

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Documentary Editing: Journal of the Association  
for Documentary Editing (1979-2011)

Documentary Editing, Association for


---

1983

## Computer-Assisted Document Control for Editorial Projects

Paul Machlis

Follow this and additional works at: <http://digitalcommons.unl.edu/docedit>

 Part of the [Digital Humanities Commons](#), [Other Arts and Humanities Commons](#), [Reading and Language Commons](#), and the [Technical and Professional Writing Commons](#)

---

Machlis, Paul, "Computer-Assisted Document Control for Editorial Projects" (1983). *Documentary Editing: Journal of the Association for Documentary Editing (1979-2011)*. 192.

<http://digitalcommons.unl.edu/docedit/192>

This Article is brought to you for free and open access by the Documentary Editing, Association for at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Documentary Editing: Journal of the Association for Documentary Editing (1979-2011) by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

# Computer-Assisted Document Control For Editorial Projects

PAUL MACHLIS

Effective organizing and indexing of historical and literary texts are essential to the work of documentary editing. Card files, typed lists, and computer printouts help us count and classify documents and provide access to materials by date, subject, owner, and other features. In doing so, they offer editors some control over the tens of thousands of documents that form each of the editions. Maintaining tools of document control is not our central work, but it is prerequisite to the publication of accurate and thoroughly annotated texts.

The following remarks concern the use of computers at four projects to produce instruments of document control that are more helpful, efficient, and flexible than their card-file alternatives. I will introduce some basic concepts of

computer-assisted control, describe each system, and offer general comment on their differences and similarities.

As editors, we are familiar with, if not actively using, word-processing capabilities of computers. The application of computing to document control utilizes different facilities, known as data processing, with which we may store and sort information about an item chosen for classification (such as a manuscript letter). That stored information is called a "record" and consists of numerous fields, or elements, of data. Many similarly structured records form a data file, which can itself be processed as a unit. Part or all of the data file can be sorted and printed in various ways. In this discussion the outcome of a document control system will be called a catalog.<sup>1</sup>

Before creating a computer-assisted catalog, a project should (1) define its goals in document control; (2) review and, if necessary, alter procedures of document acquisition, organization, and record-keeping so that the aims of document control and editing are best served; (3) determine the computer hardware (equipment) and software (programs or instructions) that are most appropriate to the task; and (4) define the records, fields, and lists to be used in the catalog. Four groups of procedures and guidelines should also be considered in producing the catalog: (1) input of records and initial proofreading; (2) sorting and selection of records from various lists; (3) output (paper, microform, or online); and (4) incorporation of revisions and additions.

Each project faces a unique set of factors in these eight areas of choice. The size and organization of the document collection, the types of documents to be edited, the urgency of need for document control, the plan of the edition, and the availability of administrative and financial support must all be considered. The choices to be made are interrelated and will require constant reevaluation during the design process. The variety of factors and options will be evident in the following descriptions of four document catalogs.

The Joseph Henry Papers (Smithsonian Institution) is preparing a comprehensive microfilm edition and a selected fifteen-volume letterpress edition of documents (mainly letters) by, to, and about the nineteenth-century physicist and first Secretary of the Smithsonian. Since 1966, the project has maintained a document control system on the Institution's mainframe (large, multi-user) computer, currently a Honeywell Series 60 level 66/20, using specially designed programs in the COBOL language as well as adaptations of the Smithsonian's "SELGEM" archival control programs. Information about approximately 55,000 of an estimated eighty to one-hundred thousand documents has been entered from worksheets prepared by project assistants; the remaining records are being added at a rate of about 3,000 per year. During a typical two- to three-minute examination of a document, the assistant records up to twenty-two fields

of information in up to fifty-five characters.

The fields chosen by the project are: the unique control number assigned to each document; a batch number, indicating the location of the project's microfilm or Xerox copy of the document; the number of pages in the document; document date, author, and addressee; the source (depository) holding the original document; up to two personal names mentioned in the document; up to five significant subjects treated in it; and a preliminary indication of whether the document will be included in the letterpress edition. Comment fields allow catalogers to indicate whether the author name, recipient name, month, day, or year have been editorially supplied, and whether the document has an enclosure, is not a letter, or spans several dates. Four lists are printed, on lettersize paper, with a Xerox 1200 system: a chronological list (subsorted by author), an alphabetical list by personal name (with writers, addressees, and mentioned persons interfiled and subsorted by date), a list by subject code (subsorted by date), and a list by control number. The current catalog is stored in looseleaf binders; an auxiliary binder guides users in the understanding of name, subject, and depository codes. A full paper printing of the catalog, including new records and data corrections, is produced every twelve to eighteen months. The data file has been used to produce specialized lists as needed, such as statistical tables and calendars for volumes of the letterpress edition, and will serve as the basis for the index to the microfilm edition.

An essential feature of the catalog is the use of codes for personal names, depositories, and subjects. Although initially name codes were acronyms (JHE for Joseph Henry), most are arbitrarily assigned. As there are only a few hundred depositories listed, depository codes roughly reflect institutional names. Subject indexing is governed by a thesaurus of 133 three-digit codes, arranged in nine groups (for instance, in the area of "Electricity and Magnetism" are codes for "Batteries," "Induction," "Magnets," "Wire," and other topics).<sup>2</sup>

PROGRAM: JHPPRT		JOSEPH HENRY PAPERS											DATE 05/25/82		PAGE 27							
NAME INDEX	CONTROL NUMBER	BATCH X M NO.	NUMBER PAGES	SYMBOL A/R	ENCL OR NON-CORR	SPAN S	M/D/Y ADDED	DATE MO DAY	YR	AUTHOR	RECIPIENT	SOURCE	NAMES TREATED		SUBJECTS					STATUS SELEC. TRANS.		
													A	B	A	B	C	D	E			
ADB	05977	M 070	01				03 16	39	ADB	JHE	SIA			866	645						002	
ADB	06766	M 071	01				03 16	39	ADB	JHE	SIA											
ADB	05978	M 070	08			S	03 19	39	JHE	ADB	SIA	TRB		866	270						002	
ADB	05979	M 070	02				03 20	39	JHE	ADB	SIA			560	284	145					002	
ADB	16682	M 122	02				03 20	39	JHE	ADB	SIA	JZM	GIC	627	145	235						
ADB	05980	M 070	01				03 23	39	ADB	JHE	SIA			284	654						002	
ADB	36201	X 066	03	R			04 16	39	ADB	ESA	PRO	JHE	CW1	407	170	449	645	866			002	
ADB	05981	M 070	02				04 19	39	ADB	JHE	SIA	JWQ		284	170						002	
ADB	17867	M 129	04				04 21	39	ADB	CRT	SIA	GIC	BID	049	284	170						
ADB	07513	M 072	01				04 25	39	JHE	HAH	SIA	ADB	SAL	654								
ADB	08274	M 098	01				04	39	JHE		SIA	ADB		170							000	
ADB	05982	M 070	03	A			05 04	39	ADB	HLL	SIA	JES	GIC	866	800	170	284	645				
ADB	05983	M 070	04				05 20	39	JHE	ADB	SIA	MFA		049	013	284					002	
ADB	05984	M 070	03				05 28	39	ADB	JHE	SIA			284	049	270					002	
ADB	00288	X 002	01				06 10	39	ADB	CCO	APS	ZOF		284	160							

Illustration 1

The Joseph Henry Papers Catalog, Name List, page 27, lists documents by, to, or mentioning Alexander D. Bache (ADB) in chronological order.

The Thomas A. Edison Papers (Rutgers University, New Brunswick, and Edison National Historic Site, West Orange, New Jersey) is preparing a selected six-part microfilm edition and a selected fifteen- to twenty-volume letterpress edition of correspondence, laboratory notebooks, patents, and litigation records. In 1980 the project implemented a document control system, closely modelled after that of the Joseph Henry Papers, with the use of an in-house Xerox 860 word processor, the Rutgers DEC 20 mainframe computer, and specially designed programs in the FORTRAN and SITBOL languages. Records include up to nineteen fields and fifty-two characters. They have been entered for about 9,000 documents selected for the first part of the microfilm edition from approximately three and one-half million pages of material in the entire archive. Project editors and student assistants prepare coded worksheets, at a rate roughly estimated at five to seven minutes per document, for later keypunching.

rections will be incorporated periodically in revised printings. Auxiliary lists of codes for names, subjects, and document types assist users of the catalog. The data file will serve as the foundation of the index to the microfilm edition.

The large number of documents has called for the processing of certain classes of documents (financial papers, inventories, brochures, and others) in groups rather than as individual items, with the identification of date, author, and recipient adjusted to apply to a collection of items. However, some documents call for more analysis than the record structure allows. When additional access for "names mentioned" or subjects is desired, a second record is created for the document. The wide variety of documents (unlike the preponderance of letters in the Henry catalog) and the desire of the editors to record all types of documents in one catalog have required guidelines for determining the authors and recipients of such items as drawings, lab notes,

Thomas A. Edison Papers														18-Mar-83	
Specific Name Code Report - JG2														Page 1	
Gr	Local	Ty	Date	Au1	Au2	Rp1	Rp2	Nr1	Nr2	Sj1	Sj2	Sj3	Status	Re1	Frme
XX	XXXXXXXX	XX	XX/XX/XX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXXXXX	XXX	XXXX X
D	7501F	01	75/09/10	JG2	TAE					280					
D	7511A	01	75/01/00	TAE	JG2			TTE		215	245		6	2	
D	7511D	01	75/07/26	TAE	JG2	742		ABC	WC2	613	215	230	7		
D	7511N	72	75/00/00	TAE				TTE	JG2	210			4		

Illustration 2  
The Thomas A. Edison Papers Catalog, Name Code Report for documents by, to, or mentioning Jay Gould (JG2).

The catalog's fields include: a call number specifying the location of the original document at the Historic Site; a two-character code to show whether the document is a letter, lab note, patent item, legal document, financial paper, or other type of material; the document date; up to two authors of the document; up to two recipients; up to two personal names mentioned; up to three subjects treated; an indication that a name or date is conjectural; and reel and frame numbers for the microfilm edition. As with the Henry project, three-letter codes are arbitrarily assigned for authors, recipients, and names mentioned, and three-digit codes are assigned for subjects within ten categories (for instance, under "Phonograph and Motion Picture" are categories for "Cylinder Phonograph," "Disk Phonograph," and "Kinetoscope," among others). The principal form of output is lists of records that match specific combinations of fields requested by project editors (for instance, all documents with a specified author and subject). Three other comprehensive lists are also currently produced: a list by file location; a frequencies list, in which names are printed with a count of their appearance in the data file; and the indexes to the volumes of the microfilm edition. New records and cor-

mortgages, stock and bond certificates, and diplomas. Because the microfilm edition and catalog currently record only those documents held at the Edison Historic Site, no depository information is given.<sup>3</sup>

The Papers of Jefferson Davis (Rice University) is preparing a selected letterpress edition of works and papers, including letters to Davis. An in-house Micom microprocessor, utilizing packaged database management software and data storage on eight-inch floppy disks, was acquired in 1980 to facilitate control over the approximately 100,000 documents of the edition. Records for some 24,000 documents related to the two volumes currently in preparation have been entered by project editors in up to eleven fields and up to 150 characters.

The fields of the Davis catalog are: date; author/recipient; type of document (report, request, memo, endorsement, etc.); one name mentioned; one subject heading; document type (ALS, LS, etc.); depository name and call number; project file location (film or folder number); an indication that the document will either be published in full or merely be listed in the series calendar; and discursive notes. Only one field is needed for both authors and recipients, because

55.12.11	f	Depierris, V.B.	query	Garland, N?	address	ALS	DNA, M567, r514	ED196	
55.12.11	f	Depierris, V.B.	query	Molinard, Albert J.S.	address	ALS	DNA, M567, r514	ED196	
55.12.11	f	Lee, John P.	submits*	Scott, Winfield	ct martial rec, 1811	LS	DNA, M221, r176	K301	ack 12.11
55.12.11	f	McClelland, R.	query	Henley, Thomas J.	clerk	LS	DNA, RG092, LsRecd	JQ413	
55.12.11	f	Totten, J.G.		* Barnard, John G.			DNA, RG094, USMA Recs	GL577	
55.12.11	f	Totten, J.G.	report	* Hulse, Isaac, Dr.	Pt Barrancas	LS	DNA, RG077, Land Papers	JH202	
55.12.11	f	Totten, J.G.		* Walker, William H.T.	USMA assignt		DNA, RG094, USMA Recs	GL577	
55.12.11	f	Zollicoffer, Felix K.	recom	Jackson, Alexander	USMA Bd Vis	ALS	DNA, RG094, USMA Recs	GM523	
55.12.11			refers		Liberty Depot stores	LbC	DNA, M444, r003, v5	Q244	to QMG
55.12.11	f	Tabor, James M.	query	*	leave, USMA	ALS	DLC, Davis	PP# 132	
55.12.12	to	Blanchard, Albert G.	ack		services, Kan; L, 12.3	PC	DNA, M221, r176	K107	
55.12.12	to	Kelly, John	reply	Wallace, Patrick	address	LbC	DNA, M006, r037	AL186	
55.12.12	to	McClelland, R.	reply	Henley, Thomas J.	clerk	PC	DNA, RG092, LsRecd	JQ414	
55.12.12	to	Vail, George	decision	Fleming, George	discharge denied	LbC	DNA, M006, r037	AL186	
55.12.12	f	Abert, J.J.	opinion*	Lester & Redington	Oswego contract		DNA, M506, r071	IL1250	JD upholds Abert decision
55.12.12	f	Atkinson, Robert J.	report	* Clark, Michael M.	mil contrib fund	LS	DNA, M221, r176	K066	
55.12.12	f	Atkinson, Robert J.	report	* Scott, Winfield	mil contrib fund	LS	DNA, M221, r176	K066	
55.12.12	f	Burnet, William E.	applic		comm; renews	ALS	DNA, RG107, Applics	KF272	
55.12.12	f	Cannon, William R.	discusse*	Brown, A.G.	pol, Miss	ALS	KyLxT	KS477	P
55.12.12	f	Cannon, William R.		* Davis, Jefferson	pol, senate	ALS	KyLxT	KS477	P
55.12.12	f	Guthrie, James	request	*	Hockaway Beach	LS	DNA, RG077, Land Papers	JH303	
55.12.12	f	Lawson, Thomas	query	*	Va Med & Surg Jour	LbC	DNA, Rg112, Ls to Sec	J1071	JD E

Illustration 3

The Papers of Jefferson Davis Catalog, sample page from a Date List, with sorting by columns two, three, and five.

Davis is always assumed to be either the writer or addressee; "to" or "f" in the preceding field indicates whether the document was to or from the person named. Editorial doubt about a name or date is indicated with a question mark; such doubtful names and dates follow those without question marks, in the various sortings. Codes are employed for depositories (NUC symbols) and types of document (ALS, etc.). Additional records are frequently created for documents when access is desired through more than one subject or name mentioned and to different versions of one document (draft, signed letter, letterbook copy). Although the project could generate full listings of the data file according to such elements as date and author, online access has so far made such lists unnecessary. Lists are produced, on the in-house daisywheel printer, in response to specific research or editorial needs. Because the computer system allows printing of documents while two users are inputting or editing data, frequent requests for printouts are not inconvenient. New records and corrections are incorporated into the data file periodically. The data base serves as an outline for the letterpress calendar and could serve similarly for a microfilm edition.<sup>4</sup>

The Mark Twain Project (The Bancroft Library, University of California, Berkeley) is preparing a letterpress edition of the complete works and papers of Samuel L. Clemens. The project implemented a control system in 1981 on a campus PDP 11/70 computer (UNIX operating system), utilizing programs specially written in the BASIC language. Information about approximately 10,000 letters by Mark Twain has been entered in records of up to twenty fields and 350 characters (with an average record length of 100 characters). A student assistant examines the project's manuscript, photocopy, or typescript texts and directly enters data through an online terminal at an average rate of one record per minute. The catalog records letters by Clemens' wife and daughters, letters by secretaries writing on behalf of Clemens, and letters known to have been written but for which no text has been located (ghost letters). Separate catalogs are in preparation for some 12,000 letters to or about Clemens, and some 5,000 writings by him.

The fields of the catalog are: a unique record number; up to two writers; up to two addressees; a date modifier (circa, before, after, postmarked, etc.); up to three dates; up to two places of writing; up to two sources of the text (to catalog composite and radiating texts accurately); and discursive notes. Lists are printed with records arranged by date, writer, addressee, place of writing, record number, day, and text source. Multiple source lists—one subsorted by date, one by addressee—allow ready collation with differing cataloging practices at different depositories. The day list groups letters by day of month (all 2 May letters together) to facilitate identification of letters listed in auction or depository catalogs without indication of year. All lists have secondary and tertiary sorting keys for records with identical primary sorting keys. Records are duplicated for particular lists: a letter addressed to two persons appears once under each name in the addressee list. Reference guides to abbreviations and source codes accompany each list. The first edition of the catalog was printed on a Xerox 6770 laser printer and bound in eight volumes. A supplement is updated monthly and will be merged with the main data file for periodic reprinting. Records are corrected with the UNIX text editor.

The catalog utilizes two- or three-letter abbreviations for writers. For the place of writing a hierarchical listing was adopted that provides sortings by state or country, then by city. Standard U.S. postal abbreviations for states were supplemented by similarly constructed codes for countries ("SH" and ship name are given for letters written at sea). The source of the most authoritative text is represented in two manners: the first, used for private and public collections, consists of a two-letter state or country code, the collection name (often in short form), and an indication whether the text source is a manuscript, transcript, or photocopy; the second, used for printed sources, consists of the publication name and pertinent bibliographic information. Uncertainty about names, dates, and places is mentioned in the notes, and unknown parts of a date are represented by asterisks. Although final printouts are relatively free of codes, the speed and accuracy of input are increased by the

OLC	to Howland, Robert M. 00677	1871.11.20	CT.Hartford	CA.MTP.ms
SLC and OLC	to Howland, Robert M. and Howland, Louise A. 04656	ca. 1877.09.03	NY.Elmira	CA.MTP.ms.
SLC	to Howland, Robert M. 07428 notes: year uncertain.	1879.10.27	CT.Hartford	CA.Daley.ms
SLC	to Howland, Robert M. 02810	1879.12.14	CT.Hartford	CA.MTP.ms
SLC	to Howland, Robert M. 04128	1879.12.29	CT.Hartford	CA.MTP.ms.
SLC	to Howland, Robert M. 02325 notes: date uncertain.	1879.12.29	CT.Hartford	CA.MTP.ms

Illustration 4

The Mark Twain Project Catalog, Addressee List, page 272, lists documents alphabetically by addressee name, with subsorting by date and writer.

liberal use of codes that are later converted to full listings by "global" substitution commands.<sup>5</sup>

Undoubtedly, computer-generated catalogs exist at other editorial projects. As they are publicly described, the variety of approaches to document control will become more apparent. Other applications of data processing, both inside and outside of the editorial sphere, will display techniques pertinent to our concerns.<sup>6</sup> Even the preceding brief descriptions of only four catalogs illustrate that somewhat different solutions to similar problems will be appropriate for different projects.

The Henry, Edison, and Davis projects supplement access by writers, dates, and other features with access to the content of documents, through preliminary subject and "names mentioned" indexing. In each case, subject indexing has been of great assistance in the preparation of selected editions, in that it "aids the editors in choosing documents for letterpress publication while making the significant information in the unselected documents readily available for use in annotation."<sup>7</sup> The Twain project chose a more modest goal of access by prominent features of documents because of an urgent need for improved cataloging. By rejecting time-consuming input of desirable but less essential information, the project completed an initial version of its catalog very quickly and relatively inexpensively.

The Twain project chose to record its three major collections in three catalogs with different record designs. The user may be inconvenienced by having to consult three reference tools but at the same time may benefit from cataloging more pertinent to the particular classes of documents: certainly the inclusion of letters with such documents as manuscripts of novels, autobiographical sketches, and poems would have required a complex and possibly cumbersome record design. The division of the edition into series (letters, sketches, etc.) also argued for separate catalogs. The other projects, each of which edits documents for strictly chronological (not series) publication, chose to cata-

log together a variety of documents (although most can be convincingly likened to letters).

Each project has chosen only those characteristics of documents that are most important to its goals. The Henry project chose not to record detailed depository call numbers, specification of document type (ALS, LS, etc.), or the identification of unclear names. The Twain project rejected as fields the type of paper, watermark, letterhead, ink color, subjects, occurrences in print, depository call numbers, and key phrases. Only the Henry project records the number of pages, only the Twain project records ghost letters and place of writing, and only the Davis project records detailed depository information and different versions of a text. The decision to exclude information from the record format may be very difficult to make. Yet the inclusion of all information that might prove useful would raise the cost and time of cataloging of any collection to unmanageable levels.

The projects use different computer hardware and software. The Davis project enjoys the convenience and independence of the in-house microcomputer but would like to have the increased storage capacity of hard disks. The Edison project utilizes the convenience and low cost of its microprocessor for record input but takes advantage of the superior processing and storage capabilities of the mainframe for other functions. The Twain and Henry projects rely exclusively on mainframe computers.

Each catalog has required careful planning of document analysis and input. Editors who know their materials well have designed catalogs that can accommodate peculiarities of their collections and a variety of unusual features. Patience and persistence in tackling the initial decisions of catalog design have saved them from inconsistencies and inadequacies in their data files, problems that could be repaired only by largescale, expensive revision of programs and data. The active involvement of all editors of the projects, including those unfamiliar with cataloging and computers, has improved both the quality and the ease of acceptance of the catalogs. Consistent cataloging has been

achieved by strictly limiting and formalizing data elements according to in-house thesauri and manuals of procedure.

Different input procedures have been used: readers at the Twain project enter data directly during document analysis while the other projects employ readers to complete worksheets for later keystroking. The "global" search-and-replace function of computers facilitates the input of highly abbreviated information, for uncoded catalogs like those for Twain and Davis, that can be easily transformed into correct form.

The projects face common problems concerning output. While all editors wish to consult data bases that are complete and accurate, they must also consider the cost and effort of continual correction and reprinting of lists. The Henry project is considering less costly microfilm output, but has found some form of "hard copy" useful for many project functions, such as producing photocopy of selected catalog information. It is also considering less frequent printings, issuance of selected lists, and online access. The Twain project is experimenting with compressed lineprinter output of selected lists. The Davis and Edison projects enjoy online access, the opportunity to revise and reprint records continually, and relatively inexpensive printing on in-house printers. Each project must devise methods of alerting users to errors in the current printings; unlike the penciled marking (and possibly refileing) of two or three index cards, correction of computer-assisted catalogs will require correction of each error in each printed listing as well as of the stored record itself.

Differences exist in specific record and field features. Twain documents span two centuries and thus cannot conveniently be recorded with two-digit years in date fields, as can the documents of the other three projects. The Davis catalog assumes the major figure (Davis) to be either the writer or the addressee; the others enter the name of their major figure for each appropriate record and also record documents neither to nor from their major figures. The Davis project includes question marks within conjectural information (with sorting order affected), the Henry and Edison projects express doubt with codes in fields assigned to that purpose, and the Twain project mentions doubt in its discursive notes. The projects differ in the number of fields allotted to various types of information, such as authors, addressees, and dates, and differ in the extent to which extra records must be created for documents requiring more than the allotted number. The Twain and Davis catalogs, containing variable length records and few codes, require little use of ancillary reference lists but more storage space, paper, and computing time than the coded files of the Edison and Henry catalogs. Coded files require constant reference to "code guides" for input, proofreading, and consultation of the catalog. These differences are relatively unrestricting for output because of the powerful search-and-replace capabilities of computer systems: any project may choose to change codes to full forms, full names to codes,

or alter the format of printouts, with an ease unimaginable with manual control systems.

Analysis of the four systems reviewed here suggests several precepts for successful computer-assisted cataloging: (1) a comprehensive knowledge of the document collection and editorial plan (more than of computer operations) is essential to produce a catalog that is most suitable to a project; (2) the adoption of limited goals and the acceptance of restrictions in time and money will produce useful reference tools in reasonable time; (3) a willingness to give ample time to initial design and document review will prevent major deficiencies and inconsistencies from arising during implementation; (4) documentation in the form of manuals and thesauri is necessary to maintain consistency and accuracy in cataloging.

Just as the publication of accurate and well-annotated texts is a desirable basis for accurate historical and literary scholarship, the preparation of tools of document control is necessary for responsible textual editing. Computers can help us produce catalogs that contribute significantly to the quality of our editorial work.

#### NOTES

1. For an excellent discussion of the application of data processing to archival control, see H. Thomas Hickerson, *Archives & Manuscripts: An Introduction to Automated Access* (Chicago: Society of American Archivists, 1981). Additional publications in the field are listed in *Automation, Machine-Readable Records, and Archival Administration: An Annotated Bibliography*, ed. Richard M. Kester (Chicago: Society of American Archivists, 1980).
2. Information about the Henry catalog has been contributed by Kathleen W. Dorman, Assistant Editor, Joseph Henry Papers, Smithsonian Institution, Washington, DC 20560.
3. Information about the Edison catalog has been contributed by Thomas E. Jeffrey, Associate Editor, Thomas A. Edison Papers, Edison National Historic Site, Main Street and Lakeside Avenue, West Orange, NJ 07052.
4. Information about the Davis catalog has been contributed by Lynda L. Crist, Editor, Papers of Jefferson Davis, Rice University, P. O. Box 1892, Houston, TX 77251.
5. Additional information about the Mark Twain catalog can be obtained from Paul Machlis, Associate Editor, Mark Twain Project, The Bancroft Library, University of California, Berkeley, CA 94720.
6. For information on the use of data processing for related purposes, see Bruce Wheaton, "A Computer Database System To Store and Display Archival Data on Correspondence of Historical Significance" (*American Archivist*, 45 [Fall 1982]: 455-466); Charles W. Polzer, "The Documentary Relations of the Southwest" (*Hispanic American Historical Review*, 58 [1978]: 460-465); *Users' Guide to Cindex . . . A Computer Indexing Program* (Papers of Henry Laurens, University of South Carolina, 1981); and Hickerson, *op cit*.
7. James M. Hobbins and Kathleen Waldenfels, *Computer Applications for Historical Editing Projects: The Henry Papers Index of Documents* (unpublished, 1980).