

PHILIP L. LONG
Associate Director
Ohio College Library Center
Columbus, Ohio

OCLC: From Concept to Functioning Network

The Ohio College Library Center (OCLC) is a not-for-profit corporation chartered by the state of Ohio to provide a means for greater cooperation among libraries of all types in Ohio and in regional library systems outside Ohio. The members of the Center are approximately fifty academic institutions and public libraries in Ohio. The origins of the Center go back to the 1950s when various efforts were sponsored by members of the Ohio College Association toward bringing about greater cooperation in the field of library service.

These various efforts came to fruition when, in the summer of 1967, the Ohio College Association "spun off" the Ohio College Library Center as a separately chartered and independent corporation, the purposes of which were to increase the availability of resources and to diminish the rate of rise of per-unit cost of library service in its member institutions.

The latter part of 1967 and the majority of 1968 was, for the Center, a time of introspection and self-definition. During that period, Frederick Kilgour, the Center's director, traveled extensively, speaking with members in order to obtain the sense of the original membership regarding what services the Center would provide and what were to be the initial priorities for the establishment of these services.

The result of this effort was the determination that the Center would concern itself initially with implementing, through the use of advanced on-line computer technology, services which would augment and aid library personnel in classical areas of library service. Programming for information services which are new to librarianship was to be deferred until after programming had been activated for already existing areas.

The initial definition of areas of activity encompasses the concept of a tightly integrated, nonredundant system based on central usage of bibliographic data in an on-line catalog. These initial systems were to be:

1. shared cataloging,
2. remote catalog access and circulation control,
3. serials control,
4. a technical processing system, and
5. retrieval of information by subject approach.

The Center's staff was augmented in early 1969 by the addition of programming personnel. At that time the decision was made to provide, for an interim period, off-line catalog production to be based on a program previously developed at the Yale University Medical Library. Several man-years of effort had been expended in production of this system, and a considerable saving in time and effort could be realized by the transferral of this programming and its associated operating system from the Yale Direct Coupled 7094-7044 system to a stand-alone IBM 7094 system which was operational at the Ohio State University.

Concurrently with programming to adapt the Yale bibliographic system for usage in OCLC member libraries, an effort was begun to define and select necessary computer equipment for the Center's long-range use.

Ten major manufacturers of large computing equipment submitted proposals to the Center for equipment. However, because successful operation of a computer-based on-line bibliographic network had not yet taken place, neither the Center's personnel nor the computer manufacturers' personnel had a clear understanding of processor requirements for such tasks. Hence the Center could not rely upon others' experience as an aid in selection of equipment. Since the construction of a realistic bench mark test was effectively impossible under the circumstances, the Center sought another means for evaluation of proposals.

The evaluation technique which was selected was simulation. The Center choose the services of COMRESS, Inc. of Rockville, Maryland. COMRESS has, since 1962, been successful in the simulation of computers in various environments, including that of real-time interactive processing. The SCERT simulation program used by COMPRESS offers the advantage that file definition, record access, and data flows are defined in a fashion which is independent of the computer and operating system configuration on which this processing is to take place. These latter items are included as two small definition decks. Thus by the expedient of changing these two small decks one can quickly compare the relative merits of computers from various manufacturers employing different operating systems.

The central question to be answered by the simulation was: What machines, if any, are available to service an extremely heavy load of communications requests made against an extremely large data base, for a

population of users capable of supporting computer expenditures in the area of \$20,000 per month?

Each manufacturer was explicitly invited to view and review the results of his own machinery in simulation. The intent of this was to obtain each manufacturer's best thinking, and to afford each manufacturer an opportunity to challenge the results of the SCERT simulator, as applied to its machinery. Several manufacturers availed themselves of this opportunity, based on simulation results, and modified their proposals. One manufacturer challenged COMRESS's ability to properly simulate its equipment. This challenge was adequately rebutted by COMRESS analysts in joint meetings with the manufacturer's representatives and Center personnel.

An unfortunate initial result of these simulations was the determination that none of the proposed equipment could, using manufacturer-supplied operating systems, handle the projected system load. This came as a distinct and unpleasant surprise to Center personnel, since we had imagined that the central processing load in such a system would be relatively light, and that the principal loading would be found on communications controllers and disk files. Careful examination of simulation results revealed that a considerable proportion of the central processing unit load would arise from overhead functions in operating systems. This gave rise to the hope that improved operating systems for these machines could be devised—operating systems which would reflect efficient programming for the specific tasks which the Center's activities encompass, rather than generalized operating systems capable of responding to a virtually infinite variety of unforeseeable processing requirements.

The Center notified all manufacturers of these results and stated that an effort would be made, by modification of the so-called factor library of the SCERT modeling system, to devise nonexistent operating systems which did not exhibit half a dozen or so almost universal inefficiencies.

After the key inefficiencies were identified and removed from the models, simulations were again run for all the machines. The result of this was an almost across-the-board reduction by a factor of four in the central processing power required to sustain the peak traffic loading anticipated for the system. Three computers were identified as requiring a sufficiently low percentage of available central processing power to make the project feasible.

At this point other features of the machinery and the manufacturers' support and financing arrangements were examined. The Xerox Sigma-5 computer emerged from this latter evaluation as being the machine of overall choice as a vehicle for solving the problem stated above.

A Xerox Sigma-5 was placed on order in February 1970. Thereafter, programming to interface the Yale bibliographic system catalog card forming

program with Ohio State University computing environment, and to provide a conversion interface between the internal format of the Yale program and Library of Congress MARC distribution service tapes proceeded rapidly. By summer 1970, over one-third of the OCLC membership was activated in the off-line catalog production system.

In that system, MARC tapes were first preprocessed on an IBM 360/Model 75 computer and then passed to a conversion and file maintenance program which operated on that machine. This latter program (CONVERT) updated the master bibliographic file according to function codes contained in MARC records, and, at the same time, selected records to be processed for catalog production from the data stream. These selections were based on requests mailed to the Center by user libraries on 80-column punched cards. The output of this program was a seven-track tape containing bibliographic records in Yale internal format, and catalog card format instructions contained in tabular form. Using this data, the reprogrammed Yale formatting system, which now ran about ten times faster than it did upon its arrival at Columbus, produced catalog cards in one of 3,000 possible combinations of options. The formatted card images which were produced by this program were taken from the IBM 7094 computer back to the IBM 360/Model 75 to be sorted in filing order for specific catalogs. A final print tape was then produced on the IBM 360 before printing the card images on special forms, in "two-up" fashion, on the IBM 1403 printer, for which the Center had obtained a modified TN print train. Approximately one-half million cards were ultimately produced by this system, at an average cost of approximately 6.8¢ per card.

Users of the system were able to begin a smooth transition from wholly manually oriented processing to the on-line interactive processing system which was to come. At the same time Center personnel were able to gain invaluable experience in the exceedingly complicated programming necessary to produce individually custom-printed catalog cards for the catalogs of participating libraries.

As the initial off-line catalog production system was coming into being, Center personnel completed a survey of available CRT plus keyboard interactive terminals for use in the on-line shared catalog system. After considerable deliberation, it was decided that no terminal was sufficiently close to the requirement of library processing, and that the Center should participate in the design of an interactive terminal which had extensive text editing capability embodied in the processor of the terminal and which had a sufficiently large character-set to encompass the vast majority of processing (in romanized form) of foreign-language materials. Spiras Systems Inc. of Waltham, Massachusetts presented the most attractive opportunity for the production of such a terminal.

Close interaction between senior personnel at the Center and engineers of Spiras was begun in late 1970. Prototype terminals were made available to the Center for testing in the Spring of 1971.

The Sigma-5 computer system was delivered to the Center in September 1970 and accepted for operations in October of that year. Programming was then begun to modify the operating system of the Sigma along the lines suggested by previous simulation as being necessary to accomplish the objectives of a network. In parallel with this activity, Center programmers began work on the on-line catalog interaction system, and on the production of third generation versions of programs for the production of catalog cards. The flexibility of the catalog production system was greatly enhanced during this period, such that some 8,000 combinations of printing options came to be available for user libraries' catalogs.

The on-line system was made available in training mode to catalogers in member libraries during July 1971. In October 1971, input of new bibliographic data from terminals in member libraries began. By February 1972, a system which permitted the smooth interaction of both member-created records and those from the Library of Congress was active for all but a very small percentage of Library of Congress records. This latter percentage represented records which very probably were for titles for which bibliographic data had already been input from a terminal, but which were such that an automatic program verification of duplication could not then be made.

During 1972, advisory committees on technical processing and serials control met regularly and produced detailed requirement definitions for these two subsystems.

In February 1973, the Center activated programming to provide an extended search capability for those titles for which the normal truncated search key produces more than two screens of main entry-title-publication date personalized catalogs. (It is from these personalized miniature catalogs that the cataloger finally selects the bibliographic record on which to base her institution's cataloging.) This technique provides a very rapid indication of whether or not material the cataloger seeks is, in fact, contained in the on-line file of reference bibliographic information.

At this moment the monograph shared cataloging system is functioning smoothly and is, for all intents and purposes, nearly completed. It currently services more than ninety terminals, of which approximately seventy-five are located in member libraries in Ohio.

To date the Center has concluded cooperative agreements with a number of like-minded groups outside of Ohio, including NELINET, FAUL, PRLC, PALINET, and CCLC. When terminals in these regions have been activated, the number of on-line terminals will have been approximately doubled.

On March 26, 1973, the membership of OCLC agreed to extend the services of the Center to regional groups outside the state on one of three regular bases:

1. that a regional group already operates a computing facility;
2. that a regional group does not now operate a computing facility, but intends to do so; or
3. that a regional group simply wishes the Center to serve their constituency, using a computer located at Columbus.

Additionally, the membership instructed the Board of Trustees and the Center management to prepare a plan whereby the Center could conclude such agreements with individual large libraries. Currently the Center is actively negotiating contracts with a number of regional groups throughout the nation.

Programming to activate the serials control system will commence in June of 1973. It is anticipated that entry of bibliographic data for serials, check-in of journals, claiming for journals, and production of catalog cards for serials will all have been activated before the summer of 1974. The Center will begin adding personnel for the purpose of programming the acquisitions component of the on-line technical processing system, starting in late July 1973.

It is currently anticipated that system definition will begin in 1974, for activation to start in 1975 of the remote catalog access and circulation control component of the Center's integrated on-line bibliographic system.

Subsequent to that will come subject approach retrieval of information, interlibrary loan, and other information services which participants in the Center may require.