January 1984

Bit by Bit: Microcomputer Applications by Archivists in Four Southeastern States

Glen McAninch

University of Kentucky

Follow this and additional works at: https://digitalcommons.kennesaw.edu/provenance

Part of the Archival Science Commons

Recommended Citation
Available at: https://digitalcommons.kennesaw.edu/provenance/vol2/iss1/4

This Article is brought to you for free and open access by DigitalCommons@Kennesaw State University. It has been accepted for inclusion in Provenance, Journal of the Society of Georgia Archivists by an authorized editor of DigitalCommons@Kennesaw State University. For more information, please contact digitalcommons@kennesaw.edu.
As archivists in Georgia, Tennessee, North Carolina, and Kentucky began to automate in the early 1980s, most found microcomputers much more to their liking than main-frame computer systems or book-oriented network systems. 1 Few archivists in the region were using computers before microcomputers were developed in the mid-1970s. The smaller computers that were marketed during this period allowed users in the region to adapt programs easily to their individual needs at minimum cost. However, the limited capacities of the first microcomputers have pushed archivists, as they move into the 1980s, to buy larger microcomputers or small minicomputers. 2

This application required archivists to learn about computer technology because procedures could not be spoon-fed from a manual devised by a network. Such expertise can be an advantage to archivists who may acquire machine-readable records generated by the personal computers of donors. In fact, archivists who braved this field can pride themselves in being part of what Alvin Toffler calls the "techno-rebels" of the "third wave." They have thus contributed to the "demassified information revolution" which Toffler feels provides an alternative to industrial society. 3

At present, standardization for archivists on the national level, despite the efforts of the National Information Systems Task Force (NISTF), may be more idealistic than practical. Furthermore, it may not be demanded by automation. Microcomputers can provide the advantages of automation without all of the demands for standardization and the problems
of integrating into a network. It should be noted that some standardization will occur when systems are developed from the same software packages or when archivists borrow ideas from one another, but this need not detract from the argument that microcomputers provide a certain flexibility that network-based, main-frame computers do not. For example, small computers provide for administrative and processing functions, such as form letter writing, label production, and report generation, that are not built into network systems.\(^4\)

It seems likely that archives in large universities will be required to provide data for book-oriented information networks. This idea is in the preliminary stages in many southeastern states. While archivists at such institutions should plan for this development, it need not conflict with the current use of microcomputers. Some network systems will use microcomputers to access network systems. It is likely that machine-readable information, particularly if it follows NISTF standards, can more easily be converted to use by networks than data that is not in machine-readable form. Current projects to convert SPINDEX to a MARC format may demonstrate the feasibility of conversion projects.\(^5\)

The use of a microcomputer by a project in Georgia proved that the smaller computers are less expensive to operate and more flexible than main-frame computers. In 1982 as part of a National Historical Publications and Records Commission (NHPRC) Needs Assessment Grant, the Manuscripts Task Force of the Georgia Historical Records Advisory Board funded a survey and directory of archival repositories in Georgia. The task force received bids from the computer department of a state university and the author who had just bought an Apple II+ microcomputer for his in-the-home business. Even though the microcomputer needed additional hardware to complete the job, the author's bid was $1000 less than that of the computer center at the state university whose staff was planning to write a program specifically for this application. The task
force also recognized the advantage of having a knowledgeable archivist inputting the data, particularly when interpretation was required.\textsuperscript{6}

For an archivist to undertake the directory project with very little training in computers meant that problems were inevitable. However, the success of the project indicates that the problems were not insurmountable, chiefly due to the relative simplicity of the microcomputer. The author selected a software package that allowed some statistical compilation, indexing, and formatting for publication. Lower case printing was not required. The package that met these specifications was a data management system called Personal Filing System (PFS) written in Pascal. The task force required that a letter quality printer be used. This added to the cost of the project and limited selection to the only low-cost, letter quality printer on the market at the time. The printer was so new that it had no proven service record; when it malfunctioned, no one could offer an effective diagnosis. In addition, in order to create the desired indexes, a second disk drive was needed.\textsuperscript{7}

Formatting data proved to be crucial, particularly in the production of printed output. This factor was not readily apparent when the project began, and at first, there was no easy way of changing the format without re-entering the data. Fortunately, part way through the project, PFS introduced an updated system which permitted shifting and lengthening of data fields and reading from one data disk to another.\textsuperscript{8} The update also contained added search capabilities and a measurement of the space left on the data disk. The data, which was gathered from a four-page survey of over one hundred fifty respondents, was so extensive that some thoughtful formatting was required in order to fit all the information on one disk.

Data fields and field names had to be limited in length and, thus, necessitated the use of abbreviations. The abbreviations made the printed output more workable though somewhat less
recognizable. Comments by the task force on abbreviation of field names were helpful. These are some of the same problems faced by institutions in networks, but with a microcomputer controlled by a small group, accommodation to the particular situation was less arduous. Though abbreviations were necessary, lengthy explanations of quantified answers to questions were recorded on the computer as appended pages to each form. This added feature of PFS demonstrated its adaptability to this unique situation.

The project's first output was a statistical compilation. The percentage figures which were easily and quickly tabulated by the computer were so numerous that more man-hours were spent interpreting the figures than in generating the statistics. The process indicates possibilities for social research in archives and the impact of computers in such a study. For example, the questionnaire found that about half of the respondents expressed "great interest" in computers. Using a Boolean logic function, it was determined that those with the greatest interest in computers came from the Atlanta and Macon (Georgia) areas. Though these calculations might have been done more easily on spreadsheet software like VisiCalc, PFS performed well without quantifying all fields.

Indexing and data editing were the most advantageous aspects of PFS in creating the Directory of Georgia Archives and Manuscript Repositories. The four alphabetical and eleven numerical indexes that went into the directory were sorted and printed in column style by the PFS Report package. The main body of the directory, to which the indexes referred, was organized and printed by the PFS package in two different formats (with and without field headings) similar to the NHPRC's national directory. Data editing required by proofreading, last minute changes, and additional information from follow-up phone interviews was facilitated by the search and organizational capabilities of the computer. Unlike SPINDEX, PFS
allows direct access to any record by searching one or more fields with whole or partial words and ranges of numbers. A plus was the ability to print mailing labels from the data in order to distribute copies of the directory.

Another Georgia project using PFS and PFS Report is the indexing of slave bills of sale by the Afro-American Family History Association. Using software that was deposited in 1983 for future updates of the archives and repositories directory, the association plans to create indexes similar to those in the directory and, in addition, use a keyword function to alphabetize names filed as multiple entries in a single data field. Statistics that can be generated on PFS by comparing fields such as sale price with age, location, etc., could prove valuable to econometric studies of slavery. The project's use of PFS was facilitated by recording data on control sheets with well-defined fields.

The projects that have been described show that the PFS data base management system coupled with PFS Report provides versatile software. Though it lacks word processing capabilities such as right-hand justifying and indented margins, the system does permit centering and margin-setting so that camera-ready copy can be produced. While it has some statistical functions and a complementary graphics package, it lacks many functions and cannot be used in conjunction with other software.

The recently established archives for Troup County, Georgia, used a Lanier E-Z 1 to prevent a backlog of recordkeeping which developed when its small staff was initially faced with processing a relatively large amount of governmental records, maps, and manuscripts. Though Lanier computers are designed primarily for word processing, the E-Z 1 has a data base management package that permits searching and sorting easily within a limited scope. The computer, which has a daisy wheel, letter quality printer and a five megabyte hard disk, was chosen primarily for Lanier's service reputation and the willingness of the local retailer to train and
suggest procedures to the archives' staff.¹²

The word processing functions, which include a spelling checker, enabled the archive to produce camera-ready copy for numerous, sophisticated newsletters, brochures, and press releases to publicize the new facility. Correspondence, particularly form letters, was easily handled by the computer. Mailing lists have been stored for use in mass mailing. Reports were generated on the computer with statistics compiled by the machine. Each of the procedures was easily learned and applied because the microcomputer and its manual were designed with these in mind. Such features as right-hand justification, proportional spacing, centering, limited searching, and many others found only in expensive word processing packages are built into the system. The repetitive tasks of generating folder and box labels also were made less tedious by use of the word processor.¹³

The Data Manager software was used to create several files. Since the Troup County Archives is the repository for local governmental records, the staff had to inventory quickly over seven hundred volumes of county records which included deed and tax information eagerly desired by genealogical patrons. The records were grouped together by the computer rather than on the shelf, thereby saving valuable staff time in physical organization. Using the sorting function of the software, an accession register was created for these records. The computer also assisted in the records management program by scheduling destruction and recording documents which had been pulled for use by the local agency.¹⁴

Though it has limitations, the Data Manager system has been used by the archives staff to track accession and administrative information, as well as to develop indexes for intellectual control of their records. Such data can be searched on-line using exact or partial match information. Tabulation of such figures as total volume accessioned or processed can be periodically compiled. Shelflist, donor list, accession record, and other administrative files can
be generated regularly by the computer, but they are not stored on disk as separately arranged indexes.

To provide information for patrons, a holdings list was created including collection title, size, record, type, processing status, brief description, existence of a guide, and inclusive dates. However, the program's limitation of field lengths printed on one line (one hundred and forty characters total per record) has made the prospective product unacceptable for patron use. Much of it had to be put in coded form which is not easily read. A more usable software package is being sought by the archive. Because the computer can support software written for CP/M operating systems, their options are many, but they will need one that is compatible with the Lanier word processing system. With an improved system, the archives staff hopes to produce name and subject indexing as well.15

The only major mechanical problem encountered by the archives was a defect in the hard disk. Due to proper floppy disk backup procedures, a potential loss of information was prevented. The storage capacity of the microcomputer is more than adequate for this newly developed archives. In fact, the large capacity of this relatively expensive microcomputer has tempted the staff to consider developing it as a multi-user system.16 The computer seems well suited to Troup County Archives' present needs.

At the University of North Carolina at Charlotte a data management package named Data Factory has been used on an Apple II computer to manipulate several of the university archives' files. University publications are entered as they are accessioned and then searched in one of as many as twenty fields or printed alphabetically in one of four different ways. The university archives' accession register, plus a list of its record groups and series, are also kept in the computer using Data Factory. Searching and printing lists of record groups, subgroups, and series numbers can be performed, as well as compilation of statistical analyses of these
categories. Monthly circulation and reference statistics for the archives are kept using the Data Factory which can do math functions in two fields.

The limitation of the system is the 48K memory of the Apple II. The computer, which was purchased in July 1982, has been superseded by the more powerful Apple IIe. For example, a comprehensive list of folders could not be handled by this small system. However, plans have been made to obtain an Apple with more than twice the memory to overcome this problem. Data Factory is quite flexible in that field lengths and format can be changed without deleting data. Files can be merged and multiline print formats can be defined. The program can read text files created by other Applesoft BASIC programs but cannot be exported to other systems. In addition, the eighty-eight page manual is sometimes unclear and contains no index.\textsuperscript{17}

An Applewriter word processing package has been used by the University of North Carolina archives and the manuscripts department to produce printed collection descriptions and box lists. Archivists with limited typing skills find they can input the information, edit, and have the computer print final copy without having to wait for clerical assistance from a typist. It is also hoped that subject and name indexes can be produced for the collection descriptions. However, a more sophisticated word processing package would be necessary to produce these indexes.\textsuperscript{18}

At the University of Kentucky in 1983, the Modern Political Collections Unit used an Apple II+ to print camera-ready copy for a guide to a governor's papers which included name and subject indexes. Box lists were also generated from the data keyed-in for the guide. The program used was ScreenWriter II, a word processing program that right-hand justifies, tabs, and generates page numbers and headings. The software alphabetized designated terms from the text and folder titles found in the guide, printing them automatically in an index which indicated the page on which the term was
found. The 48K memory of the microcomputer limited the practical amount of full text in one file to about eight pages, single-spaced. However, files can printed consecutively with the pagination, headers, and format preset for the entire document. The memory capacity also limits the number of terms in each of the two indexes to about one hundred terms from about twenty-five pages of text.19

In 1984 the Modern Political Collections Unit at Kentucky has also used an IBM PC with 512K memory and double-sided disks to create box lists, folder labels, subject indexes, and a guide to the photographs in a large collection. The software used, named PC-File, is a freeware TM data base management package with capabilities beyond more expensive software. For example, PC-File allows regularly used terms to be stored and then recalled for insertion by pressing two keys simultaneously. Field names can be changed quickly, and reports can be printed from fields in any order. Permanent indexes can be generated by copying a file and sorting it by any field. However, with eighteen fields and four hundred records per file, only a few files can be stored on one double-sided disk.20

A dot matrix printer has been used for box lists and folder labels, but access to a letter quality printer will be sought for printing the subject indexes and guides because they will be more regularly used by patrons. The thirty-page manual, which can be printed from the PC-File disk, is clear but rather brief. Files can be exported to other IBM software such as Wordstar, Mailmerge, VisiCalc, and Multiplan through a Data Interchange Format (DIF) file. Terms can be searched with partial match and linking of terms (and/or), though patron use of this capacity will be limited for the near future. The software makes excellent use of the PC's function keys. Special provisions in PC-File are made for the printing of labels for folders and mailing lists.21

The Microcomputer Archives and Records Management System (MARS) developed at Archives of
Appalachia, East Tennessee State University, was intended as a prototype archival system. With grant money received during 1980 and 1981, the archivist there planned to develop a "turn-key" or ready-to-use, integrated computer system for archives. A programmer, a consultant and literature searches helped the archivist plan for a system with three subsystems, "Accession," "Administration," and "Query Collection." When it proved too difficult to write a program for this application, a commercial data base management package was purchased along with an Apple II+ computer with 48K memory and two disk drives. DB Master, the program that was purchased, permits a more customized system than any of the data base managers previously mentioned without altering the source code. However, the plans or the system far exceeded that which was developed.

Unlike PFS, DB Master offers multi-diskette file handling, automatic data "packing" for increased disk capacity, password file protection, and "dynamic prompting" for designing instructions for the user on the screen. With these capabilities, an accession record was developed which tracked such information as accession number, title, date received, donor, record type, processing status, size, span dates, and location. At present, this preliminary record is the nearest the archives has come to having a complete record of its collections on disk. The file presently fits on one disk, though it is possible for DB Master to read two disks of the same file at one time. The file can serve as a shelflist and a donor's list as well as accession record.

The Query Collection subsystem, which would permit a patron to use partial-term and Boolean searches with as many as twenty terms to locate a full record of the collection, has yet to be developed. While DBMaster will permit such searches and can be structured for patron use, it will not allow information from the accession record to be read into the Query Collection format. This limits the system's intended capabilities as an integrated package, for many of the same fields will have to be
keyed-in on separate disks. Subject authority standards were curiously neglected when the system was first established, despite elaborate precautions to standardize procedures in anticipation of automation. This will be corrected in the plans for the new system which may offer the patrons a list of subjects by type and subheading.24

The Administrative subsystem was intended to keep track of staff work schedules, supply orders, budgets, and researcher registration. The files that actually were created were a lead file (tracks correspondence), a fund-raising file, and a researcher registration record. The computational and list handling capabilities of DB Master give these files a potential for handling some administrative duties. The registration file includes coded information on name, address, institutional affiliation, research interest, date, time, collection examined, and seat location. Much of the data for these files has not yet been converted from a paper format to machine-readable form.

A change of administration at the Archives of Appalachia has perhaps delayed the full use of the system. However, the new personnel also brought to the situation a desirable re-evaluation of a partially developed computer system. An additional use of the computer planned by the new administration is a grant-funded project to print a directory of Appalachian archives similar to the Georgia project that used PFS. The archives staff has been hampered in its production of usable printouts by a low quality thermal printer and hopes to acquire a better printer soon.25

The treasurer of the Tennessee Archivists has developed software for use with his personal Commodore 64 computer to handle the mailing and membership lists of that society. The Society of Georgia Archivists has also used a microcomputer to handle its mailing list. The program was designed to print both in alphabetical and zip code order. The former was for a list printed in their newsletter.
The latter helped meet post office requirements for bulk mailing. The chief problem was the absence of a tractor-feed on the letter quality printer. Without frequent attention the slick mailing labels slipped out of position in the friction-feed mechanism.26

The Kentucky Department for Libraries and Archives has developed a hybrid SPINDEX system for the grant-funded Kentucky guide project, an attempt to gain intellectual control of manuscript collections, as well as state and local governmental records, throughout the state. Through the use of a WANG OIS computer which telecommunicates with an IBM 360, the department can combine the capacity of the main-frame with the flexibility of a "super-microcomputer." The WANG, which is known for its word processing capabilities, has been used to handle correspondence, finance, and reports; but its most creative use has been text editing in conjunction with a SPINDEX program stored in an off-site, main-frame computer.27

SPINDEX, a program developed by the National Archives and Records Service (NARS) in the 1970s, must be run on a main-frame IBM 360 or larger. Thought it has undergone several revisions designed mostly to enhance the electronic photocomposition and indexing capabilities, it is rather difficult to access the data files. With the addition of the WANG, the department can edit and update the master file much more easily than the previous antiquated batch procedures. While the WANG does not make SPINDEX an interactive system whose files can be searched on-line by a patron, the microcomputer helps the archives diagnose the many bugs in the system and correct problems without printing all the data from beginning to end. This is particularly necessary for SPINDEX because of the poor system documentation furnished by NARS.28

The telecommunications package for the WANG also helps overcome some, if not all, of the handicaps of using an off-site, main-frame computer. Copies of the disk packs which store the master list of data are kept in the archives and at the main-frame site.
Once data is loaded at the computer center, the WANG can access it directly using little time on the main-frame. The telecommunications software should also permit the department to receive and transmit machine-readable information from other parts of Kentucky, but the details of this procedure have yet to be completed. The capability of dial up data bases in the state has also been an option.

The system allows entry of data in over nine hundred fields, only eighty of which have been used. The fields include as many as ten entries each for subjects, corporate names, personal names, media types, and geographical area. The terms can be integrated into a single index with cross-references and inversion of hyphenated terms. The limits of this system are not in its capacity to store and manipulate data. SPINDEX is designed to accommodate even item level description. Yet, the amount of time it takes to create the descriptions and type them into the machine will prohibit this kind of access except on a limited scale. Still, the project represents the use of a microcomputer toward development of a national data base. Early in 1984, the project had printouts for much of the state archives and thirty-five other repositories in Kentucky, thus showing progress on input for an estimated five thousand collections statewide.29

The current trend among archivists in these four southeastern states is to buy microcomputers with greater memory and more sophisticated data base management programs. Typical of this trend was the recent purchase of an IBM XT (with a hard disk) and D Base II program by the photographic archives at the University of Louisville. D Base II, a very popular package, enables the archivist to create an elaborate system of indexes from a single input. This powerful program requires the user to memorize what amounts to a language of commands, but its advantages are worth the effort to archivists like those in Louisville who had been using a main-frame computer.30

Many challenges face archivists who are using microcomputers. There is a need to find software
that can integrate the various functions required by archives. Few inexpensive microcomputer packages presently have compatible word processing, data base management, and spreadsheet software, though designers are beginning to market these. Despite the potential of the microcomputer for developing an interactive system, not one of the previously mentioned applications have in place an on-line system for patron use. However, archivists cited have found many useful functions for microcomputers. It is hoped that from simple steps such as those outlined that archival institutions will adopt microcomputers as a regularly applied tool and gradually build usable systems.31

NOTES

1 Two exceptions are the SELGEM program at the University of Louisville Photographic Archives and the SPINDEX program at the Kentucky Department for Libraries and Archives in Frankfort. Both now use microcomputers. Western Carolina University in Cullowhee, NC, uses a minicomputer for a manuscripts project. Apologies are due to appropriate projects that were left out of this study.


4 SPINDEX, the most widely used archival system allows users some flexibility. OCLC and RLIN, book systems that specialize in distributive processing, allow local fields, but have rigid MARC standards and high start-up costs. There have been RLIN and SOLINET (OCLC) workshops on manuscripts, but development has been slow. Nancy Sahli, "Prospectives from Archivists," in Automated ed. McCrank, 237-39, outlines the problems of networks for archivists.

5 "Newsreels," *Provenance: The Journal of the Society of Georgia Archivists* 1,1 (Spring 1983): 69, had a report that Duke University was planning to input manuscript data to OCLC. SUN Newsletter, September 1983, back page, had a description of the grant-funded RLIN to MARC conversion at Stanford, Yale, and Cornell.

6 Since the author was setting up a new business many hardware costs could be written off his taxes. Thus, in this situation it was as advantageous to use budgeted money for hardware as for labor costs. However, this meant start-up costs were sizeable.

7 It is believed that the problem stems from the printer's or interface card's inability to deal with a software package that is written in Pascal. The printer has no such problem printing the contents of Applesoft BASIC text files. PFS requires 48K memory and is also available for Apple IIe and III, plus the IBM PC, XT and compatibles.

1981), 18, notes the value of variable length fields for archives. In 1983 the data base part of PFS became known as PFS: File.

9 Multiline fields (over 40 characters) were printed in an unreadable way, and a single sided disk held less than 150 three-page records. Therefore, the text had to be abbreviated.

10 The Boolean logic function "and" searches the file for records with specified retrieval information in two or more data fields. The "not" function will retrieve all records that do not have what is specified in a particular data field. Thus Macon "and" Atlanta could be compared with all that were "not" from these cities.

11 Hickerson, Automated Access, 32.


13 Kaye Lanning, Micro PC User Survey conducted by Peter E. Schinkel and Glen McAninch (Fall 1983). The SGA Newsletter in 1983, as well as the newsletter for Troup County Archives, was compiled by Lanning on the Lanier computer.

14 Lanning, "Starting," 41; idem, spring meeting of the Society of Georgia Archivists, St. Simons Island, 26 May 1983.

15 Lanning, "Starting," 41.

16 The system costs approximately $16,000 and has 256k of internal memory.

17 Robin Brabham to Glen McAninch, 26 October

18 Brabham to McAninch, 26 October 1983; Screenwriter II for the Apple allows a limited amount of indexing of designated terms.

19 Screenwriter II is a powerful program that is complicated and sometimes awkward to use. The program itself requires much of the computer's memory.

20 PC-file version 9.1 requires only 96k memory and a single-sided disk but handles larger files with more memory and a double-sided disk drive. The latest version, PC-File III, works even faster than 9.1.

21 Douglas Clapp, "For What It's Worth," PC World, I, (6 September 1983): 154-59. The "freeware" software is designed to be copied with the provision that $35 be sent to the author, Jim Button. Clapp finds PC-file comparable to Visifile and Easyfile.


24 Kesner and Hurst "MARS", 13-15; Conversations with Archives of Appalachia staff, February 1984.

25 The archives has a word processing program
called Easywriter, which gets little use. There has been no attempt to integrate this with a DB Master.

26 Conversation with Mark Winter, October 1983.

27 Conversation with Thomas Converse 23 August 1983. The WANG OIS has 2.6MB memory and coordinates seven terminals each with 64K, a disk pack drive, several printers, a telecommunication package, and several other types of software. The computer, software, and peripherals cost about $40,000. In Washington, Oregon, and elsewhere WANGs and Kaypros are being used in similar SPINDEX applications according to the SUN Newsletter, 5 (1983). The Kentucky project also contributed to the identification of 20 percent of the state archives holdings which could be discarded.

28 Hickerson, Automated Access, 27-29.

29 Conversation with Thomas Converse, 23 August 1983; Thomas Converse, "Kentucky Guide Project", Kentucky Council on Archives Newsletter 5, 1 (Spring 1983) and 1, 2 (Fall 1979).


31 Many integrated packages like Lotus 1-2-3, are number oriented. See also SAVY for Apple II and the Incredible Jack. PFS:Write became available in late 1983 with the ability to incorporate mailing lists into form letters, but not the ability to reformat data files for publication. This is the same orientation as many recently available systems. VisiCorp and Microsoft are developing expensive integrated software. The use of DIF software and conversion of files to ASCII characters may be the best way to integrate software. Some operating systems allow integration.