

Optimizing Library Automation with a Central Dynamic Store

Too many automated library systems are focused toward a mechanized, batch-processing, card-oriented system. The technical processing areas of a large library (over 500,000 titles) are mainly business functions applicable to a Central Dynamic Store (Data Bank) concept. The implementation of such a concept utilizing an on-line computer system is discussed with emphasis on the acquisition, serials, cataloging, and circulation functions of a library.

TOO MUCH has been said about the increasing production of printed material in the world and the problems this mass of material will cause in our present libraries. Mechanization is thought to be the answer to this complex problem, but the inherent problems of mechanization have only added to the already existing complexities of the library environment.

This paper assumes that library automation is needed and discusses a new approach in library automation as distinguished from library mechanization. Library automation is the implementation of library procedures on a computer as opposed to punched-card equipment or mechanization. The approach taken in this paper is not new from a systems standpoint, but it is new in its application to a library environment. Specifically, the type of library referred to here is a university library. This does not mean that the methods discussed are not applicable to other types of libraries.

Mr. Jennings is Information Scientist and Library Systems Consultant with Arthur D. Little, Incorporated, Cambridge, Massachusetts.

The techniques of Business Information Systems utilizing a "Data Bank" store are used to eliminate the single-function machine concepts of the punch-card systems. The library as a business environment is complemented by a Central Dynamic Store (CDS) which houses selected files and which supports the simple procedures of the library and does not *pretend* to deal with all of the exceptions of the library process.

THE BUSINESS FUNCTION

The library as a business has always existed, but the library profession has not satisfactorily approached it from this perspective. The librarian can be viewed as a businessman with an information service as the product he is furnishing. Within the library, there are common business functions such as purchasing, marketing, bookkeeping, and shipping-receiving. These functions indicate a strong business attitude, and they should be treated as such for systems design simplicity.

If these business functions are defined, an approach for automating a business service environment can be taken. An initial step for definition is

the analysis of the existing practices for more centralized procedures within the business or the library. The library, like many businesses, exists for the buyer or user; therefore, its foundation must be centralized for strength and yet fluid for diversity in order to meet the demands on it.

The traditionalists have offered a solution to early business systems. This was the punched card. Without this initial media, large business systems would not have progressed to the more sophisticated on-line, time-sharing computer systems of today and tomorrow. The punched card systems, however, are becoming dated, and more effective systems are now being implemented. From the experience learned through early systems, why should libraries adopt dated card-systems for libraries? Before discussing the CDS approach for libraries, some points should be clarified about the card-system in the library environment.

THE CARD-SYSTEM LIBRARY

A majority of the library automation projects (these are mechanization projects) revolve around the unit key-punched card.¹ This approach to library automation is static. Card systems are not flexible for the selection and/or changing of specific information. Although card information is machine-readable, the validity of the information is questionable because of the difficult verification procedures employed. These and other qualities made card-systems

unsuitable for modern libraries as well as non-compatible with any other library environment than the one for which they were designed.

The mechanization of library procedures utilizing the card-system is normally implemented to a specific area of library processing (*i.e.*, circulation, ordering, and/or serials). Consequently, other areas of the libraries have suffered from the inherent problems of this mechanization. Considerations of the environmental controls and influences of the other areas within the library have not been recognized or have just been ignored. The area that is mechanized does not necessarily improve the total service or total organizational process of the library. It may, in fact, complicate or add needless problems to the library's other processing functions.

The product of a card-system is a list. The list may reflect a status of material (*i.e.*, on-order, charged circulation, or received) or financial statements for accounts. The list produced from cards is only an indication of status at the time the cards are processed, and does not reflect *current* or *now* status. The value of lists in a dynamic business environment is normally low. For purposes of control on a physical inventory as well as the normal processes of ordering, receiving, etc., a card system is *not* the answer.

The manual manipulation and storage of cards has always represented a problem. The card, like a slip of paper, is a physical piece of information. It must be maintained, routed, and stored. The problems relating to a box or tray of cards being dropped or mutilated do not have to be enumerated.

Cards can only be used as a tool to generate lists. Their use for conveyance of information to nonlibrary personnel or outside users is *null*. They are interpreted only on an internal library or processing nature. Because it is felt that library holdings must be accessible to

¹ Ralph E. McCoy, "Computerized Circulation Work: A Case Study of the 357 Data Collection System," *Library Resources & Technical Services*, IX (Winter 1965), 59-65; Carl C. Cox, "Mechanized Acquisition Procedures at the University of Maryland," *College & Research Libraries*, XXVI (May 1965), 232-36; Bruce W. Stewart, "Data Processing in an Academic Library," *Wilson Library Bulletin*, XLI (December 1966), 388-95; "The Use of Data Processing Equipment in Circulation Control," *Library Technology Reports* (July 1965), 24 p. (Widener Memorial Library, Harvard University Research Library at U.C.L.A., Research Library at the Thomas J. Watson Research Center of IBM Corporation).

outside users, information networks have been planned.² If this is to be a reality in the next few years, card systems will be incompatible in support of any information network of remote on-line nature. Current-status systems with accessible information to a remote user will be used in an information network. To ready a library for the eventuality of massive flows of information with accessible service to any user, the approach of establishing a Central Dynamic Store to support information networks must be planned. For the library card-systems in existence today, this means a repetition of initial analysis and planning to meet any on-line objective.

THE CENTRAL DYNAMIC STORE IN THE LIBRARY ENVIRONMENT

The CDS in a business as well as library environment is a working tool for the internal and external organization. It is a flexible and dynamic medium which is as fluid as the environment demands in which it is established. The characteristics of the CDS are expandability, contractibility, compatibility, and accessibility. With these characteristics, there is no question of its suitability in a library environment.

Physically, the CDS is a large-hardware, random-access memory device. This memory is a storehouse of any information desired to maintain the objectives for survival of the organization. The library business is maintained from a series of files of information that reflect the various functions or transactions of a current and historical nature. This means the CDS system can yield

in-process, to be ordered, current holdings, and historical status information for forecasting and maintaining current ordering, processing, and receiving functions.

The CDS is accessible from any communication-linked physical location in or out of the library facility. This is accomplished through a remote terminal device, preferably a cathode ray display device. Any individual employed by the library can perform the various functions of the library process on the CDS without disturbing anyone else using the CDS (*i.e.*, immediate access or no waiting). This is accomplished by operating the CDS on a time-sharing computer system.

Changes to any file or record within a file can be made at any time to reflect new information or updated information for any process within the library.

In a particular library environment the CDS contains the major files of the library. The files or information from the files is accessed within a real-time or query-response mode. The response to a query is fast enough not to jeopardize any immediate need for the response. The information gathered from the files depends upon the design of the files and the retrieval structures designed for the programming system.

To assist in the bibliographic verification of incoming orders, searching an in-process or outstanding order file along with the shelflist initially eliminates any duplicate ordering. (This was found to represent an average of 30 to 40 per cent of the incoming orders at Oregon State University.)

The ordering process is easily handled by the CDS through the introduction of the verified author-title and associated information to combine with a vendor file for the printing of the purchase order from a CDS order file on a continuous-printed form order. The printing of orders is done either through a re-

² U.S. Federal Council for Science and Technology, Committee on Scientific and Technical Information, *Recommendations for National Documents Handling Systems in Science and Technology*, U.S. Department of Commerce, National Bureau of Standards, Institute for Applied Technology, 1965; Charles J. Austin, "Medlars Project at the National Library of Medicine," *Library Resources & Technical Services*, IX (Winter 1965), 94-99.

quest to the system for immediate print, in the case of rush or direct orders, or batched ordering by vendor, on a daily or weekly basis. Once the order information is printed for mailing the original order information is transferred internally to an outstanding order file existing in the CDS. Because a historical-financial file is maintained by the CDS for serial-ordering reference, the order process incorporates serials ordering as well.

Serials processing has always been a problem in libraries because of their inconsistencies compared with monographs. Title changes, irregular publication, and noncommunication between publishers and receiving agencies can be assumed always to exist. The CDS system with its flexibility in design handles the majority of serial processing.

The check-in process of serials with the CDS system is an update function. The date of issues received is recorded after searching the serials holding record for the particular issue received. The binding and claiming functions of a serial or cataloging area are by-products of the CDS system. Internal programs hold in check all serial records to conform to predetermined algorithms. These allow the CDS system to inform technical processing personnel that a claim should be issued for a serial or that a serial is ready for binding.

Cataloging processes are a function of the CDS ordering process. If LC cataloging is available from bibliographic verification, a catalog flag is noted when the title has been placed on-order. Upon receipt of the material ordered, the system transfers all information for the shelflist record to the internal shelflist file. If a physical card catalog or book catalog is being maintained, generation of the appropriate entry is made or stored. If, during bibliographic verification LC cataloging is not available, updating of the shelflist, etc., is held for receipt and subsequent cataloging of the

document and the files are appropriately updated at that time.

The circulation function of the library business is accomplished without the standard unit card approach.³ The shelflist, merged with a student personnel file, produces a complete charge record for the material. If the library CDS system is used on a total university computer system, the student personnel records may be part of the registrar's system. Maintaining these records would not be a function of the library. Overdue notices are generated from algorithms based on prearranged or standardized checkout periods for faculty, civilians, and graduate and undergraduate students.

Another by-product of the CDS system is one of current operational statistics of the library. This by-product of the CDS is rare in today's libraries. In a manual environment, statistics are hard to retrieve because of the labor and time necessary for normal work loads. When collected, these statistics are of a historical nature and are not a current reflection. The current statistics are generated by the CDS system in patterns which are used to forecast budgets, physical storage capacities, and additional needs for a growing library.

Figure 1 illustrates a generalized approach of the CDS system for some of the processing and information needs of a library. With a little reflection, many more can be seen from the CDS concept.

CDS SYSTEM— PROGRAMMING TECHNIQUES

The CDS system is supported by a complex system of programming subsystems that is an integration of fact retrieval and storage-retrieval techniques.

³ Anne Flannery and James D. Mack, *Mechanized Circulation System, Lehigh University Library (Library Systems Analysis, Report No. 4 [Bethlehem, Pennsylvania: Lehigh University, Center for the Information Sciences, November 1966])*, 18 p. Appendices.

Because each modular routine is a separate entity, it is independent of the total program when writing, coding, and de-bugging. If modular or logical routines are intertwined as in batch-process programming, it is very difficult to modify a portion of the program without affecting the whole program. Anyone who has ever tried to modify a program written by someone else knows the difficulties of dissecting and patching another person's programming logic.

Modular programming incorporates the following methods or techniques:

Total (Main) Program. The main program executes a complete number of operations that constitute a task on input data. The main program may be an executive control program or an application program. It must make decisions governing the flow of all data to the modular routines it requires.

Multiple-Use Routines. If the same sequence of instructions is used two or more times by the main program, these instructions constitute a modular routine. The modular routine after execution of its instructions must return control to the main program.

Modular Routine Processing. A modular routine should not be dependent on any decisions made outside the routine, nor should any decisions made inside the routine determine processing outside the routine. The modular routine is responsible for all "housekeeping" for its own existence.

The construction of a modular main program in the CDS system uses an "overlay" technique and provides an efficient way for programming computer core memory programs. It conserves memory by the use of a physical part of memory for more than one modular routine. Sippl defines "overlay" as:

The technique of repeatedly using the same blocks of internal storage during different stages of a program. . . . The use of

one area in storage (core memory) to successively store more than one different subroutine or program parts.⁴

The main program calls the modular routine. After the modular routine has completed its sequence of instructions, control is returned to the main program to call the next module, etc. A mathematical description below shows this concept.

A main program resides during execution in core or main memory of a computer. The main program, P, is composed of a series of call statements,

$$\dots C_{pn}.$$

A list or library of modular routines,

$$\dots M_m,$$

is stored on some auxiliary memory such as magnetic tape, disk file, or in secondary core storage. Each of the modular routines has a return statement,

$$R_q,$$

where

$$q = n - 1.$$

During main program execution, each call statement solicits from auxiliary memory the desired modular routine. After system execution of the routine, program execution control is returned to the main program, P, for the next call statement in sequence (*i.e.*,

$$C_{p2} M_{35} \rightarrow M_{35} \rightarrow C_{p3} M_m.$$

R
3

Thus, a general equation for the total programming system (or main program) is

$$P = C_{pn}(L).$$

⁴ Charles J. Sippl, *Computer Dictionary and Handbook* (Indianapolis: Howard W. Sams and Company, Inc., 1966), p. 224.

Given: A main program, P
 A series of modular routines,
 $M_1, M_2 \dots M_m$
 A call statement, C_p
 A return statement, R_q

The structural configuration follows:

Core Memory	Auxiliary Memory
$C_{p1}M_m$	M_1 R_1
$C_{p2}M_m$	M_2 R_2
$C_{p3}M_m$	M_3 R_3
⋮	⋮
⋮	⋮
⋮	⋮
$C_{pn}M_m$	M_m R_q

Where generally,

$$P = (C_{p1}M_m, C_{p2}M_m \dots C_{pn}M_m),$$

$$m \leq n \geq m; n, m \neq 0$$

and a list, L, of modular routines,

$$L = (M_1R_2, M_2R_3 \dots M_mR_q)$$

Thus,

$$P = C_{pn} (L)$$

It has been found that programs written employing the modular programming technique are efficient both from the standpoint of core memory utilization and of program execution times. With the complexity of on-line real-time systems programming, the modular concept furnishes comprehensive and detailed documentation that may be understood by programmers at all levels of proficiency.

The CDS system is a new concept, and it is of an experimental nature today. Consequently, change and modification are characteristic. The modular

technique in programming structures meets the changing design and is necessary for the CDS system where programming logic becomes complex.

FILE ORGANIZATION

The file structures of the CDS must be planned to allow a maximum of retrieval flexibility. This can be accomplished economically by fixed field formats. This strategy allows any piece of information within any file or record to be addressed for retrieval by the programming system. The retrieved parameter can be displayed either for referral type of information or updating operations.

The fixed field format of a title record allows future expansion of the system for selective dissemination of information (SDI) purposes. This type of system function is an alarm system to announce to personnel receipt of material that they may have interest in reading. A university research environment has a high usage potential for this type of by-product from the CDS system.

For standardization, the author-title record formats for the CDS or any library system should be built around the MARC (Library of Congress) format. If file structures conform to the MARC formats, compatibility with the new-title magnetic tapes from the Library of Congress will exist.

The files in the CDS environment are accessed by main entry, LC cataloging entry, or any other parameter of information that is used for a look-up key. These parameters can be combined in a Boolean logical search request.⁵ Dynamic file structuring maintains a linking organization of files that allows retrieval of series of titles, volumes, etc.

⁵ Michael A. Jennings, "Construction of Boolean Search Operations in a Coordinate Indexing System," *Proceedings of the American Documentation Institute 1966 Annual Meeting, October 3-7, 1966, Santa Monica, California*, Adrienne Press, 1966. p. 17-27.

HARDWARE CONSIDERATIONS

To support CDS approach, an on-line, time-sharing computer configuration is needed. This is understandably a cost that would stop most libraries from implementation of this type of system. But the concept of time-sharing lends itself toward a community or cooperative ownership among a group of libraries.⁶ The shared concept helps to dissolve the total costs in maintaining a computer facility by a single library. If the library is in a university environment, the Computer Service Center of the university may be already operating a time-sharing system. The computer does not necessarily require a large core memory to support the CDS system because most operations are input and output (I/O) file processing and large computation programs are not involved.

The storage unit for this system is of primary concern. Random access storage is costly today for the capacity to hold a total file system or CDS approach for a large library environment (500,000 titles). The IBM data cell is the only mass storage equipment on the market today that approaches a reasonable storage-cost-bit. For example, the CDS system for a library of approximately 500,000 titles must employ at least one data cell of maximum capacity or approximately \$35,000 per year for storage costs.

Remote consoles in the form of both teletypes and cathode ray displays are needed to support the remote on-line inquiry and output printing-processes of the system. The number of terminals is a function of the system's environment.

It should be mentioned that in designing a system as discussed here, hardware backup should be considered. The critical importance of the service rendered

by a library utilizing the CDS system approach must be protected with backup systems. If the main system were to fail, backup storage and master files should be maintained for recovery from a failure. Backup hardware in the form of magnetic tapes is one solution. Other solutions can be employed depending on the critical importance of the functioning environment and monies available.

CONCLUSIONS

Improving manual procedures (in formulating the manual processes) with the thought of eventual automation must be of initial concern. The computer and the CDS system is the means of assisting in library procedures of a clerical nature. By no means does this paper and its concepts offer a cure-all to all complexities of a library business. But, if properly implemented, the system frees librarians from routine tasks in order to concentrate more fully on "snags" or the decision-making processes for which they are trained.

The costs are *never* reduced through library mechanization or automation. A leveling of costs may be felt if an automation system is allowed to run for about five years. After this, costs could settle to a palatable level. Expanded user service is a normal rationalization for high costs. Until deeper subject categorizing is employed this service can be practical only in a small specialized subject area or from the broad LC subject categories. Lately the expanded user service hypothesis has resulted in the *olde* practice of book catalog issue.

The notions discussed in this paper are not impractical from a standpoint of technology. Future automation projects, if they are to be meaningful and practical, will or should utilize the CDS concept. With these considerations, the compatibility with any future information network hookup on a state or nationwide level can be met. ■ ■

⁶ "New England Interstate Agency Plans Regional Library Center," *Scientific Information Notes*, VIII (October-November 1966) p. 5, 6.