

SIMON FRASER UNIVERSITY COMPUTER PRODUCED
MAP CATALOGUE

Brian PHILLIPS: Head Social Sciences Librarian and Gary ROGERS: Programmer-Analyst, Computer Centre, Simon Fraser University, Burnaby, British Columbia

An IBM 360/50 computer and magnetic tape are used in a new university library to produce a map catalogue by area and up to six subjects for each map. Cataloguing is by non-professional staff using the Library of Congress "G" schedule. Author, title, and publisher are in variable length fields, and codes are seldom used for input or interpretation. Machine searches by area, subjects, author, publisher, scale, projection, date and language can be carried out.

Simon Fraser University in Burnaby, British Columbia, opened in September 1965 to 2,500 students. The Library's book collection was small and the map collection yet to be started. To-day there are 6,000 students, approximately 350,000 volumes and 25,000 sheet maps.

When graduate work was offered in geography the map collection had to be expanded rapidly. Only a small staff was available and it was essential that any map catalogue be largely maintained by trained non-professional assistants. The circulation, acquisitions, and serials systems were automated and there was of course no sacred 3"x5" card file to be replaced. An IBM 1401 (now a 360/50) was in the Library and the University Librarian encouraged experiment. Some form of automated book catalogue was clearly indicated and work began in 1966 to develop one.

Automated or semi-automated methods for cataloguing and producing map lists have been in use for over twenty years. Very little, however, appeared in print on the subject until the 1960's. Since that time there

has been a number of articles on proposed systems and experimental projects, though only a few describe operating systems.

The U.S. Army Map Service Library has used punched cards since 1945 (1). At the time of investigation, this system was not fully automated, making use only of electric accounting machines rather than a computer. Other automated catalogues, such as those for the San Juan project (2) and for McMaster University (3), restricted the amount of information possible by using only one punched card. These systems required codes and tables for both input and interpretation.

The literature revealed other approaches, and several, such as indexing by co-ordinates (4), or using a hierarchical classification, were considered. In the former, each sheet is indexed by its centroid in latitude and longitude. This provides complete control by location; but all requests, and of course indexing, must be expressed by centroid and the extent of the search area indicated in miles. A hierarchical system, such as that suggested by Donahue and Hedges (5) or that used by McMaster University, permits a detailed subdivision of area and/or subject. There is, however, no agreement on standards, with each library developing a classification to meet its own needs.

Visits were made to the University of California at Santa Cruz and Illinois State University at Normal, Illinois, to see two systems that were being automated. Both had used the universally recognized classification of the Library of Congress. The California system was first outlined by Carlos Hagen (6) in a proposal to automate the map library at the Los Angeles Campus, and implemented with some changes at Santa Cruz by Stanley Stevens. William Easton at Illinois State has described his work in cataloguing the collection there (7). The use of codes in both cases meant a number of revisions, as new projections, publishers or other information were required. Because of format, library staff must be called upon to interpret much of the information.

MATERIALS AND METHOD

In the Simon Fraser map catalogue the Library of Congress classification "G" schedule (8) is adopted for computer use. In it each major natural or political unit is assigned a block of four numbers. The schedule starts with the world and hemispheres, then sweeps through North and South America, to Europe, Asia, Africa, and finally Oceania. Adjacent areas are thus grouped together numerically. The classification similarly groups related subjects. A single letter is used for the broad subject and an alphanumeric code for subdivisions.

In an automated system each area name must have a unique number if it is to appear in the printout under that name. To this end it has been necessary to make variations in the Library of Congress "G" schedule. Indo-China (G8010-G8014), for example, must be split to provide separate numbers for Laos, Cambodia and Vietnam. The subject classifi-

cation must also be divided to provide an alphanumeric code for each subject that is grouped under one general number in the schedule.

As commonly in map libraries, the main entry is area rather than author, which is of secondary importance. The author (engraver, cartographer, etc.) is entered on the coding form and appears in the description. In the imprint, publisher is given first, followed by place of publication. These three elements are in variable length fields.

Information from the maps is entered on a coding sheet (Figure 1) by a library assistant. Difficult sheets are entered by a librarian who checks all sheets.

As indicated on the flow chart (Figure 2), the coding sheets are sent to the library keypunching section, where a deck is made for each record. The number of cards for any particular map depends upon the quantity of information required to describe it. The cards are then sent to the Computing Centre, where they are written onto magnetic tape and used to update the current master files.

A preliminary survey determined the average length of a map record to be 350 characters, while the maximum approached the region of 700. In order to maximize the use of tape space, it was decided that four of the fields would be variable in length. These are: 1) main entry (area and title) (215 characters maximum); 2) publisher (129 characters maximum); 3) author (129 characters maximum); and 4) notes (215 characters maximum). Access to these data elements is made possible by storing their character counts in fields preceding the variable portion of the record.

Two master files are kept and updated each time a run is required. These are the area master (by L.C. class number) and the subject master. The area master contains all maps and is used to produce the classified list and the alphabetical list. The subject master contains only those maps which have been assigned an L.C. subject code. If a map has more than one subject it appears on both the list and the tape file as many times as it has subjects.

Changes and deletions are entered into the system along with additions. Status codes signal the three: A - addition, C - change, D - deletion. Change and deletion records are complete decks. The records are changed or deleted by comparing the call number on the area master and the call number and subject code on the subject master. Call number and subject code are the only fields that cannot be thus changed. Their change is accomplished by replacing the old record with the revised record.

As the only unique identifying number for each punched card would be the call number (maximum of 24 spaces), a six-digit I.D. number is assigned. It is repeated for each of the five decks. The maximum number of cards used in any deck is four (main entry and notes), though up to ninety-nine could be used if necessary.

S.F.U. MAP CODING SHEET

I.D.	NUMBER	CARD TYPE	CARD CODEZ	STATUS	AREA	SUB AREA	MAJOR SUBJECT	DATE	SHEET NUMBER	COPY	SCALE	PROJECTION	LANGUAGE	LOCATION	FORM	MINOR SUBJECTS	SIZE	CODE NAMES	
0,0	18,6,8,7	1	0,1	A	19,19,2	M	03	C	8,0	1	9,6,7	S							
I.D. NUMBER										----- MAIN ENTRY and TITLE -----									
0,0	18,6,8,7	2	0,1																
		2	0,2																
		2	0,3																
		2	0,4																
I.D. NUMBER										----- AUTHOR -----									
0,0	18,6,8,7	3	0,1																
		3	0,2																
		3	0,3																
I.D. NUMBER										----- NOTES -----									
0,0	18,6,8,7	5	0,1																
0,0	18,6,8,7	5	0,2																
0,0	18,6,8,7	5	0,3																
0,0	18,6,8,7	5	0,4																

Fig. 1. Map Coding Sheet.

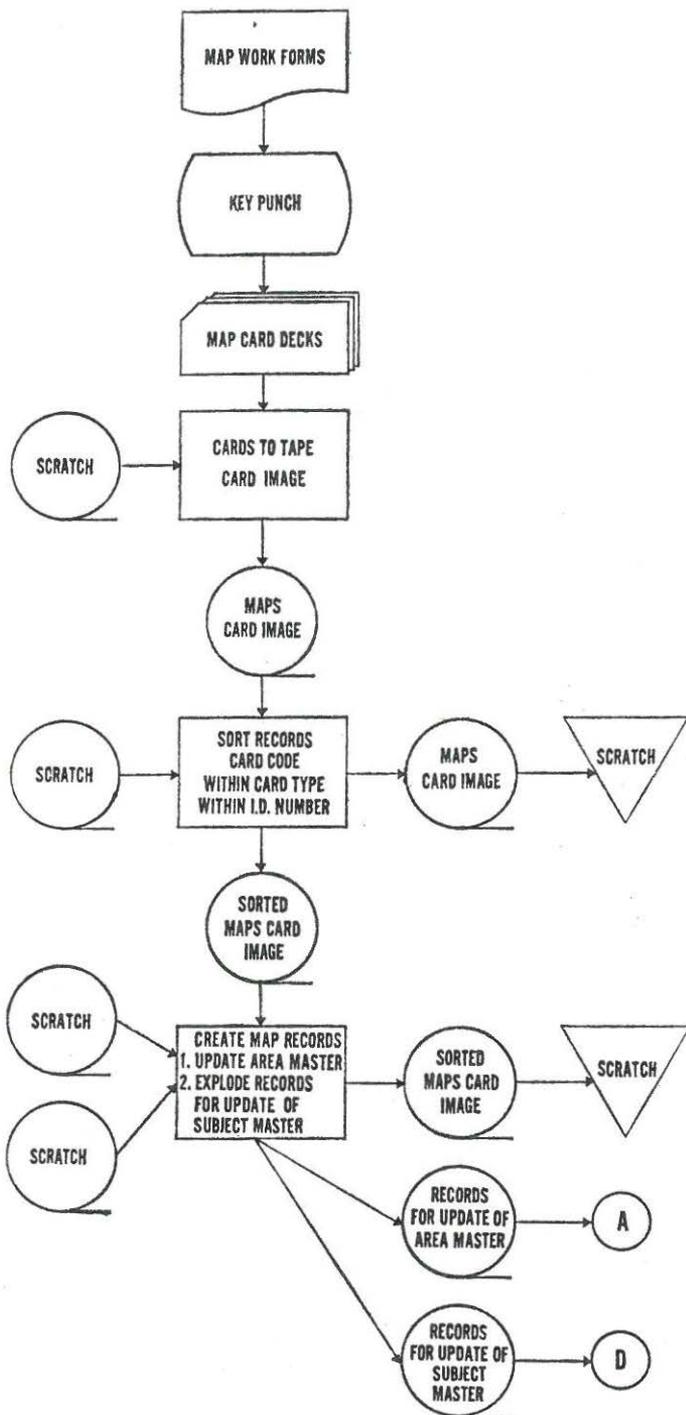


Fig. 2. Work Flow Chart.

TAPE LAYOUT - LM005

BLOCK SIZE = MAX - 2340

① SUBJECT MASTER & UPDATE TAPES

REC. SIZE = MAX. - 780

② AREA MASTER & UPDATE TAPES

IDENTIFICATION NUMBER ⑥	STATUS A = ADDITION C = CHANGE D = DELETION ①	AREA ⑤	SUB AREA ③	SUBJ. MAJ. ③	DATE OF PUBLICATION ④	SHEET NUMBER ⑩	COPY ④	SCALE ③		
1	6	7	8	12 13	15 16	18 19	22 23	32 33	36 37	39

PROJECTION ②	LANGUAGE ③	LOCATION ②	FORM ②	SIZE ⑥	CODE NAME ③	① SUBJ. MASTER SUBJ. CODE ③ ② AREA MASTER- ALPHA - CODE ④	SPARES ⑰ or ⑳
40	41 42	44 45	46 47	48 49	54 55	57 58	60 or 61 61 or 62 80

TITLE LENGTH ②	PUBLISHER LENGTH ③	AUTHOR LENGTH ③	NOTES LENGTH ③	TITLE (PLACE & SUBJECT) MAX = 215	PUBLISHER MAX = 129	AUTHOR MAX = 129	NOTES MAX = 215
81	83 84	86 87	89 90	92 93			

VARIABLE LENGTH SECTION

Fig. 3. Layout for Area and Subject Masters and Update Tapes.

The equipment used is all IBM. The cards are punched on an 029 keypuncher and verified on an 059. A computer model 360/50 is now being used, though equipment of this capacity is not necessary. During the development of the project an IBM 1401 and later a 360/40 were used. Printing is done on an IBM 1403 at 1100 lines per minute.

The Programmes and Their Functions.

The following nine programmes which were originally written in auto-coder are now in PL/I. This is a relatively new high-level language for the IBM 360 system. To have maximum efficiency from this language large core storage is necessary, though it can be used, with restrictions, on a 32K core storage computer. With the use of other programming languages the system could run efficiently on any computer.

LM001: This programme puts the card decks (from keypunching) onto tape in card image.

LM005: This programme creates and explodes each group of records on the card image tape with the same identification number to produce a subject update tape and an area update tape (Figure 3). At the same time, each record is edited; if an error is found, the record is rejected. In order for a record to be valid, the following conditions must exist:

- 1) Numerical Identification Number.
- 2) Valid card type, i.e. (1, 2, 3, 4, 5) (See Figure 1).
- 3) No duplicate cards for the same map.
- 4) Card codes successive.
- 5) Area being 'G' followed by four numeric digits.
- 6) Numerical date.
- 7) If scale absent, 'Z' (not printed).
- 8) General information card and title card for each map.

LM010: In this programme, the area master is updated with the area tape. An error message is printed and the record rejected if there is already an addition on the master file or if there is a change or deletion having no corresponding record on the master file. Also the area number is checked against a table to see if it is valid. If it is invalid, an error message is printed but the record will appear in the master file.

LM015: This programme lists the alphabetical geographical master.

LM025: This programme lists the area master geographically.

LM030: This programme updates the subject master. A message is printed and the record rejected if there is an addition which already exists on the master file and if there is a change or deletion which has no corresponding record on the master file. Also, the subject code is checked against a table to see if it is valid. If the subject code is invalid, an error message is printed but the record will appear in the master file.

- LM035: This programme lists the subject master. At the same time, a tape is produced with each subject heading and the page number it appears on.
- LM036: This programme lists a table of contents for the master subject list.
- LM037: This programme lists an index for the master subject list.

THE CATALOGUE

The catalogue is a book catalogue produced in three sections, unburst and top-bound in loose-leaf binders. The first is the classified or shelf list section, which brings maps of adjacent areas together. Within each L.C. number or area, maps are by subject code. In L.C. order, general maps are followed by those with subject emphasis, then by those showing political divisions, ending with cities. Area names and numbers are in bold type. All pages are numbered and there is a table of contents giving area name and L.C. equivalent. There is also a list of subjects with code numbers. Section two is the same list in alphabetical order by area name (Figure 4). Section three (Figure 5) is the subject listing. Maps are arranged by L.C. subject code rather than alphabetically, which gives the advantage of grouping related subjects together. Within each group maps are in class number (ie. area) order. In the format for this section the L.C. alphanumeric code is given first, with the subject name in bold type

ANTIGUA. . . . G5047-G5050

G5047 1959

ANTIGUA, WEST INDIES (ANTIGUA ISLAND).

1:25,000. TRANSVERSE MERCATOR. GREAT BRITAIN.

DIRECTORATE OF OVERSEAS SURVEYS, LONDON, 1962.

SET OF 2 MAPS.

Fig. 4. Alphabetical Area List.

J80 INDUSTRIAL AGRICULTURAL PRODUCTS

G8481 J80 1959

RHODESIA AND NYASALAND - TOBACCO (TOBACCO PRODUCTION . . . RHODESIA AND NYASALAND).

1:3 MILLION. RHODESIA AND NYASALAND. DIRECTOR OF FEDERAL SURVEYS, SALISBURY, 1961.

FEDERAL ATLAS MAP NO 20.

K FORESTS AND FORESTRY

K10 FORESTRY IN GENERAL

Fig. 5. Classified Subject List.

for major groups and in regular type for the subdivisions. An alphabetical index of subjects refers the user to the page where his subject begins.

The call number in all three lists includes only the major subject of the map, but since a map may cover several, up to five additional or "minor" ones may be included when cataloguing. A single sheet may therefore appear under several headings in the subject section. This method is also used to catalogue a single sheet containing several separate maps.

EVALUATION

Although some modifications may yet be made to the system, the catalogue has proven highly successful and possesses a number of advantages over existing manual and automated systems. Its clear format and lack of symbols make it easy to use. It is issued each trimester in three copies, of which one is kept in the Library. One copy is sent to the Geography Department, and one to the History Department.

The work form is simple enough to be used by skilled non-professional help and as all punched cards are verified there are fewer errors than with card typing. Some errors do occur, but in almost all instances the record is automatically rejected and corrections made. Filing errors are non-existent. Few codes are needed for input and only the L.C. number, form and location are not readily understood in the printed catalogue. Although language, scale, projection and subject are entered in code or short form, they appear in full on the lists. The codes used for form, language, and projection are very simple and reference to them is seldom necessary.

Main entry, author, and imprint are in variable length fields, allowing complete information to be given without codes or abbreviations. As main entry is the area name and imprint is by publisher rather than place, a gazetteer-index, and a list by publisher, as well as an author list or index, can be produced when required.

Although not envisaged as an important element, the provision of a punched card for notes has been most valuable. If a map is withdrawn from a journal, or has an accompanying brochure filed elsewhere, this is stated. Any further explanation necessary for an understanding of the map is also given.

Since all elements on the first card are in fixed fields it is possible to obtain lists on demand by subject, date, scale, language projection, etc. Although the extent of the Simon Fraser collection makes this impractical now, its potential for preparing bibliographies and machine searching is apparent.

An analysis shows that initial costs were not excessive. The programming time of two months was the largest single item at \$2,400.00. Computer time and forms to produce the three listings totalled \$110.00. The projected cost based on the present size of the collection is \$280.00 per

year, a figure which will increase as the collection grows. Key punching and verifying time is approximately 2½ minutes per map. While this is of course a cost factor, it is done at slack time by the Cataloguing Department, whose operators are paid from \$360.00 to \$400.00 per month. In a manual system, an additional clerk at \$3,564.00 would have been needed to type and file cards, and furniture for the cards would also have been required.

The disadvantages are now more evident than upon receipt of the first lists in June 1968. Use of the classified section has been slight except by the Library staff and it will be issued only once each year. The alphabetical area section is the most heavily used, but the arrangement of entries by L.C. code under each area is confusing. As the number of maps increases from one page to many the user finds it increasingly difficult to locate a thematic map. The third section, by subject, helps overcome this problem, but here again the list is by L.C. code, not alphabetical within each subject.

Topographic series and sets were catalogued with one entry so the number of records was considerably less than the 25,000 sheets in the collection. Archival and facsimile maps acquired since the system was designed have presented problems.

The Librarian and Library Assistant were new to map work; consequently the number of errors was high, and corrections and patching-up were time consuming and therefore costly.

CONCLUSION

Despite the less than perfect product, however, the results are worthwhile. First time users experience some difficulty with the classified arrangements but only a simple explanation is needed, and thereafter students are able to identify and locate most maps with little reference to staff. The Geography Department, and to a lesser extent the History Department, do make use of their copies of the catalogue. Telephone enquiries for holdings are minimal and some faculty have asked that they be given their own copies.

The Simon Fraser system is not expensive to operate, the catalogue could be issued more frequently at little extra cost, and the system uses a widely accepted classification scheme that is updated periodically. The programmes employed could be adapted by other libraries with few, if any, modifications and the system could be run on any computer.

There will be more sophisticated map catalogues, such as that of the Library of Congress, using MARC II format, and others which will take greater advantage of computer capabilities. Extensive and costly research, however, will be needed to develop these systems. The Simon Fraser system is operating now, was developed in a very short time, and has had a successful first year of use.

REFERENCES

1. Murphy, Mary: "Will Automation Work for Maps?" *Special Libraries*, 54 (November 1963), 563-567.
2. Thomas, Kenneth A., Sr.: "The San Juan Island Project: Cataloguing Maps by Mechanized Techniques," *Special Libraries Association, Geography and Map Division. Bulletin*, 54 (December 1963), 8-12.
3. Donkin, Kate; Goodchild, Michael: "An Automated Approach to Map Utility," *The Cartographer*, 4 (June 1967), 39-45.
4. Stallings, David Lloyd: Automated Map Reference Retrieval. A thesis submitted in partial fulfilment of the requirements for the degree of Master of Arts. (Seattle, University of Washington, 1966), p. 71.
5. Donahue, Joseph C.; Hedges, Charles P.: "CARES - A Proposed Cartographic Retrieval System," *American Documentation Institute. Proceedings*, (1964), 137-140.
6. Hagen, Carlos B.: "An Information Retrieval System for Maps," *UNESCO. Library Bulletin*, 20 (January-February 1966) 30-35.
7. Easton, William W.: "Automating the Illinois State University Map Library," *Special Libraries Association, Geography and Map Division. Bulletin*, 67 (March 1967), 3-9.
8. United States. Library of Congress. Subject Cataloging Division: *Classification, Class G: Geography, Anthropology, Folklore, Manners and Customs*; 3rd. ed. (Washington: Superintendent of Documents, 1954), p.502.