A Quarter Century of Advanced Data Processing in the University Library

Jim Segesta and Rod Hersberger

Advanced Data Processing in the University Library, by Louis Schultheiss, Don Culbertson, and Edward Heiliger, the first book about computers in libraries, was published a quarter century ago. It tried to show how computers might be used in libraries and how libraries might change from conventional to computerized operations. It presented its view of library computerization more fully and systematically than anything published previously. This retrospective review attempts to consider how accurate that view was.

A quarter century ago, in 1962, the rent for an IBM 1401 was $105 per hour or $5,500 per month, and all computers were mainframes that came with programmers, not software. The word software was so novel that its use required an explanation. Digital Equipment Corporation was still a year or two away from producing the first minicomputer, the PDP 8. Innovative libraries were acquiring Xerox machines; patrons were learning to love the convenience of copying whole pages mechanically, and the model 914 showed promise as a way to make catalog cards. In that year, Advanced Data Processing in the University Library, by Louis A. Schultheiss, Don S. Culbertson, and Edward M. Heiliger, was published by Scarecrow Press. It was the first book specifically devoted to the use of computers in libraries. This paper is a retrospective review of that book, which will be referred to here as ADP. It seems appropriate after a quarter century to take another look at the first book about computers in libraries. Rereading ADP leads one to consider to what extent librarians’ initial expectations for the computer have been met and, where the expectations differed from the outcomes, to speculate about the reasons.

Since ADP was published, we have never lacked books about computers in libraries. ADP was like the bugle call that signals a cavalry charge. The next year, 1963, saw the publication of the King report, Automation and the Library of Congress. The first of the long series of Clinics on Library Applications of Data Processing was held at the University of Illinois in Urbana. Also in 1963 the Airlie Foundation Conference on Libraries and Automation was held, the proceedings of which were published in 1964. In that year Medlars became operational, ERIC was founded, and the first volume of the American Society for Information Science’s Proceedings of the Annual Meeting was published. The literature of the library’s computer age has poured forth ever since.

Library automation (as opposed to computerization) was no novelty in 1962. Punched card and paper tape technology had been pressed to fairly high states of development in a few libraries. More than a decade earlier, in 1951, Kings County Public Library in Washington had produced a book catalog using punched card
equipment. By 1955 the Los Angeles County Public Library had published a book catalog listing 161,000 titles in twenty-four volumes, and by 1960 there were more than fifteen library book catalogs produced by punched card equipment. At the University of Missouri punched cards were used to prepare orders, and catalog cards were produced on Friden Flexowriters, which were electric typewriters driven by a punched paper tape. If a typist at a Flexowriter typed one main entry card complete with tracings, the machine could then produce an entire set of properly headed cards.

In a related area, information retrieval systems using computers had been under development for almost a decade. As early as 1954 in, oddly, the Mojave Desert, one of the earliest electronic IBM computers had been harnessed to a Uniterm file for the world’s first bibliographic search by computer. This was the line of development that led over the years to the computerized databases and to Dialog, BRS, and others. Even in 1962 the computerization of information retrieval was understood to be related to, but separate from, the computerization of libraries.

"ADP was the first systematic, comprehensive, published description of how computers might be used in libraries."
Clearly, the book looked good to librarians in the early sixties. Events moved so fast, however, that by the end of the decade few bibliographies on library automation included *ADP*. It was influential, but only briefly; a flood of writings describing more sophisticated systems and more advanced technology quickly replaced it. Too, the accumulating experience of libraries with computers revealed some problems with the book’s analysis of the computer’s role in libraries, and at Florida Atlantic University what was understood to be an attempt to implement the system proposed in *ADP* failed, with no good results for the book’s reputation.

*ADP* came into being in this way. Edward Heiliger, the director of the University of Illinois’ Chicago Circle (UICC) campus library, had an early interest in the application of computers to library work. In 1959 he invited Cloyd Dake Gull of General Electric to speak to some Chicago librarians about computers in libraries, and that meeting led to a project to create a flowchart of the UICC library’s paperwork systems, which led in turn to an application in October 1960 to the Council on Library Resources (CLR) for a grant “to investigate the possibilities of a total system of mechanization of routines in a university library.” CLR promptly approved a $50,000 grant: *ADP* is the record of the project funded by that grant.

The UICC librarians had more reason than most to investigate the computer. New libraries, now rare, were common then. In 1962, the year *ADP* was published, librarians in California, for example, were beginning to assemble opening-day collections for no less than three new University of California (UC) campuses and two California State University campuses. While the UICC campus was not new, it had a small library that would grow very fast. Any librarian in that situation at that particular time was duty-bound to consider whether the computer offered superior alternatives to the traditional ways of building libraries. The potential gain was huge: if the computer could be used from the beginning, the mammoth files of a large library could be created in a machine-readable form. There would never be a need, in that library, to convert paper records to electronic forms. Naturally, librarians at such new academic libraries as UICC, UC San Diego, and Florida Atlantic were among the pioneers in applying the computer to the work of the library.

The book is not easy to describe; as noted above, Wasserman, who admired *ADP*, called it a “mélange.” *ADP* is both a plan and a description of the process that produced the plan. The product was a paper, not an operating system. The book is very clear about that. (Later in the sixties, publications that described plans as if operational appeared frequently enough to draw complaints.7 No charge of that sort could be leveled against *ADP*.) Further, *ADP* was a team effort involving the authors, other UICC librarians, and the consultants from General Electric, Gull, and G. P. Williams. It is hard to tell who contributed what. The consultants were responsible for much of the content of the key chapter that described the proposed systems.

The planning process was better than the plan. The book could still be consulted with profit as a guide to planning an innovation in a library. UICC librarians began with first principles: the first chapter reviewed the history of library automation to that time, and the second chapter was entitled “A Machine Age Library Philosophy.”

The philosophy could almost serve today. Very briefly, the librarians wanted a new system that would provide all the information then available through conventional card catalogs and order, serial, and circulation files. In addition they wanted current awareness lists and subject lists compiled and printed by the computer. They understood that computers would be able to combine and manipulate subject headings and other parameters in ways that were and would always be impossible with conventional files.

On the other hand, they did not seek subject indexing in more depth than was then provided by the Library of Congress and the standard indexing services. They also waived the mechanical retrieval of in-
formation and understood that, in the long run, Library of Congress cataloging in machine-readable form would set standards for format. The final point in the UICC philosophy was that centralized collections would serve campuses better than networks of departmental libraries. Although not mentioned in the statement of philosophy, costs were very much on their minds; the authors hoped that the computer could reduce or contain the costs of the technical services, and that was cited repeatedly as a justification for library computerization. Finally, throughout the book the authors described their method as a "total systems approach." Allowing for the differences between batch processing and online computing, "total systems" means approximately the same as integrated systems.

Another thing well planned was staff involvement—the authors seemed very sensitive to the concerns of their co-workers. Regular meetings with the professional staff were held before and throughout the project. One appendix includes reprinted essays that the UICC librarians had written on their expectations for the project. At one stage the librarians even debated a list of questions about basic library procedures: "Resolved, we should do away with all fines"; "Resolved, each faculty department needs a catalog," etc. The staff appears to have been enthusiastic about the project. In their preface to ADP the consultants, Gull and Williams, presumably encouraged by their experience at UICC, ventured to predict, "the stereotype of librarians as conservatives opposed to technological change will vanish whenever librarians are faced with real opportunities to introduce sound technological improvements." 16

A chapter "Present Methods" analyzed unit costs for acquisitions, serials, circulation, and cataloging for the year 1959–60. The work load was small—2,549 titles cataloged, for instance—and the unit costs were high: the authors estimated $13.51 to acquire, catalog, and process one book at UICC. 9 Reviewers were struck by those figures; from England, C. A. Crossley grumped that his "American colleagues have become absorbed in routine to such an extent that they fail to see that machine and computer techniques are not the only way out of the wood—they could cut down some trees." 10

Reviewers were also impressed, but more happily, by the flowcharting—ADP contained almost 100 pages of flowcharts. The technique was not widely known among librarians then, but the reviewers could see how useful it would be in describing the operations of technical services. As noted above, Schultheiss delivered a paper on the subject the next year at the first Clinic on Library Applications of Data Processing at Urbana. 11

Chapter 7, "Proposed Systems," is the heart of the book. Taken from the consultants' report, it describes and estimates costs for an integrated system whose main files would be in book form. The processing information list, to be printed weekly, would record the books on order and in cataloging. When books were ready for use, their records would be transferred from the information list to the monthly edition (supplement) of the catalog. This edition would be revised monthly for a year or more; then its records would all go into the total holdings edition of the catalog. There would thus be only two alphabets to search in order to determine the library's holdings of cataloged books and a third to search for the books on order and in process. The costs for producing those two lists were estimated at $33,000 per year at a work load of 2,000 titles per month, and that excluded the cost of LC copy and keypunching. Arranging the individual records was a problem. As the authors noted, "It is easy enough to say that a computer will put an unordered list in order by author or by title. However, the computer program for doing this does not exist." 12 Sorting codes were a possible solution. The difficulty was certainly not insurmountable, but four out of the authors' list of eight tasks to be accomplished before an operating library could be fully automated involved filing and alphabetization. 13

At the time, many false trails beckoned library innovators. In the late fifties and early sixties various mechanical devices had been mated with microfilm databases
to provide a kind of automated retrieval of information. Systems like the Rapid Selector and Walnut received a good deal of attention; the authors of *ADP* considered but quite properly dismissed them. Facsimile transmission was more exciting. They visualized the transmission, over coaxial cable or by microwave, of pages from books and journals in Urbana. There is no mention, though, of transmissions between computers and remote terminals linked by telephone lines. Within a decade OCLC had made the MARC records accessible to midwestern libraries using that technology, and the consequences were revolutionary; not surprisingly, there is no glimmer of that development in *ADP*.

"The authors tried to show how computers might be employed in libraries and how libraries might change from conventional to computerized operations."

The authors tried to show how computers might be employed in libraries and how libraries might change from conventional to computerized operations. They presented their view more fully and systematically than anyone who had published previously. How accurate was that view?

One assumption was that a small, new library was a good place to start computerizing. This assumption depended on another: that the research and development necessary to produce better or cheaper alternatives to conventional methods were not beyond the reach of such a library, given reasonable funding. Tied up in the same bundle of assumptions was the notion that computerization was something that each library would have to develop for itself. (That had been the case with the punched card systems of precomputer library automation.)

Later in the sixties hundreds of libraries, many of them small or new, did make moves toward computerization. In 1970 the Lark Association surveyed 3,000 libraries about their automation projects; 347 reported projects either planned or operating. Many of these projects were small, single-purpose applications; few, if any, were as comprehensive as the system proposed in *ADP*; the more ambitious were trouble-prone. The library landscape of the late sixties was pock-marked with crashed and ruinous automation projects. Libraries that should never have ventured so far into automation at that stage of development wasted some sizable sums. Pioneers are more likely than other people to leave either their names on maps or their scalps on teepees, and never was that truer for library pioneers than in the automation efforts of the sixties.

The CLR’s *Fourteenth Annual Report*, in 1970, stated the lesson plainly: “It has become increasingly evident that the average library will never be able to ‘go it alone’ in some aspects of the new technology—automation for example. The level of investment required to reap the benefits of the emerging national machine-readable databases exemplified by MARC is far beyond the individual budgeting capacity of any but the very largest libraries. Agreement is growing that the only possible solution to the dilemma—especially for the medium-sized and small libraries—is for them to band together in local, state or regional consortia and thus pool their assets and efforts.”

One might wonder if the book, too optimistic, had encouraged librarians to enter the water when it would have been better to warn them away.

Other, much grander appraisals of the cost of automation appeared shortly. The very next year, 1963, the King report estimated $50 to $70 million for the automation of the Library of Congress. By the end of the decade a proposal for five national computer projects spoke of $400-$500 million over a four-year period, over and above all present funding.

Not only were the research and development to prove much costlier than anyone dreamed in 1962, but when the money came it flowed through doubtless well-worn channels to the largest universities, just like the funding for research in other fields. Writing in 1973, a decade after
ADP, Herman Fussler reported that Columbia University had spent $1,105,000 on automation from 1966 to 1970; that Stanford had grants for library automation of $1.2 million from the Office of Education, $1.13 million from NSF, and $650,000 from CLR and NEH; and that the University of Chicago had received almost $2 million in grants for library automation in the late sixties and early seventies. As is to be expected with research and development, many very expensive projects were failures. Among the more successful results of this research investment are NOTIS from Northwestern, VTLS from Virginia Tech, and Stanford’s BALLOTS, a forerunner of RLG’s RLIN. Although many small and new libraries experimented with computerization, the important developments took place elsewhere.

ADP also cried the second coming of the book catalog—book catalogs printed by computing equipment and supplemented regularly were to replace card catalogs. That, too, was, on the whole, an error. In their rematch the card catalog again bested its rival, but the book catalog did win a few rounds. Computer-printed book catalogs have superseded card catalogs in many county libraries and in other locations where multiple copies of a book or COM catalog could replace a number of separate card catalogs. The book catalog’s little cousin, the serials holdings list produced by computer, is a library staple, and the online catalog may well become the premier development of library computerization.

Interestingly, the King report, published a year after ADP in 1963, contains an appendix dated August 1962, which describes and proposes an online catalog. This appendix, “A Cost Analysis of an Automated System for the Library of Congress,” by Herbert T. Spiro and Allan D. Kotin of the Planning Research Corporation of Los Angeles, was actually the largest part of the King report. The concept of Spiro and Kotin was quite like that in UC’s Melvyl and in the OLPACs, which in 1987 seem certain to become the catalogs of the future. The two potential products of the cataloger’s computer, that is, the book catalog and the online catalog, were thus opposed at the very beginning of library automation. Why did the Library of Congress consultants recommend an online catalog while those at the University of Illinois recommended a book catalog? Money was undoubtedly a factor, and so was time. The King report mentioned $50 million (in 1963 dollars) as a reasonable price for automating the Library of Congress, and spoke of the research and development necessary to perfect the terminals that would be used to search the computerized catalog. That money and effort were out of the question for any other library. In Chicago the UICC librarians faced the prospect of an imminent acceleration in growth. If the computer were to help, it would not be through the development of a computerized catalog. The large sums of money and the lengthy period of research and development required were too costly. The UICC librarians and their consultants looked instead at batch processing and fast printing and made their choice accordingly.

“The total systems approach is perhaps the most enduring of all the concepts in ADP.”

The total systems approach is perhaps the most enduring of all the concepts in ADP; the concept was and is appealing. A record created at the time a book was selected for acquisition would roll through the technical services like a growing snowball, acquiring buying information first, cataloging information later, and circulation records ultimately. Creating the record would be like assembling a machine moving along an assembly line, parts being added at each stop. In ADP’s plan the record would first appear in the processing information list, a printed list of books on order and in process. The list would be revised weekly, adding new records and adding new information to existing records. When the cataloging information had been added, the record would be transferred to the next issue of the monthly edition (the cumulating supple-
ment) of the printed book catalog. Later the record would appear in the total holdings edition of the book catalog. The daily circulation list would be a brief listing of all books not on the shelf; it would be produced from punched book cards removed when the books were borrowed.

Total systems have evolved into integrated systems, and those are coming slowly—a quarter century after ADP, libraries are only on the threshold of integrated systems. The systems we seem likely to implement in the next few years are conceptually much like that proposed in ADP, except that today’s systems provide or allow for interfaces with the book trade, the Library of Congress, and the interlibrary loan networks. Technically the difference is that the modern systems are online. Just as in ADP’s plan, technical processing and circulation are included in the integrated systems, while the reference department is a user of both the library’s integrated system and the bibliographic services supplied by vendors.

The one criticism that might be made of ADP’s version of the total systems approach is that it supposed libraries to be much more independent than they really are. In actuality the main departments of a library, except circulation, are themselves parts of larger systems outside the library. Acquisition departments are bound up in the book trade; catalog departments are tied to the sources of their cataloging data, principally the Library of Congress; and reference departments depend on their “toolmakers,” the producers of the reference books and bibliographies, printed or machine-readable. The big payoffs in library computerization so far (again excepting circulation) have come from systems that have computerized the links between the cataloging and reference departments and the suppliers of the information in which they deal. The computer lets catalogers and reference librarians tap the collections of data that they use in their work more effectively. The computer could and should do the same for acquisitions librarians, but that particular development has sadly lagged.

Computerization has surely dramatized and highlighted these connections between the departments of the library and the outside world with which the departments interface, but the links should have been visible enough in 1962. Why did ADP tend to overlook or at least minimize them? Why did the reviewers see no problems in the absence of connections between the library’s computer and the suppliers of the library’s data? One reason, no doubt, was the precedent of precomputer automation. The punched tape and punched card systems of the time were all developed individually, each library working on its own. Too, the long-distance transmission of data that allows computers to be consulted from great distances was still in an experimental stage.

Another possible contributing factor is that special libraries were a misleading model. In the Proceedings of the first Clinic on Library Applications of Data Processing in 1963, Burton Adkinson praised the total systems approach to library automation:

Thus machine records produced in one operation, acquisitions for example, can be used in others such as cataloging, circulation control, announcements, selective dissemination, or information retrieval. Examples of activities using this approach are National Reactor Testing Station, Lockheed Missiles, Douglas Aircraft, and Sandia Corporation. Certainly this systems approach is good.

It is significant that his examples of successful applications for the total systems approach were all special libraries. Several factors tended to propel special libraries into computerization earlier than academic or public libraries, so that anyone looking for examples or models of computer applications in libraries in the early sixties would need to look at special libraries. But it is characteristic of special libraries that they must develop their own sources for acquisitions, their own cataloging, and their own homemade files and indexes. Special libraries depend less on the book trade, the Library of Congress, and the publishers of reference books and bibliographies than do academic or public libraries. To whatever extent special libraries were the models for the total systems approach espoused in ADP, a correction should have been made to take into
account the differences between them and academic libraries.

The authors of ADP recognized and avoided certain dead ends that looked promising to some at that time. They understood that computerized information retrieval would be developed outside the library. At the time there were a number of new products based on hybrid technologies mixing microforms and computers or punched cards; the authors realized that these were not in the mainstream of library development. They also saw that commercial processing, then drawing considerable attention for several reasons, was not to be significant in the future of academic and public libraries; they understood, too, that expedients like the keyword indexing of titles would not replace subject cataloging and that "semiautomated" transitions, approaches to computerization via punched card technology, were not the right path to follow.

ADP was perhaps the earliest published attempt to foresee, comprehensively and in detail, the changes that computers would bring to libraries. Its view, because of the authors' situation, was deliberately short-range: its message was "Let's start—now!" Many librarians of the time welcomed this message and took it to heart. The book impressed its readers, who were the library leaders of that time. Some of its ideas stood the test of time rather well, others did not. No one should be surprised that in many ways the computerization of libraries developed differently than projected in ADP. The power of the minicomputer and the pervasiveness of the microcomputer today, for instance, were unimaginable in 1962. It is interesting, after a quarter century, to try to account for the divergences and to identify the circumstances then operable that led librarians to expect developments that never materialized and to minimize circumstances that were to become very important. That is what this paper has attempted. To keep a fair perspective in judging a book written twenty-five years ago at the very inception of a new technology, it is necessary only to turn in the other direction and attempt to imagine how matters will stand a quarter century hence.

REFERENCES AND NOTES

9. Ibid., p.104.
19. Charles M. Goldstein, in "Integrated Library Systems," *Bulletin of the Medical Library Association* 71, no.3:308-11, credits Dr. Fussler and the University of Chicago Library, ca. 1965, as the principal progenitors of integrated library systems. While ADP used the phrase "total systems," the concepts seem similar.
20. The Library of Congress, monolithic and governmental, moved with what now seems admirable celerity in the sixties to produce MARC. OCLC and the other intermediaries immediately developed the necessary links with the consumer libraries. The result has been a revolution in cataloging procedures. The publishers of indexes each independently developed machine-readable versions of their bibliographies, which BRS, Dialog, and SDC have exploited for the common good. Such systems have been operational for ten or fifteen years. The book trade, on the other hand, entrepreneurial and multifarious, has never collaborated successfully in automation. After a quarter century of automation it should be possible to identify any recent book at a terminal, receive a display of the latest price and availability information, add local ordering information, and direct the resultant request to wholesaler A with instructions to the computer that if that firm cannot supply the book within three days, it should transmit the order to wholesaler B and thence, if necessary, to the publisher. Of course, it is impossible to do this now, even though the same operation has been routine for years in interlibrary loan. Could anyone, twenty-five years ago, have predicted success for automation in cataloging and bibliography, failure in acquisitions?
Of course, both the Library of Congress and the publishers of the bibliographies had enormous databases that had to be keyboarded and printed before the computer came along. The transition to machine-readable records was easier because the computer could help with the printing.
22. Schultheiss, Culbertson, and Heiliger, p.33.
23. Ibid., p.34.