Implementing the New System: Conversion, Training and Scheduling

HILLIS L. GRIFFIN

WHEN ALL of the analysis has been done, and the studies have disclosed that it is time to implement a computer-based procedure to secure the desired objectives, several problems arise. One problem is that of building the data file which will form the basis for the new procedure. This may involve placing a substantial amount of information in machine-readable form, or of converting (and perhaps augmenting) an already existing machine-readable file. The staff must be trained to use the new system effectively and not to let it use them. Another problem is that of scheduling—how and when and where the jobs will be run, and how this schedule will mesh with the needs of the users of the system is another problem area. There is the story, for example, of the computer-based circulation system that was a winner in every way, except that the daily output always arrived a week late.

CONVERSION OF THE INFORMATION

The information upon which the new system is to operate must be assembled, put into machine-readable form, and validated. This may consist of translating the information in a single file of cards into machine-readable form, or of merging and selecting appropriate information from several manual files into the new data file. In any event, the information must somehow be transcribed (keyboarded or keypunched) into machine-readable form.

The source information for the new system may come from one or many files, and certain elements of information may be duplicated among them. The serial holdings of a library, for example, might be reflected in varying degrees of correctness in the binding file, a serials

Hillis L. Griffin is Director of Technical Information, Argonne National Laboratory, Argonne, Illinois.

check-in file, a shelflist, the official catalog, or a line-dex display. One file might have better bibliographic information, and another might have better holdings information. Experience shows that there will be less than perfect agreement among the files on any item of information, and that ultimately a physical inventory of the holdings may be required simply to validate the information which is assembled from the manual files. The problem then becomes one of finding the easiest way to gather the information upon which to base the inventory, and to use as much of this information as possible in the final file.

The input equipment available to the user will strongly influence strategy in converting the data to machine-readable form. The best method will be that which provides the highest input rate with the greatest accuracy at the least cost, highest reliability, greatest ease of operation, and with a character-set which satisfies the job requirements in the most reasonable manner. Among the devices to consider are the keypunch, optical character recognition (OCR) equipment, intelligent and unintelligent terminals, typewriters which produce some machine-readable output (magnetic cards, magnetic tape, punched paper tape, punched cards), and others. Mark-sense and magnetic ink character recognition (MICR) are of such limited utility for bibliographic data that they need not be considered.

The keypunch is the most commonly available device. It produces punched cards which can be read by almost any computer. The IBM 029 printing keypunch has a substantial repertoire of special characters (punctuation) and, while it punches only upper-case letters directly, can rather easily represent both upper- and lower-case by use of a "shift code" to signify upper-case and nonstandard characters. In practice this limitation poses no great difficulty. Good design of the input format of the punched cards, together with a suitable computer program to read and reconstruct the input information and to provide validation listings, can furnish a fast input medium for the smallest installation at minimum cost. The keypunch is not a very sophisticated device, but it is completely independent of computer and telephone line difficulties. It functions with an extremely high degree of reliability and it can withstand a great deal of punishment.

OCR input is prepared by typing the information on ordinary paper using a typewriter whose type font, character and line spacing are compatible with the OCR input reader to be used. The typewriter is usually an IBM Selectric with an OCR typeball and proper line and character spacing. The user must be certain that the character-set to be used (including lower-case and special characters) is available on the typeball and can be recognized by the OCR reader. The OCR reader is connected to the computer and, like any other input medium, requires a computer program for its operation. OCR readers are not inexpensive, and are not commonly a part of a computer system. An advantage of OCR input is that keying errors are easily corrected, and that less training is required for typists to use the system than for other input devices. If the equipment is not already available, most libraries by themselves will not have enough work to justify the expense of adding OCR equipment to the computer system they will be using. Like the keypunch, the OCR input is read into the computer in a batch rather than an on-line mode. The keyboard devices are always usable regardless of any computer or telephone line problems.

The "unintelligent" terminal is generally a typewriter connected by a telephone line to a computer across the street or across the country. A typist types information on the typewriter and it is transmitted, line by line, to the computer, which stores the information on secondary storage (disc or tape) for later use. The terminal user can interact with the computer program to edit, validate, or search the information file. Various text-processing systems are available, among them ATS (Administrative Terminal System), Wilbur, Supercomp, and DataText, and many others are available on a service bureau basis, or on the user's own computer. Such systems are very "typewriter-like" and require relatively little operator training for keying, although the operator must learn a repertoire of simple commands to operate the system and interact with the program. Input is not as fast as with the typewriter, and delays of 5 to 10 seconds per line may be encountered as system saturation occurs. Editing is somewhat clumsy because it is oriented to lines, rather than to sentences or paragraphs and it takes some time to type sentences back for editing purposes. Video terminals (a television tube display and associated keyboard) improve the visibility of data by displaying several lines at once, but many of these too have difficulties in accommodating changes which involve more than a single line of text. Telephone line or computer difficulties make the terminal unusable until they are resolved. Costs of these systems typically involve rental of the terminal, equipment for attaching the device to the telephone line, telephone line charges while the system is active, costs for special telephone lines (if the ordinary lines are too "noisy" for normal operation), computer time charges for time hooked up to the system, and charges for data storage.

The "intelligent" terminal is typically a video terminal with a small computer inside. It may operate under an editing program similar to ATS, or under a user-provided editing program which is tailored to the application. The data may be written on a cassette of magnetic tape (like those used on small cassette recorders), or the terminal may interact with a large computer much as does the "unintelligent" terminal. If the information is gathered on cassettes, the terminal is completely independent of the large computer, and is quite portable. It can, however, communicate with the large computer via telephone line to transfer the collected information from the cassettes to some other storage medium associated with the large machine. In this case the terminal requires only computer and telephone line time sufficient to transfer the accumulated information, which is much less than the time taken to accumulate it. This leads to some saving in telephone line and computer time charges and gives a greater degree of independence from the large system as well. The video terminal has the inherent capability of rapid response since there are no moving parts required to display characters, and it can normally display an entire bibliographic entry in a single screen—a benefit for validation. Unfortunately, there are presently only a few video terminals which are both reasonably priced and satisfactory for text processing. These are adequate, however, and new and better equipment is rapidly being made available.

A more powerful extension of the intelligent terminal is the combination of an unintelligent video terminal (or several of them) with a minicomputer, disc storage device, tape drive, line printer, and appropriate telephone line adaptors. Such a system can accumulate data onto normal 7- or 9-track magnetic tapes or onto cassettes or disc for later transfer to 7- or 9-track tape, can send or receive data across high-speed telephone lines, and can print validation (or other listings) on a medium-speed printer, as can many intelligent animals. Such a system can handle several terminals simultaneously, and it is in this situation that it is most economical. This is the basis of the commercial key-to-tape and key-to-disc systems which are popular replacements for keypunch systems. Unfortunately, most of the commercial key-to-tape or key-to-disc systems have input devices which do not display the results of the input to the operator (much like a typewriter which does not print what has been typed).

The typewriter which produces some machine-readable output should also be considered. There are several devices of this kind, each producing a different product. The IBM MT/ST (magnetic tape Selectric typewriter) produces a magnetic tape cassette which can be converted by appropriate devices to the usual magnetic computer tape or used directly as computer input with the MT/ST cassette reader that attaches directly to an IBM system 360/370 computer. The reader is not priced for the low volume user, however. Another device for off-line conversion to magnetic computer tape is the Digi-Data converter. Both the Wang Laboratories and Texas Instruments Co. have announced keyboard devices which produce hard copy and cassette tapes. The Wang typewriter is rather like the MT/ST, with editing capabilities, plus the ability to transmit over telephone lines. The Texas Instruments machine produces copy on a special paper with a thermal (rather than an impact) printer. It is priced considerably lower than almost any other comparable device. Also available is the IBM MC/ST (magnetic card Selectric typewriter) which records data on the edge of fanfold cards with a magnetic coating. It also has an associated data-transmission device.

Also available are punched paper tape typewriters, which record each key-stroke on punched paper tape. They require a device either to convert the punched paper tape to punched cards or magnetic tape off-line, or to read the punched paper tape into the computer. While punched paper tape readers are often found on very small computers, these computers generally cannot produce magnetic tape or punched card output. The larger computer systems seldom have a punched paper tape reader as an associated input device. Furthermore, punched paper tape is difficult to correct. So, considering the various difficulties in using it and the cost of input devices to get the information into the computer, there are easier, cheaper, and more reliable ways than punched paper tape to do the job.

Which one should be selected? The choice is often made simply by the equipment which is locally available. If the computer available does not have OCR equipment, then one will probably not be able to justify its addition to the system solely for the library. The computer available also may not be of a type capable of supporting on-line terminals. If there is no money for equipment, perhaps time can be stolen on a locally available keypunch, and if there is only a little money, renting a keypunch is an option. With access to a data entry system via telephone line, another alternative is open. At some point, line and connect-time charges will make an intelligent terminal very attractive. Another alternative is to have the job done by a service bureau of some sort. This imposes a new set of decisions to be discussed later.

INPUT DATA PREPARATION

Next comes the task of putting the input information in some form which is usable for the keyboarding operation. The traditional method is the worksheet. This is a sheet, filled with squares, onto which the data is manually transcribed in block letters so that the typist or keyboard operator can then key the information. Another expedient is to type worksheets which can be used by the input clerk for data entry.

One advantage of doing the conversion job within the library is the familiarity the library staff has with the information. Ever since the acceptance of the typewriter as a piece of library equipment, clerks have been transcribing catalog cards on the typewriter from hand-written copy (written on the back of a proofsheet), or a "revised" proofsheet or LC cards. Through some miracle the clerk is able to transcribe this information onto a typed card in the proper format. But somehow, when people begin to deal with computer input, they suddenly become more formal. If one considers that keying information is infinitely faster than writing it in squares in block letters with a hard lead pencil, then it is reasonable to refrain from filling in worksheets by hand. If they must be typed, why not just key the information into whatever is to be used for computer input, rather than prepare a worksheet? It is faster and more efficient, and helps to reduce the cost of labor for conversion. Since labor will be the major cost in conversion of a system, every possible way should be sought to conserve labor.

One of the traditional reasons for worksheets is to allow the keypunch operator to key-verify the information that has been punched. This process consists of re-keying the information on a verifier machine to see that the holes in the card agree with the characters on the worksheet. This involves two keying steps. Key verification is certainly the best way to validate typical accounting data—account numbers, part numbers, quantities, hours worked—all difficult to proofread visually. But bibliographic data can be proofread more rapidly than it can be re-keyed.

The conversion then might well be done within the library, using library staff to do the keying, and using raw data directly from the library records, rather than worksheets. Given a good input format (on the keypunch) or a "fill-in" interrogative program (on a terminal system) the clerk should be able to enter the information rapidly, even using more than one file to supply the required information. After a short time, the operator will be able to proceed more rapidly without the prompting of the "fill-in" program on the terminal system.

The computer can be of considerable assistance in proofreading. It can place the input in sequence, print a prooflist which will display the information in a form which makes it easy to use, and perhaps also in the input format as well. But the program can do much more. In the case of bibliographic information, for example, it can detect many of the mechanical errors in the input and let the proofreader check for errors in spelling, etc. Mechanical errors include such things as a missing call number, an all numeric LC classification number or Cutter number, a mixed alphanumeric Dewey classification number, a publication date later than this year, or earlier than, for example, 1900 (or whatever is chosen), the absence of the main entry or the title, or the appearance of two titles (one of which probably should have been the *added* title), etc., there are many things that the program can look for and identify. The proofreading program should make things *easy* for the proofreader, because this will increase the accuracy of the proofreading task.

If manual files are at hopeless odds, what can be done? One method is to convert the best file to machine-readable form, and use a computer-produced worklist to note corrections, additions and changes. The worklist can be split into sections so that several people can work simultaneously on different parts of the alphabet, and with different files at different places. The revised worklist forms an excellent basis for entering the accumulated changes and corrections, and assures that they will not be lost or duplicated. The wrong way to do this is to take time to type a work file on 3 by 5 cards and use this file to resolve problems. The keystrokes could have been better invested preparing the actual input to the machine-readable system.

It might appear that the easy way out of the travail of file conversion would be to farm the job out to some enterprising and utterly reliable contractor who would, by some magic, and without removing records from the library, effect the file conversion. There are, for example, contractors who will microfilm your input and send the film overseas, where it will be keyed and returned on magnetic tape. Costs are generally much lower than for domestic labor, but turnaround time must be allowed.

There are also domestic service bureaus which will accept conversion work to be done on keypunch, terminal, or OCR devices. There is also the campus data processing center which may agree to keyboard your records on a "time available" basis. This usually turns into a bitter and frustrating experience for all concerned.

All of these schemes assume that the data is in a format which will allow it to be keyed as it stands. If it is not, then the data will have to be transcribed to worksheets which can be used by the keyboard operators. If the library has to do the transcription, then it might be better to hire and train some temporary help, or begin training the affected staff to do their own keyboarding. Since nobody is as interested in a project as those it directly affects, the best way to guarantee its success is to give it close and sympathetic supervision by people who are interested in and responsible for its success, and who will have to live with it later.

STAFF TRAINING

If the new system is to work properly, the staff must acquire new skills to learn how to use it. The conversion process may require training of staff even before the final system is operational, but this gives them a good opportunity to participate in the development and growth of the system, and for their skills to grow with it. The conversion period is generally a more relaxed time so that learning can proceed without undue pressure.

Those who are directly involved in entering information into the system will need to learn how to operate the keypunches, or the terminals. Keypunch instruction using programmed courses commercially available on magnetic tape has been very successful. A course of instruction covering 2½ days turns out excellent keypunchers. Terminal systems often have a tutorial mode to indoctrinate new users. In any system the machines are the easy part—the knowledge of the library and its ways comes harder. Mastery of the mechanical procedures will enable the operator to give full attention to the protocol of entering information—the format, the rules, the procedures. The best system will be one which conserves labor, maximizes speed of throughput and minimizes errors—that is, one which puts most of the burdens on the computer rather than on the input operator.

The people who work directly with the system need to know other things about it as well. They should have some idea of its goals, how it does its job, and their part in all of this. And, finally, they should actually watch the computer at work. No matter how well the process of catalog card production on the computer is described—the speed of input, of sorting, of alphabetizing the records (the reverse of the manual procedure) and then of printing—nothing is quite as awe-provoking or as ultimately descriptive as seeing the computer do it.

The new users also need to know how to recognize problems and solve them. Any new system has problems; this should be admitted to the staff along with the information that they are part of the solution. Key punching errors that are not corrected threaten the reliability of the system. So do programming errors, and both should be corrected without delay. The people actually using the system will suggest additions, revisions, or improvements, and they should be heard and heeded. If the problems are not resolved the users suffer, the system is not effective, and the library is the ultimate loser.

Training in programming and systems analysis should not be a library concern. Programming classes are offered on many campuses. The problem will be in applying the language, usually presented in a mathematical (FORTRAN, PL/1) or business (COBOL) context, or as an entity in itself (assembly language) to the library and its very diverse needs. It is unfortunate but true that the best environment for library programming is within the library by programmer-analysts who have had training and appropriate experiences as librarians. They know why libraries do the peculiar things they do, and they have some interest in doing things better. If they are on the library staff, they are visible and face the prospect of living with satisfied or dissatisfied users —a good incentive to make the system work well.

The training program is probably best operated by the library itself. It will be a continuing responsibility because of personnel changes and systems changes. To direct the training program, oversee the conversion and handle day-to-day problems, an articulate, understanding person who knows how to use the equipment, understands the system, and can communicate this knowledge to the library staff in their own language will be invaluable.

Levels of training are related to the needs of the job, and the skills of those selected for training. Experienced keypunch operators will often need to learn new techniques to cope with the large amount of free-form alphabetic information. Typists will find out very quickly that you do not backspace and "x-out" mistakes on the keypunch. Given a choice, the best candidates will be those from the library with adequate keyboard skills who are alert and able to proceed independently.

SCHEDULING

The library will need some appropriate guarantees that its jobs will be run on the computer in a scheduled and timely manner. In the case of an on-line system this also means that the appropriate library files will have to be on-line (and consuming some substantial part of the system resources) when they are needed. The on-line circulation system is a case in point; it cannot be on-line only on alternate Tuesdays from 4 to 5 in the afternoon. Batch operations also need to operate on schedule if they are to be reliable and useful.

Most libraries will find it satisfactory to have their batch processing accomplished at night or on weekends, often at less cost. In this way they can fully utilize normal working hours to prepare input to the system, making it current to the close of the day. The output, run that night and returned the next morning, will be current. But it must come back the next morning if people are depending on it to do their jobs.

Long computer runs may be scheduled for an otherwise slack period in the computer time schedule. The work load in a business-oriented computer center is often cyclic and there are also cycles in a scientific computer center. Such scheduling requires planning and coordination with the computer center, but will assure prompt attention to these jobs.

This work was performed under the auspices of the United States Atomic Energy Commission.