COMPARATIVE COSTS OF CONVERTING SHELF LIST RECORDS TO MACHINE READABLE FORM

Richard E. CHAPIN and Dale H. PRETZER:

Michigan State University Library, East Lansing, Michigan

A study at Michigan State University Library compared costs of three different methods of conversion: keypunching, paper-tape typewriting, and optical scanning by a service bureau. The record converted included call number, copy number, first 39 letters of the author's name, first 43 letters of the title, and date of publication. Source documents were all of the shelf list cards at the Library. The end products were a master book tape of the library collections and a machine readable book card for each volume to be used in an automated circulation system.

The problems of format, cost and techniques in converting bibliographic data to machine readable form have caused many libraries to defer the automation of certain routine operations. The literature offers little for the administrator facing the decisions of what to convert and how to convert it.

Automated circulation systems require at least partial conversion of the accumulated bibliographic record. The University of Missouri, like many libraries, has been converting the past record only for books as they are circulated (1). Southern Illinois University (2) and Johns Hopkins (3), on the other hand, have converted the record for their entire collections. The Southern Illinois program is based upon converting only the call number. Johns Hopkins has converted the call number, main entry, title, pagination, size, and number of copies. And Missouri has recorded call number, accession number, and abbreviated author and title.

Several methods of converting the record have been described. Missouri employed keypunching; Southern Illinois marked code sheets which were scanned electronically and converted to magnetic tape; Johns Hopkins, working from microfilm copy of the shelf list, used special type font and typed the records for optical scanning. An IBM report on converting the National Union Catalog recommended an on-line terminal as the best method of conversion (4).

Studies at Michigan State University led to the conclusion that acquisition, serials, circulation, and card production contained certain routines that might well be automated. Once automation of circulation was decided upon as our initial effort, decisions were necessary as to the conversion. It was recommended that a portion of the bibliographic record for all items in the shelf list should be converted. Information other than the call number is being used for other programs (5).

Cost figures for converting library records are scarce. In only two instances are figures available. The IBM report on the National Union Catalog shows that the average entry in NUC contains 277 characters, with an estimated conversion cost ranging from $0.3531 to $0.417 per entry. The proposed conversion method employs an on-line terminal, a technique not available to most libraries.

The Johns Hopkins conversion of "about 300,000 cards" was accomplished by optical scanning and cost $18,170 (3,p.4). This figures out at about $.06 per record. Later in the report it is stated that the conversion "is at a rate of $.0038 per character converted" (3,p.25). At $.06 per card and $.0038 per character, the converted record would consist of 16 characters!

In the study herewith reported every effort was made to arrive at comparative cost figures for the three methods of conversion that are readily available to most research libraries: keypunching, paper-tape typewriting, and optical scanning as accomplished through a service bureau.

METHODS OF STUDY

The shelf list records of the Michigan State University Library were divided into three sections by numbering catalog drawers in sequence: 1,2,3; then 2,3,1; then 3,1,2. All the drawers marked with number one became one sample group; those marked two and three made up the other groups. This method of numbering the drawers gave samples from each area of the classification schedule for each method of conversion.

The bibliographic data were taken directly from the shelf list without transferring information to worksheets. A sample of the shelf list shows that 74 per cent of the cards are Library of Congress cards or copies of Library of Congress proof slips. Of those cards produced in the library, only 12 per cent of the total were abbreviated records.

The keypunch operators, the typists, and the service bureau were in-

structed to extract information from the shelf list record. All differences in type—capitals, italics, etc.—were to be ignored; transliterated titles were to be used in those cases where entries were in non-Roman alphabet; accents and diacritical marks were ignored, except where it made a difference in filing, as with umlauts; all numbers in title and author fields were to be spelled as if written.

---

① ②

QD
941
.A413

Agranovich, Vladimir Moiseevich.
③ Spatial dispersion in crystal optics and the theory of excitons [by] V. M. Agranovich and V. L. Ginzburg. Translated from the original manuscript by Literaturprojekt, Innsbruck, Austria. London, New York, Interscience Publishers [c1966. ④

c.1 ⑤
c.2
c.3

xxxi, 316 p. illus. 24 cm. (Interscience monographs and texts in physics and astronomy, v. 18)
Translation of Кристаллооптика с учетом пространственной дисперсии и теория экситонов (romanized: Kristallooptika s uchetom prostranstvennoĭ dispersii i teoriia ĕksitonov)
Bibliography: p. 307-313.
1. Crystal optics. 2. Exciton theory.   I. Ginzburg, Vitaliĭ Lazarevich, 1916–   joint author. II. Title.   (Series: Interscience monographs in physics and astronomy, v. 18)
QD941.A413   548'.9   66-29947

Library of Congress   [8]

---

① JX
1417
.W47

②

Whitaker, Urban George, 1924–   ed.
③ Democracies and international relations: can ours survive? San Francisco, Chandler Pub. Co. [1961, ④

c.1 ⑤
c.2
c.3 BUSINESS LIB.
c.4 UNDERGRADUATE
LIBRARY

Fig. 1. Shelf List Cards.

Information that was transcribed is marked in the example, Figure 1. The complete call number 1) was included. Author 2) was typed through 39 spaces, including dates, if possible. In cases where author entry was lengthy the operators were instructed to stop at the end of 39 spaces. Title 3) was recorded as completely as possible through 43 spaces, but not to extend beyond the first major punctuation. Date 4) was included as shown. Only one copy 5) was shown on each entry. In the example of abbreviated form in Figure 1, five separate records were required, with change only in copy number.

The master book tape includes the call number, which occupies 32 spaces; 3 spaces are allowed for copy number, 39 for author, 43 for title, and 4 for date of publication. On the book card, Figure 2, which was generated by the computer from the master book tape, the format is as follows: 32 spaces for call numbers, 3 for copy number, 11 for author, 26 for title and 4 for the year published. The remainder of the card is for machine codes used in the circulation system.
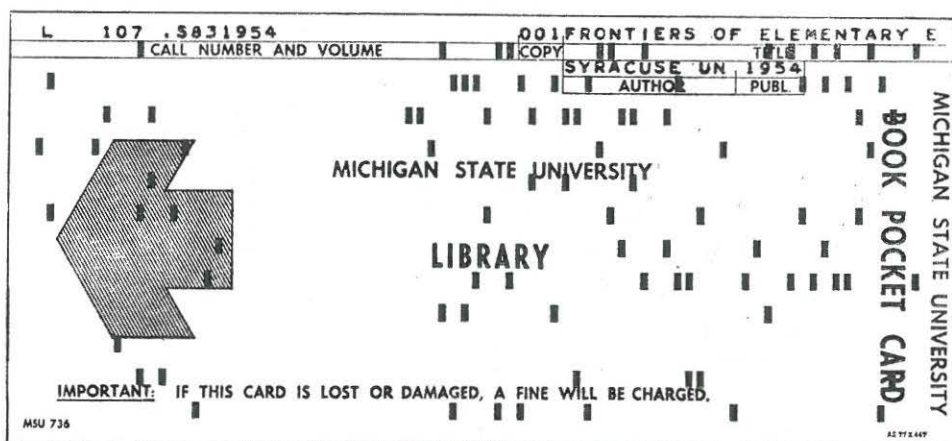

Fig. 2. Book Pocket Card.

The book card alone can be created directly by the keypunch. However, if a library has equipment available for a more complete program, it is useful to prepare information in a format to create a master book tape. Programs have been written so that the master tape can be added to or deleted from at a later date.

Four operators worked on the project at Michigan State University. Two of them were average keypunch operators with little typing skill, one was an expert typist, and the other was an expert keypunch operator. The first two operators were trained to use both the keypunch and the Flexowriter. The purpose in using a variety of typists and operators for the job was to arrive at average figures for the conversion project. The data show great variance of output among operators.

The outline of the methods used is shown in Figure 3. The keypunch method recorded the bibliographic data by use of an IBM 026 keypunch. The punch cards were transferred to a magnetic tape and the book cards were generated by the computer.

The paper-tape typewriter information was punched in paper tape by the use of a 2201 Flexowriter. A portion of the sample was converted directly to magnetic tape. Since some libraries will not have a paper-tape to magnetic-tape converter, the remainder of the paper-tape sample was converted to punch cards and then to magnetic tape.
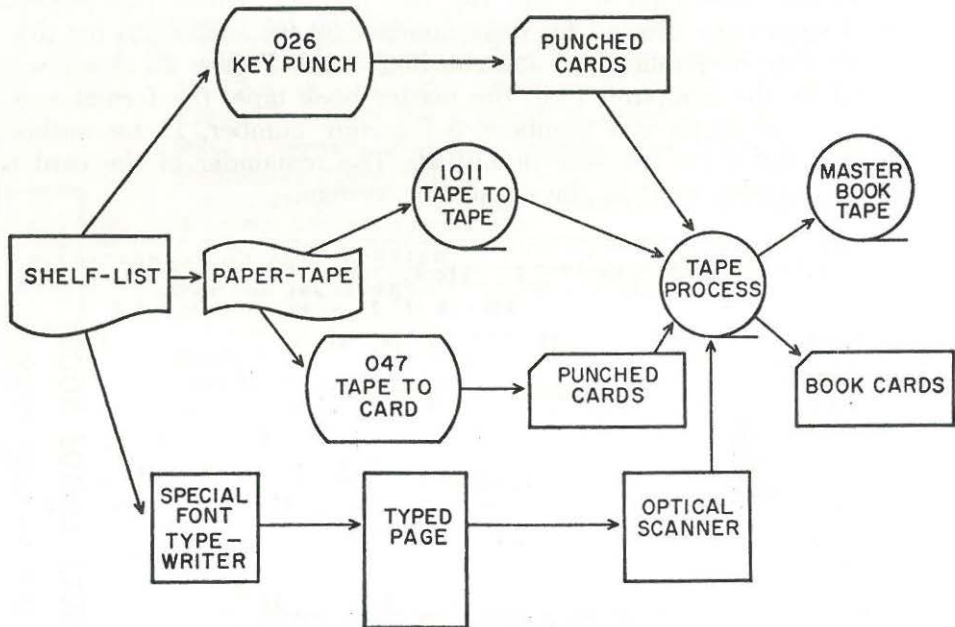

Fig. 3. Flowchart of Shelf List Record.

The optical scanning method was handled by Farrington Corporation's service bureau, Input Services, in Dayton, Ohio. The service bureau assigned 10 to 15 employees to transcribe the shelf list. They used IBM selectric typewriters, with special type font. Special symbols were used to designate end of field. The data were recorded on continuous-form paper. The typed record was then edited and scanned, producing a magnetic tape. After the tape was used for production of book cards, it was added to the master book tape.

The first batch of cards sent to Dayton was gone from the Library for approximately four weeks. After the personnel at Dayton became accustomed to the format and to library terminology, the turnaround time was approximately two weeks. The 255,000 records which were converted by the service bureau were sent off campus in four separate batches.

Machine verification of the record was not required. Each operator was instructed to proofread her own copy. Machine verification was considered, but the idea was discarded because of the extra cost involved. Also, since book cards were to be inserted in all volumes, final verification would result when the books and cards were matched.

RESULTS

In the conversion keypunching cost 6.63 cents per record. Paper-tape ran slightly higher—7.07 cents; this higher cost was due to the added cost of machinery and the added cost of going from paper tape to magnetic tape. Optical scanning, through a service bureau, was exactly the same as keypunching—6.63 cents, including the programming costs. Cost details are shown in Table 1.

Table 1. Average Cost Per Shelf List Record Converted

	Keypunch	Paper-tape Typewriter	Scanning, Service Bureau
Labor (1)	$.04073	$.03960	
Salary	.03723	.03620	
Fringe Benefits	.00350	.00340	
Equipment	.00322	.00888	
Rental (2)	.00280	.00840	
Computer	.00042	.00048 (3)	$.00030
Supplies	.00003	.00052	
Overhead (4)	.02232	.02172	
Contractual Services			.06600 (5)
TOTAL	$.06630	$.07072	$.06630

(1) Average costs for all operators based upon salary of $2.10 per hour, and fringe benefits of 9.4 per cent.

(2) Rental time to Library of IBM 1401 computer is $30.00 per hour, including personnel costs.

(3) Includes $.000089 for tape-to-tape conversion and $.000091 for tape to card to magnetic tape conversion.

(4) University charge of 54.87 per cent of salaries, for space, utilities, maintenance, etc. This figure does not include cost of training and supervision.

(5) $.057 per record plus .009 per record for programming costs.

Late in the study we observed that a seemingly inordinate amount of the Flexowriter time was consumed by the automatic movement of the typewriter carriage to the pre-determined fixed fields. In order to circum-

vent this the operator was instructed to strike one key to indicate end of field, and then she no longer had to wait for the carriage movement. By using the manual field markers, as opposed to automatic fixed fields, the cost of the Flexowriter operation was reduced to 6.672 cents per record. The disadvantage of the manual field-marking system was the increased chance of operator error, which amounted to 3.13 per cent more than the fixed-field method. For this reason, and in spite of the economy of the manual method, the use of pre-determined fixed fields for Flexowriter conversion is to be preferred.

In the comparison of the salary costs for keypunching and for the use of Flexowriter, great variations were shown among operators. Two participants were asked to use both the keypunch and the Flexowriter on varying days, with tallies of their output accounted for throughout the entire project. Operator 1 was essentially a skilled keypunch operator who had some background in typing. Her salary cost per record during keypunching was 3.98 cents; her salary for the paper-tape typewriter was 7.92 cents. Operator 2 was a skilled keypunch operator who was also sent to typing class for one term to raise her typing skill. Her salary cost was 3.92 cents per record on the keypunch and 3.79 cents per record on the paper-tape machine. Operator 3, who was a skilled keypunch operator, averaged 2.32 cents per record for salary cost. Operator 4, who was a typist and not a keypunch operator, produced records on the Flexowriter at a cost of 3.56 cents per record. The above figures indicate salaries only, and do not include overhead, fringe benefits, and other expenses which are reflected in the total conversion cost shown.

A letter from Farrington Service Corporation stated the following information about the scanning operation: "1) Our typists produced an approximate total of 7,950 typing pages in the course of this conversion. 2) Each typist averaged from 3.6 to 3.8 pages per hour. 3) We processed an average of 800-1,000 (shelf list) cards, per girl, per day. 4) The total man hours expended in this project was 2,144. 5) The amount of error detected as a result of sight verification varies significantly from girl to girl. The average, however, ran approximately 2.8 per cent (of records to be corrected)."

Comparison was made of actual records converted per eight-hour day by each of the methods. The service bureau, with skilled typists, was able to convert approximately 100 records an hour for each typist. The most efficient keypunch operator averaged about 75 records per hour, which was noticeably more than the average. The paper-tape typist, using pre-programmed fixed fields, reached 65 records per hour, but was able to produce 73 records per hour by manually typing the field markers.

A short-run sample was stop-watch-timed to give an indication of the differences in results for each method when only minimum changes in certain fields, such as copy number or volume number, were required.

On the keypunch machine an operator consumed 34.6 seconds in typing the initial record and 20.4 seconds in duplicating the basic information and changing data in one given field. The operator with the automatic program Flexowriter consumed 47.2 seconds typing the initial record, including 13.2 seconds in shifting fields and automatically firing the record marks, and 24 seconds duplicating the record. When she manually indicated the field information, she was able to convert the initial record in slightly less time—30 seconds; and she took 22.8 seconds to duplicate the data with a change in one field.

Final verification will be completed only when all cards are matched with the proper books. For those books that do not circulate, this may never be accomplished. A sample of cards was selected to reflect the three methods of conversion. The service bureau cards contained fewer errors than those produced by keypunching and paper-tape typewriting. Production of records that were not acceptable to the computer in an edit program occurred in 1.75 per cent of the sample for keypunching, 0.93 per cent for paper-tape typewriting, and 0.16 per cent for service bureau. Operator errors, discovered while matching cards with books, showed a higher percentage: 4.62 per cent for keypunching, 3.60 per cent for Flexowriter, and 0.35 per cent for service bureau.

CONCLUSIONS AND RECOMMENDATIONS

1.   The cost of converting a portion of the bibliographic record is relatively inexpensive when compared to the total cost of automated library programs. One reason for our delay in entering into the field of an automated circulation program was that of making the book cards. Now that this task has been completed, it is obvious that conversion is a one-time cost that can well be absorbed. If the library cannot afford the original conversion, at a cost of 6 or 7 cents a record, then the library cannot afford to proceed with automated programs.

2.   There is no difference in cost between keypunching a machine readable record and in having the project undertaken by a service bureau. The use of paper-tape typewriter for conversion costs more than the other two methods.

3.   Large scale conversion of records to machine readable form might well be done by an outside organization. In order to get the task completed in a short period of time, a library would be required to hire a number of short-term clerical employees. In the case of Michigan State, situated in the small community of East Lansing, recruiting and training a large number of employees for short-term projects is most difficult. It is rather certain that the overhead for such a program would bring the cost beyond that of using a service bureau. On the basis of our experience it is recommended that the conversion be sent to a service bureau.

4. A library can get along without portions of a shelf list for short periods of time. One of the predicted problems of sending material off campus to be converted was that of losing the availability of the shelf list records. Although there were some inconveniences, it was found that the library could carry on its operations and function without the shelf list. Certainly, this could not be done if the shelf list cards were gone for any length of time.

ACKNOWLEDGMENT

REFERENCES

1. Parker, Ralph H.: "Development of Automatic Systems at the University of Missouri Library," in University of Illinois Graduate School of Library Science, *Proceedings of the 1963 Clinic on Library Applications of Data Processing.* (Champaign, Ill.: Illini Union Bookstore, 1964), 43-55.
2. Southern Illinois University. Office of Systems and Procedures: *An Automated Circulation Control System for the Delyte W. Morris Library; the System and Its Progress in Brief.* (Carbondale, Ill.: Southern Illinois University, 1963).
3. The Johns Hopkins University. The Milton S. Eisenhower Library: *Progress Report on an Operations Research and Systems Engineering Study of a University Library.* (Baltimore: Johns Hopkins, 1965).
4. International Business Machines. Federal Systems Division: *Report on a Pilot Project for Converting the pre-1952 National Union Catalog to a Machine Readable Record.* (Rockville, Maryland: IBM, 1965).
5. Chapin, Richard E.: "Administrative and Economic Considerations for Library Automation," in University of Illinois Graduate School of Library Science, *Proceedings of the 1967 Clinic on Applications of Data Processing.* (In press).