Modrak

The papers which constitute this Record were presented at a meeting of the On-Line Users Group of the Library Association of Australia (ACT), held on 22 July 1981 at BMR.

The information contained in this report has been obtained by the Bureau of Mineral Resources, Geology and Geophysics as part of the policy of the Australian Government to assist in the exploration and development of mineral resources. It may not be published in any form or used in a company prospectus or statement without the permission in writing of the Director.

Record 1981/56

COMPUTERISED BIBLIOGRAPHIC DATABASES IN BMR

Compiled by K. Modrak

The papers which constitute this Record were presented at a meeting of the On-Line Users Group of the Library Association of Australia (ACT), held on 22 July 1981 at BMR.

Contents

	<u>Page</u>
Glossary of database terminology	i
BMR computer-based information systems - a history (A.J. Barlow)	1
System software for BMR's Hewlett-Packard computer (R.P. Lorenz)	3
Database organisation	3
Database programs	4
LIBLN - the Library database (M. McGeehan)	6
Functions	6
Database schema	7
Keyword set (THKEY)	7
Updating	8
Retrieval of information (Transfer Files)	9
Output	9
The Stratigraphic Index part of GEODX (K. Modrak)	11
Other computerised bibliographic databases in BMR (R.P. Lorenz)	18
Canberra Bibliography (CABIB)	18
Oil Shale Bibliography (OSHAL)	20
Volcano Bibliography (VLBIB)	21
Remote Sensing Bibliography (DIPORS)	22
Palaeontology Bibliography (PLBIB)	24
Palaeontology Database (PALEO)	24
Murray Basin Bibliography (MURRAY)	25

List of Figures

- Figure 1. Database organisation
- Figure 2. Example of program IMUP
- Figure 3. Example of program IMFN
- Figure 4. Diagram of LIBLN schema
- Figure 5. Example of a borrower's loan card
- Figure 6. Example of a coded sheet for a monograph
- Figure 7. Example of a coded sheet for a serial issue
- Figure 8. Example of completed indexing sheets for a reference containing stratigraphic names.
- Figure 9. GEODX database structure diagram
- Figure 10. Examples of retrieval output using files GEOBIB, VLREFS and SNREF
- Figure 11. Example of a Canberra Bibliography coding sheet
- Figure 12. Example of a Volcano Bibliography coding sheet
- Figure 13. Example of a Remote Sensing Bibliography coding sheet

Glossary of database terminology

- Data item - The smallest accessible data element in a database.
- Data item field - An area of fixed length within a data set, which contains individual data items.
- Data item name - A name given to describe the data item value.
- Data set (Data file) - One logical group of data items.
- Hashing - A predetermined series of calculations used by the computer to determine the physical location of items in the database.
- Key item - A data item which may be used to directly locate entries from a data set.
- Line printer - A device used for printing reports away from the terminal.
- Link item - A key item in a master set common to more than one data set, used to link information in separate data sets.
- Master set - A data set of key items through which entries may be accessed directly.
- Random access - The ability to find a data item without having to search serially through all data items in the data set.
- Schema - A description of the user database.
- Transfer file - A computer file containing a set of commands to run a program.

BMR computer-based information systems - a history

A.J. Barlow

BMR first became involved in digital computing in 1959 when a digital data acquisition system was designed and built for recording airborne data in digital form in survey aircraft. During the 1960s the use of digital computers grew and the Geophysical Branch used the Control Data Model 3200 computer at Monash University for geophysical processing - mainly gravity data.

When the Geophysical Branch moved to Canberra in 1965, most work was processed on the CSIRO Control Data Model 3600 computer. In the late 1960s several information storage programs were written by S. Henley and others; some data were also stored on the INFOL database system which was available on the 3600 computer. BMR still has two databases on this system; however, because the system has only a purely serial retrieval mode, it is not very efficient.

The ADP Group was formed in 1974 and one of its first tasks was to rationalise computer storage of information. Investigations of the various database management systems then available extended over a period of about two years, on a part-time basis. Several possibilities were examined. CSIRO had commenced the design of FORDATA which at that time had serious limitations in that much of the data had to be presorted and no real retrieval language had been designed. Database management systems available commercially were in general too expensive to operate, particularly as a service centre computer. SYSTEM 2000 was tried on the Australian Bureau of Statistics computer but was difficult to use and although variable-length fields could be used, the overheads, both in storage space and retrieval time, overrode the advantage of the variable field length. TOTAL appeared to be most suitable for the type of information to be stored by BMR. However, there were no computers in Canberra supporting this system, and line and related charges using a service centre computer in Sydney or elsewhere were too expensive to consider.

In 1976, IMAGE was released by Hewlett-Packard for Hewlett-Packard computers. It is nearly identical to TOTAL in database structure and the software was much less expensive than other commercial systems. At this time too, BMR had an in-house Model 2100 Hewlett-Packard computer installed, primarily for laboratory data acquisition. Although the disc storage space was very limited, it did provide a cheap method of gaining experience in database management systems. It was only a matter of time before larger discs became available.

The first two reasonably large databases were the Library Loans System and the Palaeontological Specimen Index. The latter was simple in concept, as online retrieval was not then required and fast access unnecessary. Some catalogues of specimens have been prepared from the data and more are to be prepared shortly. The library database was the first where more or less continuous online access was required, as well as the production of lists, catalogues, and loan return notices.

There are now 30 active databases in BMR, ranging in size from 100 K characters to 50 M characters. The capacities of the Library database (LIBLN) and the Stratigraphic Index database (GEODX) are each of about 50 M characters, and each is about two-thirds full.

A number of generalised-application programs have been developed to assist in adding, updating, and retrieving information; the IMAGE database system suffers from some limitations and the programs have been written to overcome them. One limitation is in retrieval from multiple data sets, which the program IMFN has been designed to overcome.

In future we expect to increase the size and scope of these databases. As well, BMR has been given a charter to become a national geoscience data source. Minicomputers may not be large enough to handle the possibly greatly increased range of data that BMR may be required to hold, so thought is being given to alternative systems, particularly in the direction of an in-house medium-scale computer.

System software for BMR's Hewlett-Packard computer

R. Lorenz

The software sub-system that enable terminal users to communicate with the system is called TMGR. TMGR controls user language and prevents users from interfering with each other's programs and files.

Logging-in procedures are simple: the user merely presses any key on the terminal and the system responds with a request to log in. Each user has an individual code name that allows the user access to selected programs and databases. Once the user is cleared there are several commands available to instruct the computer on the procedure required. Some common commands are:

RU - runs a program
TR - runs a program via a transfer file
ED - edits a file
ED,- - creates a file
LN - sends the information to the line printer
EX - logs out from the terminal

And when the user wishes to stop a program in midstream:

BR - breaks the program, stopping the routine being performed
(the computer then awaits further instructions)
AB - aborts the program and leaves all files open

Data base organisation

All bibliographic databases within BMR are organised using the Hewlett-Packard IMAGE/QUERY database management system. This system works by grouping data elements into files called data sets. The sets have abbreviated names, e.g. monographs (MONOG), loan records (LOANS) etc., and are linked using the same reference number of item number (Fig. 1). Sets within a database are linked to each other via a master set using a common item field(s), e.g. a reference number field.

If random access is required, i.e. if it is necessary to retrieve a given data element quickly, the system makes special provision for this by designating certain data elements as key items and establishing a master set for them (Fig. 1). The master set contains one of each different value of a key item and all the addresses for the item in the data sets. Entries are located quickly by a 'hashing' technique carried out by the system.

The choice of sets and arrangements of fields is made within the constraints laid down by the Image System, taking into account factors such as economy of space, type of retrievals required, etc.

Database programs

Various programs have been written for use in maintaining and accessing the databases; they fall into three categories:

1. Updating: IMUP, IMAD
2. Search : QUERY, IMFN, IMUP, IMCHK
3. Special-purpose: IMDAL, RUNEM

The two most commonly used programs are IMUP and IMFN.

IMUP operates on only one set at a time for the following: adding new entries, correcting entries, deleting entries, listing entries. It prompts the user for the name of the database, the set name, the data item field, and the type of operation to be carried out (Fig. 2). Updating may be done interactively or in batch mode. Batch mode produces a print-out of the items corrected, added, or deleted. A version of IMUP, called IMAD, is used for adding new entries to several sets at one time in batch mode only. Print-outs of the created files are produced and corrections may be done before these files are run against the database.

IMFN (Fig. 3) is a version of QUERY which allows interactive searching across sets in the database. QUERY itself only allows for the interrogation of one set at a time. With IMFN there is slightly less program prompting than with IMUP, and so the user needs to be reasonably familiar with the system before using the program.

HEWLETT-PACKARD
IMAGE/QUERY
DATA BASE ORGANISATION

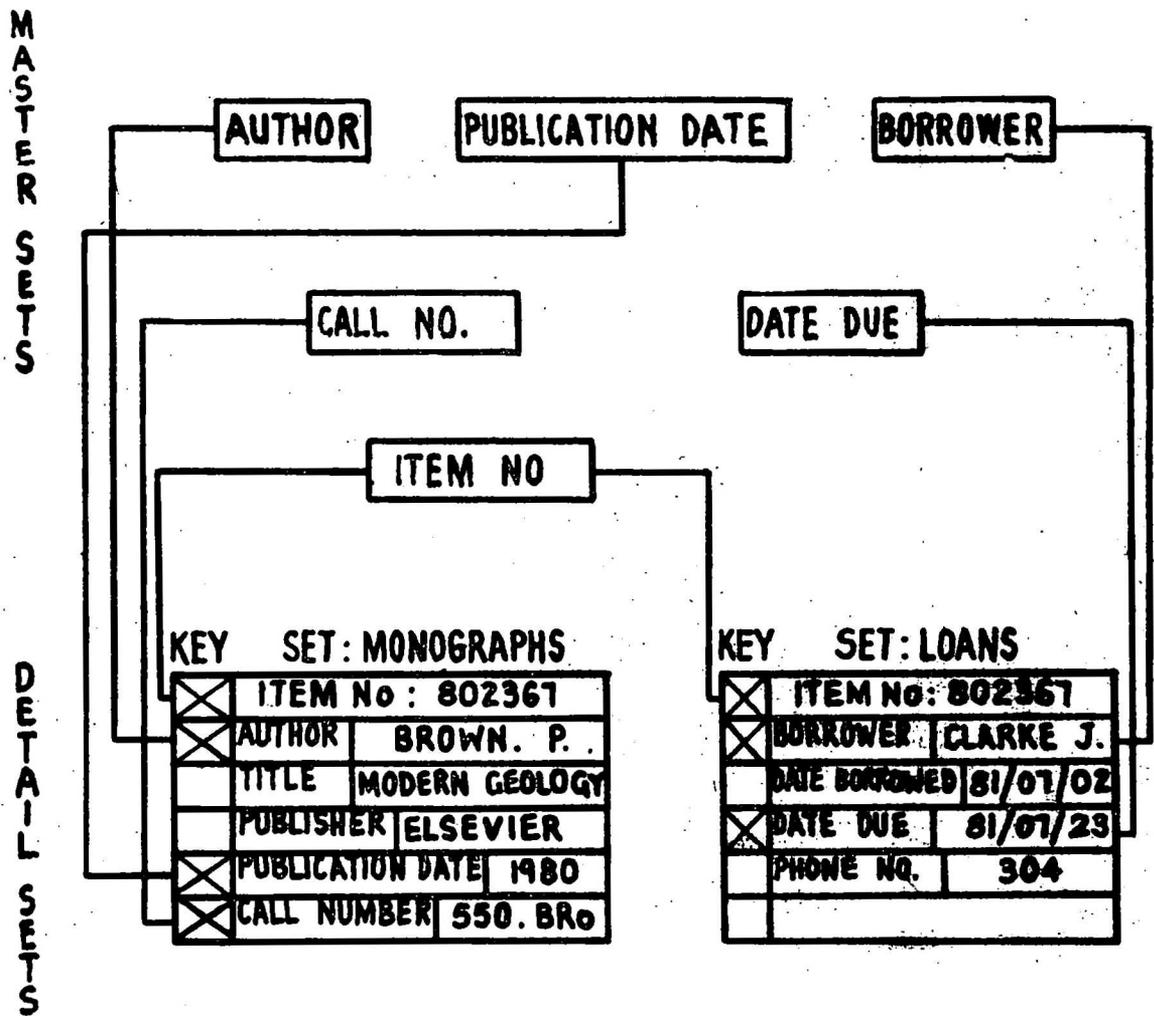


Figure 1. Database organization

EXAMPLE OF PROGRAM IMUP

PLEASE LOG-ON: JACK.ADP

SESSION 20 ON 11:16 AM FRI., 10 JULY, 1981
PREVIOUS TOTAL SESSION TIME: 02 HRS., 30 MIN., 43 SEC.

:TMGR
SECURITY CODE? 8065

!RU,IMUP,1

DATABASE:SECURITY =GEODX:8065

IMUP OPERATIONS

- 1 UPDATE data-entry
- 1 UPDATE data-chain
- 2 DELETE data-entry
- 2 DELETE data-chain
- 3 ADD data-entry
- 4 LIST data-entries
- 5 EXIT IMUP

OPERATION NUMBER (or /?)=1

UPDATE operation

SET NAME (or /?)=SUBJT

SEARCH VARIABLE (or /?)=KEYWD

KEYWD VALUE =Micropaleontology

1 ENTRIES IN DATA-CHAIN

REFN1 :81/21791

KEYWD :Micropaleontology

IS THIS ENTRY TO BE PROCESSED? (YES/NO)=YES

REFN1 :/

KEYWD :Micropaleontology

DATA-ENTRY 10692 UPDATED

1 ENTRIES PROCESSED

KEYWD VALUE =/e

FURTHER PROCESSING? (OPTION or /?) =/E

TRANSACTION FILE J1020A:8065:60 SAVED FOR 7 DAYS

Figure 2. Running the program IMUP. The user's input is underlined.

EXAMPLE OF PROGRAM IMFN

```
RU,IMFN,1,!  
DATA-BASE:SECURITY=?GEODX:8065:  
NEXT ?LINK=MSREF:  
NEXT ?FIND KEYWD IS "Niobium" END;  
00001 ENTITIES QUALIFIED  
NEXT ?REPORT NAME=STREF;  
k1SGEODX BIBLIOGRAPHYKOS PAGE 1
```

```
79/19610 Withnall I.W.  
Mines and mineral deposits of the Georgetown 1:100 000 sheet area,  
Queensland.  
Geological Survey of Queensland. Report 100  
1978
```

```
END OF FILE  
NEXT ?EXIT;  
!EX  
:EX  
FMGR 071  
:EX,SP  
$END FMGR  
FMG20 REMOVED
```

```
SESSION 20 OFF 11:28 AM FRI., 10 JULY, 1981  
CONNECT TIME: 00 HRS., 07 MIN., 20 SEC.  
CPU USAGE: 00 HRS., 00 MIN., 07 SEC., 820 MS.  
CUMULATIVE CONNECT TIME: 02 HRS., 41 MIN., 31 SEC.  
END OF SESSION
```

Figure 3. Running the program IMFN. The user's input is underlined.

There are two basic commands - FIND and REPORT.

FIND outlines the parameters of the search, using operators IS, IGT (is greater than), ILT (is less than), INE (is not equal to), and CO (contains). Logical connectors AND and OR can also be used in the FIND statement. The search value is placed in parenthesis, and the FIND statement is completed with an END statement. All statements within IMFN and QUERY must finish with a semicolon.

REPORT allows for the sorting and formatting of information resulting from the FIND statement. It enables the setting out of bibliographic data and tabulation of numeric data. It is possible to send the information to the terminal, the line printer, or save it as a file for later use. This is done by using certain parameters when logging into the program, or specifying them before the REPORT statement.

The command, EXIT, closes all files and terminates the program.

To save time in frequent similar searches, Transfer Files are set up that enable additional prompts to be given to the user, run the required searches, and automatically output the result. In the following example a transfer file, KEY, is set up. This file in turn uses a file, SEARCH, that prompts the user to name the field to be searched and the item required. All other instructions are taken from the SEARCH file.

```

TR, KEY
RU, IMFN, SEARCH, 1, 1
ITEM NAME = KEYWD
OPERATOR (IS, IGT, ILT, INE) = IS
VALUE = Niobium

```

The result of the search is automatically sent to the terminal.

IMCHK, IMDAL, and RUNEM are less frequently used programs. IMCHK provides an alphabetical list of entries in master sets and their corresponding number of entries in the data set. IMDAL is the special-purpose program used to format the Library overdue notices. RUNEM is text formatting program which will produce reports fully formatted into chapters and sections, with contents list, index, and page numbering.

LIBLN - The BMR Library database

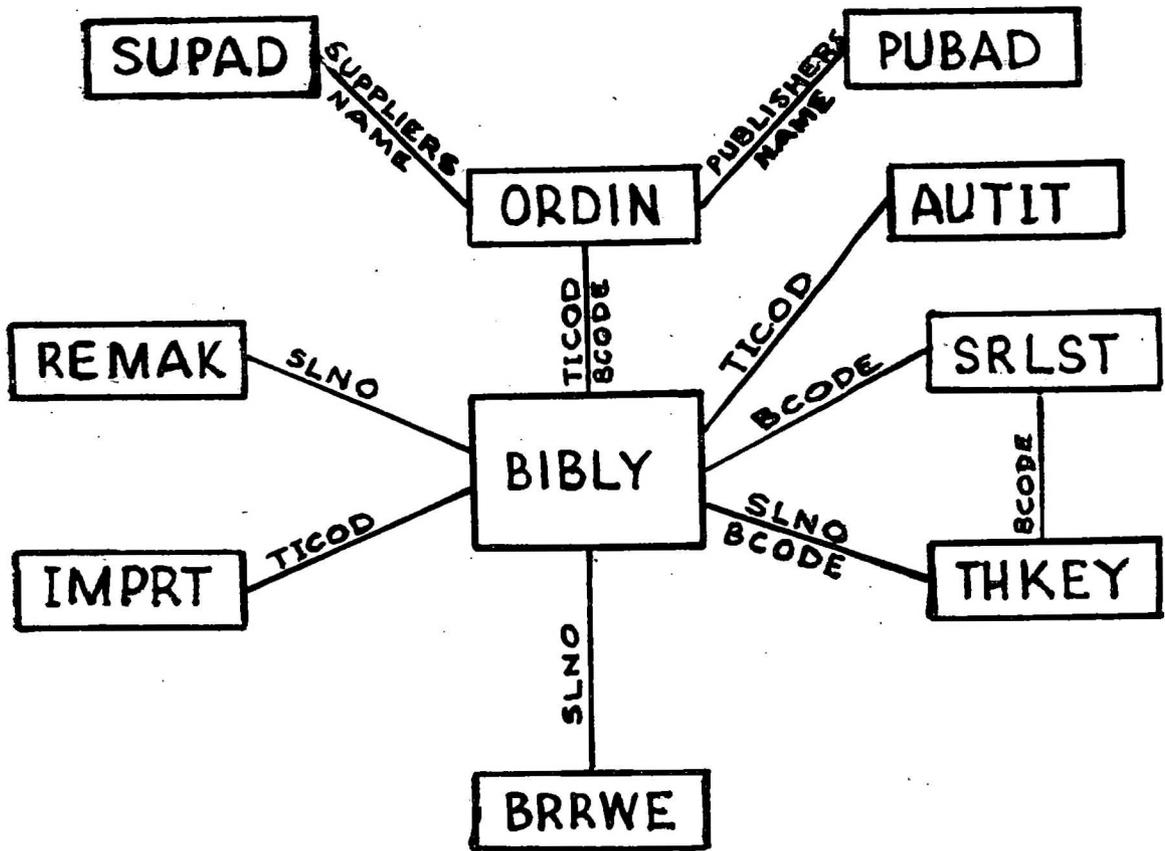
M. McGeehan

The Library Loans computer database, LIBLN, is the largest bibliographic database in BMR in terms of computer disc space occupied. The database originated at the request of the Departmental auditors for better control over library records, and came into operation in October 1978. On 30 June 1981 the database contained 48 066 items in set BIBLY, 7193 loans, 18 330 keywords, and 2644 serial titles.

Functions

LIBLN began primarily to record library loans. To carry out this function, information is recorded from the following sources:

- (a) Accessions of all serial issues received by the library since October 1978. A unique code is used for each title of a serial; this is known as the ACODE or BCODE depending on which data set it appears in.
- (b) Accessions of all monographs received since October 1978. Provisions are made for bibliographical details, call number, and subject headings and at this stage the database does not cater for added entries and multiple authors. It therefore cannot be used to produce a library catalogue. Consequently, the card catalogue is being maintained until funds can be allocated for further development of the database.
- (c) Items for loan received before October 1978. These items are coded and assigned a computer number at the time of the loan.
- (d) Loans from other libraries are entered into the system, thus forming a record of inter-library loans.



SLNO = computer number
 BCODE = code for serial name
 TICOD = title code

Figure 4. Diagram of LIBLN schema.

As the database was developed, other functions were added:

- (a) The ordering system. Provisions are made for publishers' and suppliers' names and addresses.
- (b) Information retrieval using keywords.
- (c) Generating master lists of serials held in the library.
- (d) Recording the loan frequency of items. Fields are available to record this information, but the process is not in operation at this stage.

Database schema

Figure 4 is a diagram of the database schema. The data sets used are as follows:

BIBLY	- Bibliographical details, including call number.
ORDIN	- Ordering details.
SUPAD	- Suppliers' names and addresses.
REMAK	- Remarks, e.g. microfiche.
IMPRT	- Imprint, collation, and accession date.
BRRWE	- Borrower's name, issue date, and loan type.
THKEY	- Keywords.
SRLST	- Serial information, including visible index instructions.
AUTIT	- Author and title.
PUBAD	- Publishers' names and addresses.

Entries into all sets are in upper-case except for the monograph collation data set and 'see' entries (i.e. cross references) in SRLST.

Keyword set (THKEY)

THKEY is the set which is most used and thus probably creates the most interest for this meeting.

Keywords are assigned to the coding sheets at the end of each day, and the sheets are then sent to the data processing operators for input to the database. In practice, the coding sheets are sorted into two categories.

- (1) Monographs and titles in series - these have an author and title field.
- (2) Serials - no data for author and title fields.

Keywords are assigned in the first category using the 'Keyword in Context' (KWIC) method. Newly acquired monographs are excepted, having been assigned subject headings at the time of cataloguing.

The keyword entries are controlled with a 'thesaurus' which has evolved from terms already in the database. The thesaurus is periodically checked to eliminate duplications and errors and to maintain consistency. Where confusion exists with similar terms the Thesaurus published by the Australian Mineral Foundation is consulted. No abbreviations are used.

Emphasis is given to keywords particularly useful for BMR subject areas. These include: rock type, formation, geological age, process term (as in mineral processing), and place name. For Australian place names an entry is also made for the State; for overseas place names an additional entry is given for the country.

Updating

LIBLN is updated daily. The frequency is dictated by the need for borrowers' loans to be recorded promptly. Borrowers cards are maintained; they are filled out and then signed by the borrower, a procedure required by the auditors.

As each item is accessioned or borrowed it is assigned a unique number and is stamped by a numbering machine. The same number is given to the coding sheet and, in the case of a loan, to the borrower's card.

BMR LIBRARY LOAN SLIP																				
Database : LIBLN										Set : BRRWE										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
CSLNO	504406																			
BRWNM	STONE A.																			
ISDAT	8/06/30																			
DCDAT	8/07/21																			
WVTP	3W																			
RES																				
Borrowers signature																				
																				
																				M(P)-244(b)

Figure 5. Example of a borrower's loan card.

DATABASE: LIBLN																																								
Set: AUTIT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
TJCOD	[REDACTED]																																							
AUTHR	[REDACTED]																																							
AUTHS	[REDACTED]																																							
TITLE	[REDACTED]																																							
TITLF	[REDACTED]																																							
TITLG	[REDACTED]																																							
Set: BIBLY	[REDACTED]																																							
ASLNO	0	8	3	3	2	2	[REDACTED]																																	
BCODE	IE75 [REDACTED]																																							
ISSUU	NO. 529 [REDACTED]																																							
TICOD	[REDACTED]																																							
PUBDT	MAY 1981 [REDACTED]																																							
DTRCD	81/06/12 [REDACTED]																																							
CPYNO	[REDACTED]																																							
ACCNO	[REDACTED]																																							
CALNO	5338.45 (94) AUS [REDACTED]																																							
LNFRQ	0 [REDACTED]																																							
Set: REMAK	[REDACTED]																																							
DSLNO	[REDACTED]																																							
REMKS	[REDACTED]																																							
Set: IMPRT	[REDACTED]																																							
PUBLR	[REDACTED]																																							
PLACE	[REDACTED]																																							
COLLN	[REDACTED]																																							
ACCDT	[REDACTED]																																							
TKCOD	[REDACTED]																																							
Set: THKEY	[REDACTED]																																							
GSLNO	[REDACTED]																																							
GCODE	[REDACTED]																																							
KEYPH	[REDACTED]																																							
KEYPH	[REDACTED]																																							
KEYPH	[REDACTED]																																							
KEYPH	[REDACTED]																																							
KEYPH	[REDACTED]																																							
KEYPH	[REDACTED]																																							

Figure 7. Example of a coding sheet for a serial issue.

10

Figures 5, 6 and 7 show a borrower's card and completed coding sheets for a monograph and a serial.

At the end of the day the sheets are passed onto the Data Processing Operators for input into LIBLN. Computer printouts are then obtained for checking and correcting.

Retrieval of information (Transfer Files)

Four standard transfer files have been set up for ease of access to the database for information. They are as follows -

- | | |
|------------|--|
| TR, MONO | - retrieval by one item only e.g. keyword, call number. |
| TR, MONO 1 | - retrieval by an author's name. |
| TR, BRWNM | - lists all items held by a specific borrower. |
| TR, SERIAL | - retrieval by a combination of two items, such as two key words or, as most often used, serial BCODE and and issue. |

Output

The output serves several needs other than the basic information needs:

1. Audit:

A list of accessioned items is printed in number order, for use by the auditors at stocktake time.

2. Reminder notices:

Each fortnight a list of reminder notices is printed for overdue books. On each notice is recorded the borrower's name and a list of books and their dates of issue.

3. Generating a keywords thesaurus:

LIBLN can generate a useful up-to-date thesaurus of keywords.

4. List of serials and their BCODEs

5. Accessions lists:

The generating of regular accession lists is a project being developed in BMR Library.

The Stratigraphic Index part of GEODX

K. Modrak

The geological information database, GEODX, which began operating early in 1979, uses the BMR's in-house Hewlett-Packard 2100 computer and the IMAGE database system.

The system has been shown to be more useful and efficient for stratigraphic nomenclature purposes than the card indexes. Articles indexed today appear on the database, at the latest, one month later compared to the turnaround time of about six months which was common for the card indexes. A high level of accuracy for the information input can be maintained and retrievals can be virtually carried out on any indexing field. It is the latter aspect which has made the Stratigraphic Index input a potential information source for references to Australian geology.

The information in the database has two immediate main sources in BMR:

1. the Information Section,
2. the Stratigraphic Index group.

The Information Section indexes all BMR publications, and unpublished material such as BMR Records. The Stratigraphic Index indexes all published literature on Australian geology received by the BMR Library. It is this second source of information which is the basis of this paper.

Articles relevant to Australian geology, whether they contain stratigraphic names or not, are indexed under the following fields:

- | | |
|-------------------------|--|
| AUTHR: Author(s) | - Surname followed by initials. |
| PYEAR: Publication year | - |
| TITLA,B,C: Title | - There is a limit of 216 characters;
some abbreviation may be necessary. |

- DOCTP: Source journal - If the journal title is unique, and thus can stand on its own, it is recorded in this form. Otherwise the layout follows AESIS, i.e. organisation name followed by the type of publication. A master list is periodically generated for consultation.
- STATE: State(s) - This field records the State(s) to which the article is relevant. If it is of a general nature it is indexed as 'Aust'.
- MAPNO: Map sheet(s)
HUNMP - An attempt is made to record the 1:250 000 Sheet area and the 1:100 000 Sheet to which the article is specifically relevant.
- MDPNM: Detailed location - If possible the detailed geographic location is recorded, allowing retrievals to be very specific.
- BASNP: Basin/structural province(s) - This field contains detailed geological locations.
- KEYWD: Keywords - The article is assigned keywords obtained from the first edition of the Australian Thesaurus of Earth Sciences and Related Terms. In addition specific names of minerals, fossils and techniques which are not included in the Thesaurus can be used.

If stratigraphic names are found in the article the following information is also indexed:

- SNAME: Stratigraphic Name - Spelt exactly as in the article.
- SSTTE: State(s) - The State(s) to which the name is relevant (which is not necessarily the State to which the article is relevant).
- REFNC: Literature occurrence - The main location of the name within the article is recorded, e.g. Table 4, Page 5, Figure 3.
- TYPER: Mode of occurrence - A description of how the name was used, i.e. whether the unit is defined, fully described, described, briefly described, or only mentioned.
- MTHYR: Date - As an aid to compilation of the Variation List, the month and year in which the name was indexed is recorded.
- GLAGE
AGEDC: Geological age - A geological age is only recorded if the stratigraphic name appears on a map or is defined.
- CMMS: Comments - This field is reserved for any additional useful remarks about the usage of the name.

These indexing fields have been followed since the commencement of computerisation. Literature indexed before then is on cards; the data from the cards are being input into the database by contractors as finances allow. Unfortunately, complete continuity of fields of information will not occur because the fields HUNMP, MDPNM, & BASNP did not exist before the index was computerised.

Figure 8 shows the format of an indexed article as abstracted onto a data sheet.

Once a month the data sheets are numbered sequentially and passed to the Data Processing Operators for input into the database. The program IMAD is used for the bibliographic details and program IMUP is used for the stratigraphic name information.

The data base structure diagram (Fig. 9) shows the relationships of the various sets and fields. The more commonly used fields are by now the following sizes:

Keywords	- 1500 keywords have been used.
Authors	- 1700 authors have been recorded.
Source journals	- References have been indexed from about 300 different journal series.
Structural provinces	- 420 have been used.
Detailed geo- graphic locations	- 1600 have been noted.

Once the information has been added to the database it is exhaustively edited, as detailed below:

1. The input is checked against the data sheets.
2. Program IMCHK is used to list unique values of each set.
3. Once corrections have been made from 1 & 2, step 1 is repeated.
4. Stratigraphic names used in each article are listed and checked for duplication in indexing.
5. A list of all the names input is checked against the Central Register to detect new entries to the database. It is at this stage that recourse is made to the literature to check any discrepancies of spelling or style.

Information is available for retrieval as soon as it is added to the database. Generally though, as the editing process takes about two weeks, retrievals at this early stage are carried out only if essential.

Several files have been set up to present the output from a retrieval (Fig. 10):

- (a) GEOBIB - A listing of all information on Bibliographic data sheet No. 1 (Fig. 10, example 1).

BIBLIOGRAPHIC DATA SHEET No.2
FOR STRATIGRAPHIC INDEX

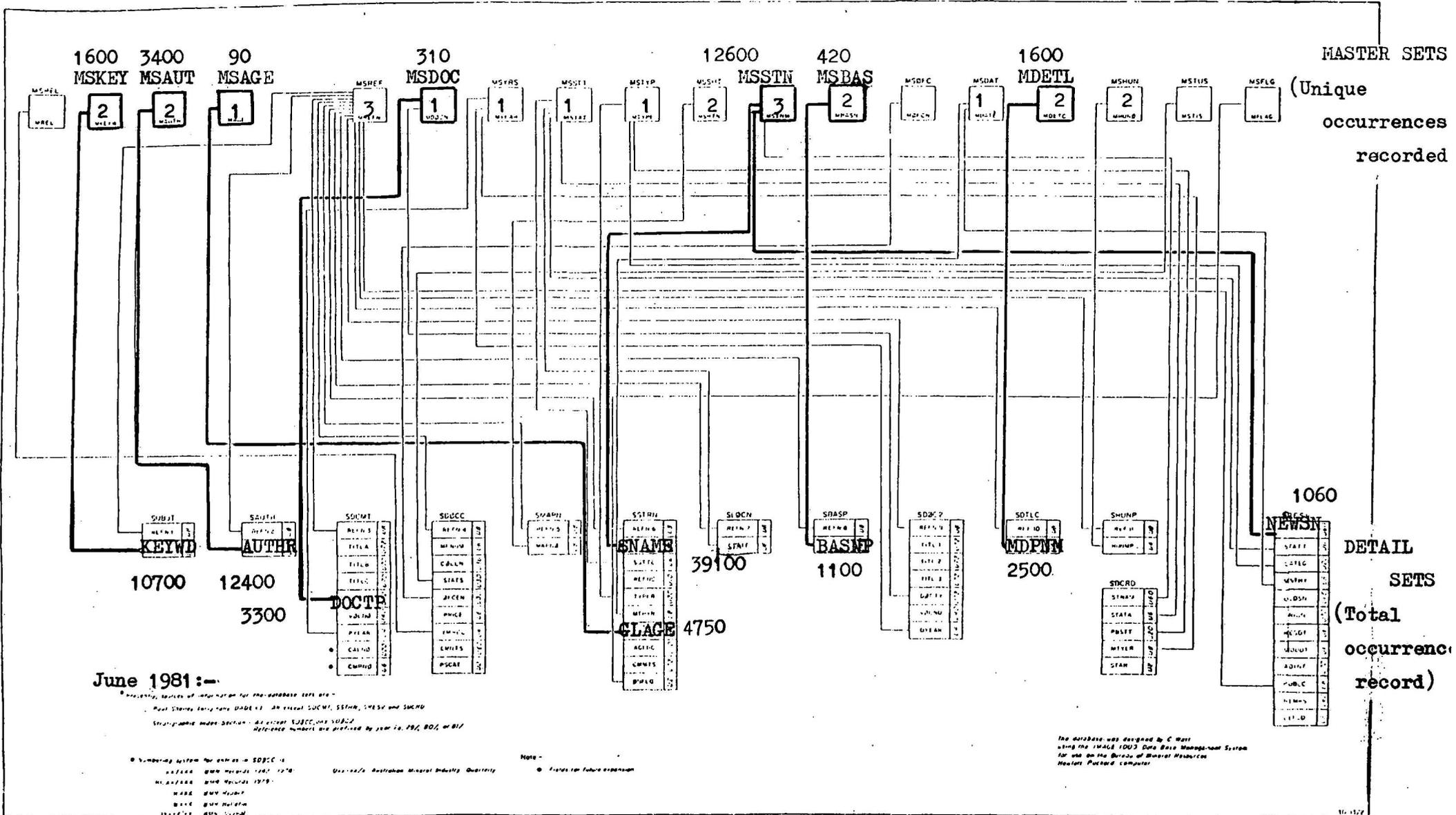
Database : GEODX	Program : IMUP	Date : 10/4/81	Page 2 of 5
Set : SSTRN	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40		
REFN6	81/21571		
SNAME	Candlew Formation		
SSTTE	Qld		
REFNC	P8		
TYPFR	Briefly described		
MTHYR	Apr. 1981		
GLAGE			
AGEDC			
CMMS	See also Figs. 10 & 11.		

REFN6	=		
SNAME	Etheridge Group		
SSTTE	=		
REFNC	P1		
TYPFR	Briefly described		
MTHYR	=		
GLAGE			
AGEDC			
CMMS			

REFN6	=		
SNAME	Robertson River Formation		
SSTTE	=		
REFNC	Fig. 2		
TYPFR	Mention		
MTHYR	=		
GLAGE			
AGEDC			
CMMS	Refers to distribution of gold deposits.		

REFN6	=		
SNAME	Cobbold Dolerite "		
SSTTE	=		
REFNC	P5		
TYPFR	Mention		
MTHYR	=		
GLAGE			
AGEDC			
CMMS			

Figure 9. GEODX database structure diagram.



Where - MSKEY/KEYWD = Keywords MSSTN/SNAME = Stratigraphic name 1 = Checked once a month
 MSAUT/AUTHR = Authors MSBAS/BASNP = Basin/Structural province 2 = Checked every 3 months
 MSDOC/DOCTP = Document type MDETL/MDPNM = Detailed location 3 = Checked every 6 months

1600 = No.occurrences in set.

20
Figure 10. Examples of common types of output.

Example 1 - GEOBIB format.

GEOIDX BIBLIOGRAPHY

81/21571 ,Bain J.H.C.
1981
Some new ideas on the age and origin of the Etheridge Goldfield, and
their exploration implications.
AusIMM. Proceedings 277, P1-10
STATE(S) : *Qld
1:250 000 MAPS : *SE/54-12
1:100 000 MAPS : * 7561 * 7661 * 7560 * 7660 * 7659
DETAILED LOCN. : *Etheridge Goldfield
BASIN / PROVINCE: *Georgetown Inlier
SUBJECT KEYWORDS : *Mineral exploration *Ore genesis *Mineral deposits *Geochronology

Example 2 - VLREF format.

GEOIDX REFERENCES
(Sorted by Reference Numbers)

81/21571 1981 ,Bain J.H.C.
Some new ideas on the age and origin of the Etheridge Goldfield, and
their exploration implications.
AusIMM. Proceedings 277, P1-10

Example 3 - SNREF format.

GEOIDX REFERENCES
(Includes Stratigraphic Names Used)

81/21571 1981 ,Bain J.H.C.
Some new ideas on the age and origin of the Etheridge Goldfield, and
their exploration implications.
AusIMM. Proceedings 277, P1-10
STRATIGRAPHIC NAMES:-
Candlow Formation , Etheridge Group , Robertson River Formation
Cobbold Dolerite" , Forsayth Granite" , Croydon Volcanics
Esmeralda Granite , Yarman Formation , Malacura Sandstone
Langdon River Siltstone , Heliman Formation , Townley Formation
Bernecker Creek Formation , Einasleigh Metamorphics ,

- (b) VLREFS - An abbreviated version of GEOBIB showing the author, title, and source journal (Fig. 10, example 2).
- (c) SNREF - Similar in format to VLREFS, but includes a list of stratigraphic names which appeared in the article (Fig. 10, example 3).
- (d) An assortment of other files set up for specific needs of clients.

The computerised Stratigraphic Index now has five primary types of output -

1. Variation List - A bimonthly list, distributed to State and Territory Geological Surveys, Geology Departments of universities and other tertiary institutions, interested mining and exploration companies, and BMR personnel, containing:
 - (i) Names reserved in each State, including details of variations and deletions.
 - (ii) Reserved names in each State that have recently been published, who reserved the name, name of author, the bibliographic reference and a short comment on the way the stratigraphic name was used.
 - (iii) Other names, i.e. names published without prior reservation, the bibliographic reference, and a short comment on the way the stratigraphic name was used.
 - (iv) Published definitions of units, by name and the reference.

The generation of the Variation List via the computer was carried out in stages, starting with the July-August 1979 List (No. 35) and culminating with the May-June 1981 List (No. 46), which was entirely computer-generated.

2. Deletions List - Proposed stratigraphic names are reserved for a period of five years only. Each year a list of reserved names to be deleted from the Reserved Name Index is circulated and only those names for which retention is requested are retained for a further five years.

Deletions List No. 7, 1980, was the first to be entirely computer-generated.

3. Selective Dissemination of Information - The information cards, sent out over many years to State Surveys, were phased out in December 1980, and have been replaced by bimonthly computer printouts (coinciding with the Variation Lists), which are formulated as closely as possible to the client's needs. The first batch was sent to the Western Australian Geological Survey in February 1981.
4. Current Awareness Service for BMR staff - This new service, started in March 1981, is provided on a bimonthly basis; two types of lists are produced:
- (a) General lists, distributed to Sections within BMR.
 - (b) Specialised lists, provided for project groups.
5. General Enquiries - Information on stratigraphic names, references, reservation of names, etc., is available. Replies to such requests from non-BMR enquirers generally are available within a week. If the information sought is current and held on the database, the turnaround time is a day or so; the longer elapsed time is needed if the card system has to be used.
6. Cumulation Lists derived from the Variation Lists - To promote the orderly use of stratigraphic names some Stratigraphic Nomenclature Subcommittees have published cumulated lists of names relevant to their State or Territory, derived from the Variation Lists (see Records of Geological Survey of NSW 18 (2) pp.201 - 208, and Queensland Government Mining Journal 82 (955) pp.299-230). It is intended that in future these

lists will be generated from GEODX and be made available to all Stratigraphic Nomenclature Subcommittees.

At the end of June 1981, the database contained all the literature indexed since December 1978¹ - a total of 2508 references - plus another 1900 previously indexed². These references contained 12 597 different stratigraphic names resulting in a total of 39 119 occurrences.

Further information on the Stratigraphic Index and its relationship with GEODX can be obtained from a handout available free of charge from the Stratigraphic Indexers, BMR.

¹ i.e. all Australian and foreign literature received in BMR Library from about that date.

² Earlier data held in the index cards is being put into the GEODX database by contract as funds permit. At the end of June 1981, bibliographic references with author names A to Gill, E.D and back to the beginning of 1970 had been transferred or prepared for transfer.

Other Computerised bibliographic databases in BMR

R.P. Lorenz

There are nine databases within BMR that are either bibliographic or have a bibliographic component and were set up using the IMAGE/QUERY system:

LIBLN - Library
CEODX - Australian geoscience literature
CABIB - Canberra bibliography
OSHAL - Dil shale bibliography
VLBIB - Volcano bibliography
DIPORS - Remote sensing bibliography
PLBIB - Palaeontology bibliography
PALEO - Palaeontology database
MURRY - Murray Basin bibliography

LIBLN and GEODX (stratigraphic index component) in separate papers in this Record. The remaining seven are described below.

Canberra Bibliography (CABIB)

BMR is the principal authority concerned with all aspects of the geology of the Canberra area. Enquiries and requests for assistance concerning geological problems are directed to the Geological Branch, and to the Information Section and Library of the Operations Branch. In order to meet some of these requests, R. Lorenz (the Reference Librarian) and R.S. Abell, compiler of the Canberra 1:100 000 geological map sheet, obtained provisional approval to set up the bibliography and a trial database became operational in September 1978. The project was approved in October 1978 and data entries commenced soon afterwards.

Size. By the end of 1979, 89 references had been indexed and entered into the database; a further 72 were added in 1980. The bibliography is expected to total at least 500 references.

Content. All references relating to the geology of the Canberra 1:250 000 Sheet area, SI/55-16, are included, together with annotations. Much of this material is in the form of unpublished BMR Records.

Structure. Features a main author field which enables listings of authors with papers in chronological order. Broad subject categories as well as keywords are used. Provision has also been made for a collation field.

Standards. Data are entered in upper case for some fields, and in upper & lower for others, to produce variety when printed out (Fig. 11)

The library shelf number is included in the main bibliographic reference set. Where an item not in the library is indexed, a ** is placed in that field. Abbreviations are used, where necessary, only in the collation field. Authors initials are used.

Indexing is at four levels

- (1) Broad subject categories. These are chosen from the Australian Mineral Foundation (AMF) publications 'Aesis Quarterly' and 'Earth Sciences and Related Information'.
- (2) Keywords. Specific subject keywords are chosen from the AMF 'Thesaurus of Earth Sciences and Related Terms' 1976. Also included are descriptors such as the names of particular engineering projects and public buildings.
- (3) Localities. These are listed if the work deals substantially with a particular geographic locality, e.g. Tuggeranong, Red Hill, etc.
- (4) Stratigraphic names. Indexed exactly as in the published form, together with the geological age, if stated.

Updating. To date there have been two annual batches of updates. These were checked and then corrected by the Data Processing Operators. Before each issue of the bibliography was run, checks were carried out on the contents of each set, using the IMCHK program.

Format Files. Two basic bibliographic listings can be obtained using IMFN. The files used with this program are:

- (1) FCAB - sorted by author and publication date
- (2) FCAUT - sorted by reference number

Other files have been designed to produce author, locality, keyword, and stratigraphic name indexes. The file, CHECK, gives all the details for each reference, including localities, keywords, and stratigraphic names.

Documentation. The 1979 references in the bibliography have been issued as BMR Record 1980/57 (unpublished Microfiche). The 1980 and 1981 references (i.e. those input in those years) will also be issued as a BMR Record (unpublished Microfiche). Hand-copy printouts can be provided on request.

Oil Shale Bibliography (OSHAL)

The Oil Shale Bibliography was set up in response to a need by the Petroleum Exploration Branch for a comprehensive listing of all references on oil shale, in Australian and overseas. The database OSHAL was set up in late 1979, and entries were made from cards held by D. L. Gibson of the Geological Branch. After a search on GEOREF (the American Geological Institute database), a number of additional references were entered into the database.

Size. At July 1981 there were 197 references in the database: authors 186, keywords 107, localities 64.

Content. At this stage the bibliography is limited to published material, but eventually it is expected to include unpublished material. It will probably be limited to Australian references, to limit its size. The following aspects will be covered: geology, processing, economics, environment.

Standards. Data are entered in a combination of upper-case and upper-&-lower-case (as in Canberra Bibliography). Keywords are not necessarily assigned from a thesaurus, although the AMF Thesaurus is used as a guide. Abbreviations are not used, except for authors initials.

Updating. Batches are entered whenever a new edition of the bibliography is planned. At this stage only that batch is edited.

Format Files. There is only one format, sorted by entry number, which gives only basic bibliographic details. The file used is called OSBIB.

Documentation. There have been two issues of the Oil Shale Bibliography, both in the BMR Library Bibliography series (Nos. 9 and 11). Other issues are planned, as time permits.

Volcano Bibliography (VLBIB)

Compiled by R.W Johnson & W.D. Palfreyman, the Volcano Bibliography began with a collection of pre-1945 reprints obtained from Robin Cooke, an Australian volcanologist who was killed on Karkar Island in 1979, while investigating an eruption of Karkar Volcano. Initially a card index was maintained. Subsequently it became apparent that a considerable amount of other pre-1945 information existed that could be collected and indexed. The aim is to collect this material, index it, and make the bibliography available on request.

It was decided to set up a computer database; this was done in late 1980, and all the information on the card index is now in the database.

Size. The database contains 522 references, as at August 1981.

Content. Material containing references to volcanoes, earthquakes, and tsunamis in Papua and New Guinea prior to 1945, with annotations (including reference to translation, if any).

Type of material indexed is: (a) published books and journals; (b) newspapers; (c) unpublished government documents, personal letters,

manuscripts, diaries, reports, photographs, and theses; (d) maps; (e) transcriptions of personal interviews and radio programs.

Standards. Upper & lower case for all fields (Fig. 12). In the author field the author's initials are used, and any foreign name prefixes (der, von, van, etc.) are placed after initials, e.g. 'Dam, A., van'. A thesaurus has been generated from the entries in the database. The keywords are primarily geographical terms, including broad districts (e.g. Bismarck Archipelago, Bougainville, Gazelle Peninsula), detailed locations, plus the terms volcano, earthquake, tsunami, etc. No rock names are used.

Updating. New references are entered in batches, on an irregular basis, by the Data Processing Operators. The aim is to send a batch at the beginning of the week, add it to the database and have any corrections carried out by the end of the week. After each batch is input, a file is created for that batch, and corrections made from it. Lists of authors, journals, etc., generated via IMCHK are produced, to check for errors.

Format Files. Searching is carried out with program IMFN. Report files vary according to type of sorting: (1) by entry number, VLRPNO; (2) by author, VLRPAU; (3) by publication year, VLYEAR.

All files retrieve full bibliographic details.

Documentation. A circular was sent to interested people around Australia requesting information for the database. As yet no listings have been produced for any outside person. It is planned to produce a BMR Report (microfiche) containing the bibliography.

Remote sensing bibliography (DIPORS)

The Database for Digital Image processing of Remote Sensing Data (DIPORS) is a bibliographic database set up by R.F. Moore to assist in BMR research on digital image processing systems. The bibliography began on cards in 1979 but this became inefficient and a database was set up in December 1980. All data from the cards were entered and subsequent items were coded and entered into the database.

Size. To July 1981, the database contained 708 references.

Content. The bibliography covers English language material on image processing, including: landsat; medical images, e.g. X-rays; military application, e.g. digitisation of air photos; X-ray images of building structures; computer vision; microscopic images, e.g. electron microscope, cytogenetics, etc.

The type of material indexed includes monographs, serials, unpublished theses, technical reports, and manuscripts.

Standards. Upper & lower case is used for entries, except for monograph titles in the publication field, which are in upper case (Fig. 13). In the author field, initials are used instead of full first names. Foreign name prefixes are all in lower case, e.g. 'van Dam, A.'

In the locality set, there are three levels of geographic names: country, State, town or place name.

Keywords are assigned from a specialised thesaurus devised specifically for the database. At least two are assigned to every item - a broad term, followed by a specific term. The thesaurus has a hierarchical component for ease of use.

Updating. New entries are batched into groups of 10-40 and entered. These are checked and edited after each batch. Checks of all entries are made about every two months.

Format Files. Searches are carried out using program IMFN. There are four report files, each providing a different arrangement: (1) RSNOT, not sorted; (2) RSAUTH, sorted by author; (3) RSYEAR, sorted by publication year; (4) RAPUBL, sorted by the publications field.

Documentation. A paper on the database was presented to the 2nd Australasian Remote Sensing Conference, 31 August-4 September at ANU. It is intended to produce a BMR Record on the database. Listings can be produced on request.

Palaeontology Bibliography (PLBIB)

This database was originally set up in 1969 on the CSIRO CYBER computer, using INFOL; and was transferred to the BMR Hewlett Packard in December 1980. PLBIB is based on a card system, maintained by S.K. Skwarko of the Palaeontology Section of Geological Branch; the card system was started in 1960 and contains references on Ordovician graptolites and Mesozoic molluscs. Most of the data on the cards have now been transferred to the database.

Size. As at July, 1980: subjects 4700, taxa 9841, ages 4372, locations 3655, total references 3488.

These figures are total values, not unique item values.

Contents. Family names of fossils described in publications are indexed, including geological age, and geographic location. World-wide literature, in any language, is indexed so long as it is considered as published.

Standards. Only upper case is used. Initials are used for authors given names. Names of serials are abbreviated but no particular standard is used. A system of code numbers is used for family names of fossils, geographic localities, and geological ages.

Updating. Carried out once a year, in batches which are then checked.

Format Files. Only one format file (STANPT) is available; data are arranged numerically by reference number. At present the only key item is the reference number.

Documentation. BMR Record 1969/75 has been issued, giving details of the database on the INFOL system. No further documentation is planned.

Palaeontology Database (PALEO)

This database was set up in late 1978 for the storage and control of information on the Commonwealth Palaeontological Collection (CPC)

which contains about 22 000 specimens. The CPC is a collection of Australasian type fossils and other fossils illustrated in publications; it is held at BMR. No systematic catalogue exists in the collection. The database has concentrated to date on the Order Foraminiferida and now contains information collected since 1960 on foraminifera. Other Orders will be added as time permits.

Size. At July 1981, 7073 specimens and 4593 references had been indexed.

Content. Most of the database is information about specimens in the collection. This includes: class, order, family, genus and species names; country and State of origin; well name; map sheet; location; basin (including grid references); formation; age.

Standards. Upper-&-lower-case is used, except for authors which are in upper case only. Journal titles are abbreviated.

Updating. Weekly batches are entered and checked. All the input and updating is carried out by the Data Processing Operators.

Format Files. One only (PALFM) can be used for output, sorted by reference number.

Documentation. None planned at this stage.

Murray Basin Bibliography (MURRAY)

A card index to Murray Basin references was set up in mid-1979 to assist with the Murray Basin Project*. Later it was decided to enter the data in GEODX, reference numbers being prefixed by MB/. However, this decision was reviewed as some of the data were not compatible with those of GEODX. An identical database was then set up specifically for Murray Basin references.

* A joint project with the Geological Surveys of New South Wales, Victoria, and South Australia, and the Water Resources Commission of New South Wales.

Content. Most of the material is in the form of unpublished reports from companies and State Geological Surveys, (confidential material is excluded). Some of the references go back as far as 1880. The database concentrates on general geology but some references to hydrology, soils, palaeontology, and some fringe topics are included. No data have been entered since November, 1980.

Standards and format files are the same as GEODX. Updates are carried out periodically.

Documentation. Lists were compiled in November 1980 and sent to participating members of the project.