

AN INTERACTIVE COMPUTER-BASED CIRCULATION SYSTEM: DESIGN AND DEVELOPMENT

James S. AAGAARD: Departments of Computer Sciences and Electrical Engineering, Northwestern University, Evanston, Illinois.

An on-line computer-based circulation control system has been installed at the Northwestern University library. Features of the system include self-service book charge, remote terminal inquiry and update, and automatic production of notices for call-ins and books available. Fine notices are also prepared daily and overdue notices weekly. Important considerations in the design of the system were to minimize costs of operation and to include technical services functions eventually. The system operates on a relatively small computer in a multiprogrammed mode.

INTRODUCTION

Although the Northwestern University library had given some consideration to the adoption of data processing techniques over a period of many years, it was not until planning for a new library building started that this consideration became serious. An associate university librarian and a systems analyst were added to the staff with specific responsibilities in the "automation" area. The recommendation of the systems analyst was that an on-line system should be designed to integrate all library functions. Two areas were isolated for initial development: technical services, including ordering and cataloging, and circulation control.

Several other decisions were made at about the same time (Fall 1967). Perhaps most important of these was the choice of computer. The acquisition of a dedicated library computer was ruled out on the basis of cost, leaving the choice to be made between a Control Data 6400 soon to be installed in the university's Computing Center, or an IBM 360/30 in the administrative data processing department. It was clear that the IBM 360 would have to be upgraded considerably to handle an on-line system, but the decision was made to use it, based on the facts that it was already installed and operating, that the machine itself was more adaptable to text processing

applications, and that the library was an administrative application. A small programming staff was available, and it was decided to use that staff rather than have the library develop its own programming capability. The university's engineering and science libraries were administratively divorced from the rest of the Evanston campus libraries to serve as a pilot location for development and testing.

One final decision was made by the programming staff. Since there was reason to believe that the use of a real-time system by the library might generate similar requests from other users of data processing services, the system should be capable of extension to other applications, if possible. Design then began on a general-purpose file maintenance system. A detailed description of this system will be presented in another paper.

Actual programming started in spring 1968, and in about a year the teleprocessing system was essentially complete and work was started on various subsidiary programs, to be run on a daily or weekly basis. These included programs for producing catalog cards, purchase orders, and similar materials.

However, at this time the realization came that the opening of the new building was less than a year away (construction was on schedule, unlike the situation with several other buildings), and the new library administration felt very strongly that it would be desirable to have an operational circulation system. Work then was suspended on the technical services part of the system, but it is important to note that the system developed up to that point provides the on-line inquiry capability to the circulation operation. It is also true that this capability is more sophisticated than would be needed for circulation applications only.

After the basic system design for the circulation system was completed in spring 1969, the massive job of preparing nearly a million punched cards for the books was started in the summer. This was done using student operators, working from the shelflist. The most expensive and time-consuming part of the job, however, proved to be the insertion of the cards in the books and this was not completely finished before the new building opened in January 1970. The computer circulation system was not ready for operation until December, so that it was tested in the pilot library for only a three-week period, hardly enough for a complete cycle of book charges and discharges.

Operation in the new building was complicated by several factors besides a new and unfamiliar circulation system. The building itself was not quite finished, all of the books were not in place, all of the remote terminals were not installed, and there was a large backlog of work which had accumulated during the moving period. The most serious problem, however, was that a decision had been made to continue the old manual circulation system in parallel with the new one, and with the other problems this became too much of a burden on the library staff. When it ap-

peared that there were no problems with the new system which could not be worked out, the manual system was quickly abandoned. After this point operations began to improve rapidly, and within a few months the system was running quite smoothly. The systems and programming staff has now returned to the implementation of the technical services system.

GENERAL DESCRIPTION

Functionally the Northwestern University library circulation system may be viewed as consisting of three parts. The first of these is a book charge/discharge operation using the IBM 1030 series of terminals. The second part is the general-purpose file maintenance system, originally developed for technical services; and the third part is a group of programs which are run in "batch" mode and thus have no direct interaction with the remote terminals.

The teleprocessing program operates in a partition of 36,864 bytes of storage on a 65,536-byte computer. The IBM Disk Operating System is used, which requires 8,240 bytes, leaving 18,432 bytes for batch programs. The Basic Telecommunications Access Method is used for remote terminal input-output operations. Data storage is on an IBM 2321 Data Cell.

The present terminal configuration consists of five pairs of 1031 input terminals and 1033 printers, and four 2740 model 2 typewriter terminals (two of which are used for technical services development). The partition size is probably adequate for one or two more terminals. All of the 1030 terminals share a common telephone line, as do all of the 2740 terminals. Two of the 1030 terminals are master units, connected directly to the telephone line, while three 1030s are satellites, operating through a master terminal. One master 1030 and one 2740 are located in the Technological Institute library; the remaining terminals are in the main university library.

BOOK CHARGE SYSTEM

Each of the 1031 terminals will accept an 80-column punched card which is kept in the book pocket, and also a punched plastic user identification badge. The three satellite terminals are located in the stack area of the library, one on each of three floors adjacent to the elevators. They are used for self-service charge of books. A master terminal, which also includes a manual entry keyboard, is located at the circulation desk on the main floor and is operated by library staff. The keyboard on this terminal allows the staff to perform additional functions, such as charging books to users without badges, charging for periods other than the standard loan period, processing renewals, and discharging books which have been returned.

The 1033 printer associated with each terminal is really just a modified electric typewriter. When a book is charged and the transaction is ac-

cepted by the computer, the printer creates a date due slip which shows the call number of the book, the identification number of the user, and the date due. This is placed in the book pocket and serves as the borrower's pass to carry the book past the exit guards. To make this a reasonably secure system, the guard must verify that the call number printed on the slip corresponds with the number on the book, and the user number on the slip corresponds with the number on a valid university identification badge. Note that this system permits a user to bring the book back into the library and take it out again as often as he wishes during the loan period.

The 1030 terminals are associated with a small, specialized part of the computer teleprocessing program which accepts the information from the terminal, reformats it so that it is compatible with the general-purpose file maintenance system, and enters it in the file, checking, of course, that the same book is not already in the file. (Transactions which are invalid for one reason or another result in an "unprocessed" message on the printer, and the user must then go to the circulation desk to have the problem resolved.) This portion of the system also processes renewals and discharges from the terminal at the circulation desk.

INQUIRY SYSTEM

Inquiry requests use the general-purpose file maintenance system originally developed for technical services. This gives an operator at the circulation desk the capability to query the file about the status of any book, and also to make certain changes in records, such as indicating renewals and saves. The terminal used for this purpose is the IBM 2740, which is similar to a typewriter in operation.

In order to facilitate the inquiry operations, the key to each record is the actual call number of the book, rather than some arbitrary accession number. The call number is divided into two parts, which we call the search key and the key extension. The search key includes the Dewey classification number, Cutter number, and up to four work letters; the key extension includes other information such as edition number, volume number, or copy number. A few compromises were necessary with the key extension in order to adapt it to the limited character set that the 1030 terminals can process, but these changes have not caused any difficulty.

When a record is first entered in the file, the search key is used to calculate a position in the file. This is done by taking the alphanumeric characters in the search key, performing some mathematical operations to reduce the number of characters to four, and then treating the result as if it were a number. This resultant number is divided by a constant number which is chosen to be the largest prime number which is smaller than the number of tracks on the storage device for the file (the data cell). The remainder from this randomizing computation gives a location where an attempt is made to place the record. It is quite possible that this place

in the file will be occupied by some other record, which might have the same search key or a quite different one. Additional steps are provided to find an alternate location in such cases. Additional information about the file organization is given in the appendix to this article.

When a record must be located in the file, the same randomizing procedure is followed. Then, the complete keys of any records which are found at the calculated location are examined to find the desired record. However, the operator at the inquiry terminal has the option of requesting a search for records based on search key only, and this is often useful.

Even though the individual records for each book are relatively short (sixty-seven characters), the cost of maintaining a record in the file for every book in the library would be prohibitive. For this reason the philosophy has been adopted that an entry will be made in the file only for a book which is not in its proper place on the shelves. Thus the file includes entries for books which have been placed on reserve, loaned to library departments, or reported lost or missing. There are probably more records of this type in the file at any given time than records representing books borrowed by individuals.

Unfortunately, the library does not have the terminal or personnel capacity to verify the status of all books before the user leaves the card catalog area (adjacent to the circulation desk) so that if he doesn't find a book on the shelves he must return to the circulation desk to request an inquiry. In many cases, of course, the user may be satisfied with a nearby book on the same subject, or perhaps he just intends to browse in a general subject area. It is hoped that at some future time it will be possible to obtain inquiry terminals which can be user-operated and located in the stack area.

An additional feature of the system is the capability to place a "save" on a book which is on loan. This is done by the operator at the inquiry terminal, who adds the saver's identification number to the record of the book. The save request triggers a call-in notice, discussed below. This procedure has the minor disadvantage that a save cannot be entered so that the first copy of a particular book which is returned will be held; the operator must select a copy if more than one is out (or enter the save on all copies). This has not proved to be a serious problem.

The overall teleprocessing system includes a file of records, called the "transaction file," which is written sequentially. Records from other files which can be accessed by remote terminals are written in the transaction file under certain circumstances. In the case of circulation records accessed by the inquiry terminal, writing of the transaction file is under the control of the terminal operator, who will always request that the book record be written in the transaction file when entering a save on a record. Records processed by the 1030 terminals are entered in the transaction file only when they represent books which are discharged and which possibly involve a fine or contain a save.

BATCH PROCESSING

The third part of the complete system includes a number of "batch" programs, which are run periodically and are independent of the real-time program. However, they may be, and usually are, run at times when the real-time program is also operating. These programs use either the transaction file or the main circulation file.

Each weekday morning, a series of programs is run which processes the data entered into the transaction file since the previous run. Three types of printed notices are prepared by these programs as shown in the accompanying figures. One is a fine notice for books which were overdue when returned but for which a fine was not collected. (If a fine was collected the fact is indicated by a keyboard entry on the 1030 terminal when the book is discharged.) There also may be circumstances when the regular fine was collected, but a penalty fine is due because the book was called in but not returned in the specified time.

The second type of notice is the call-in notice. This is prepared from records which were placed in the transaction file as a result of a save entered by the inquiry terminal operator. The third type of notice is the book-available notice, which results when a book with a save is discharged. (When the actual discharge is performed the real-time computer program prints a message on the 1033 printer so that the book will not be returned to the stacks.)

After the selection of records from the transaction file, name and address information is added from student and personnel files maintained by the data processing department, and a four-part notice is prepared. The same printed form is used for all three types of notices; the additional copies of the fine notice are used in a manual follow-up system if the fine is not paid immediately. Processing of these notices generally takes less than ten minutes of computer time a day.

On a weekly basis, another series of programs is used to process the entire file of outstanding books. A considerably longer time is required because of the size of the file which must be examined. At this time all records in the file are also transferred to a backup file, which provides some protection in case of damage to the prime file (it has not yet been necessary to use it). Information is also extracted about books which have had transactions in the past week, and a circulation statistics report is prepared. Finally, records for books which are overdue are extracted and processed similarly to the daily notices, resulting in a one-part overdue notice.

On a quarterly basis the entire file is examined and lists are prepared of all entries which otherwise do not qualify for overdue notices. These include charges to reserve, lost and missing, other libraries, carrels within the library, and to faculty (who are not charged fines). These lists are distributed for verification that the books are actually located where the

file says they are and then returned to the circulation department for any further processing which might be necessary.

BACKUP SYSTEM

In designing a real-time circulation system it was obviously necessary to provide for those occasions when some equipment malfunction prevents normal processing. It was not felt necessary to provide any backup to the inquiry part of the system, and book discharges can be allowed to wait for several days if absolutely necessary. To process charges during a period when the computer system is not operational, a Standard Register Source Record punch is used. This device can accept the same plastic user badge and book card as the 1030 terminal, and transfers the punched information to a two-part form. The first part of the form serves as the date due slip, while the second part, a standard 80-column card, is used to enter the transaction through the 1031 when the real-time system is again operational.

This system has the advantage that it is completely independent of the real-time system, except of course for the building electric power. If for some reason the Standard Register punch cannot be used for a transaction (or for a book without a book card under any circumstances), the missing information can be handwritten on the two-part form. The second card part is then keypunched and the transaction entered through the 1031 terminal. This ultimate backup system is avoided, however, as it is very susceptible to transcription errors.

COSTS

Any attempt to determine the cost of the on-line circulation system is complicated by several factors. The cost of the terminals is an obvious item, and fairly easy to determine and allocate. However, the cost of the communications adapter which connects the telephone lines to the computer, as well as the cost of running the teleprocessing program, must be shared by all users. These now include technical services as well as circulation services, and in the future may include other nonlibrary university users. Finally, even if this allocation could be made, there is still the problem of separating the costs of the teleprocessing program and the batch programs being run in the data processing department.

However, since poor information may be better than none at all, the following figures are presented. They include monthly charges for the real-time program for both circulation and the part of technical services which is now operating, but do not include any charges for running of batch programs.

1030 Terminals

Master 1031A & 1033 with manual entry

2 @ \$251

\$ 502

Satellite 1031B & 1033	
3 @ \$155	465
2740 Model 2, 600 bps	
4 @ \$170	680
2701 Data Adapter with 2 lines	450
Telephone line charges	25
Data Cell Space, 5 cells	
(1 circulation; 4 technical services)	1,400
Core storage allocated exclusively to teleprocessing	1,700
Estimated share of CPU and disk costs	1,400
Special operator charge	350

FUTURE PLANS

Although the present real-time program includes a list containing a limited number of invalid user numbers (lost badges or users who are guilty of repeated violations of library rules), it would be more satisfactory to have a list of valid numbers instead. This might even be expanded into a self-regulating system, where users with a sufficient number of "demerits" would be prevented from charging additional books, and which might reduce the need for fines.

Another desirable addition to the system would be if some simple and inexpensive display terminals could be obtained and placed in the stack area to provide a self-service inquiry capability. This would relieve the circulation desk staff of some additional work, as well as reduce the number of trips the users must make to and from the stack area.

ACKNOWLEDGMENTS

The success of this circulation system is due in no small part to the constant support and encouragement from Mr. Howard F. Smith, director of Administrative Data Processing, and Mr. John P. McGowan, university librarian. Mrs. Velma Veneziano was responsible for establishing the library's requirements and rendered invaluable help during the implementation of the system. [*Editor's note: Mrs. Veneziano is preparing an updated (1972) summary of the actual application of this design, which it is hoped will be published shortly.*]

APPENDIX 1

Detailed Description of File Structure

The following information is included in each circulation file record:

Key

Dewey classification number—11 characters. (The decimal point is included in the record but not punched on book cards.)

Cutter number—5 characters.

Work letters—4 characters.

Key Extension

Volume, editor, copy—17 characters.

Location code—2 characters. This code indicates whether the book is in the main library or in a branch. It also provides for a subsidiary location code if required.

Large book indicator—1 character.

Charge code and date—3 characters.

Borrower number—5 characters.

Renewal code and date—3 characters.

Discharge code and date—3 characters.

Save date—2 characters.

Saver number—5 characters.

Due date—4 characters.

Terminal identification code—1 character.

Reserved for system use—1 character.

All dates and user identification numbers are packed to save space.

Because of the random organization of the file it is necessary that storage space be allocated for considerably more records than the expected maximum file size. This allocation conceivably might have been based on a simulation study of the file operation, but this could not be justified since not even a good estimate of maximum file size was available. (An important unknown factor was the change in usage patterns expected to result from the move to a new building.)

The only available estimates indicated a maximum file size of about 60-70,000 records, and on this basis sufficient space was allocated to hold 144,000 records. With twelve records to a track on the IBM 2321 Data Cell, the file requires 12,000 tracks, or 60 percent of one cell. An equivalent would be about one pack on an IBM 2314 disk storage unit.

Of this total space, five percent (about 7,200 records) is set aside as an overflow area; the remainder is the prime area. When the randomizing algorithm leads to a cylinder (twenty tracks) which is completely filled, the record is written in the overflow area and the cylinder at the prime location is flagged. Using this system, failure due to "running out of space" will be gradual; it is likely that system performance will be seriously degraded before the overflow area is completely filled. After more than a year of operation the file contains almost 63,000 records, of which about 230 are in the overflow area. Consideration is being given to minor changes in the handling of overflow records which should reduce the time required to search the overflow area.