

The Historian and the New Technology

By JAMES B. RHOADS

National Archives and Records Service

LAST December, at the luncheon of the Society of American Archivists held during the annual meeting of the American Historical Association, John Hope Franklin delivered a paper that was warmly received by the many who had gathered to hear it.

It was not my good fortune to be present at the reading, but I heard numerous reports, all phrased in terms of "stimulating," "provocative," "entertaining," "scholarly," and words equally discomfoting. Discomfoting to me because I was the act standing in the wings, and the accolades confirmed what I had all-too-well expected: "Dr. Franklin is a tough act to follow."

Already committed to make this appearance today, I considered fashioning my own presentation after his, in the hope of imitating what would be difficult or impossible to "top." As everyone who heard Dr. Franklin will recall, he recounted what might well be called the "adventures" of his University of Chicago seminar, which spent part of one semester in archival research at the North Carolina Archives in Raleigh. Dr. Franklin cleverly titled his paper "Archival Odysseys: Taking Students to the Sources." For my own paper today, I considered the title "Archival Oddities: Taking Sources to the Students." Wisely, I am sure, I abandoned the ill-conceived notion that I might be able to match the stride of the illustrious and inimitable Dr. Franklin, but I ruminated somewhat fondly over that parodied title—"Archival Oddities: Taking Sources to the Students." Aside from the fact that it is in the style of John Hope Franklin, it smacks of an incongruity that is appropriately gross for my topic today. It suggests that knowledge about archival sources has not been given to researchers. And while that is certainly not true literally, it begins to have some degree of validity if we are allowed to speculate about information systems and capabilities of the near future.

Before getting into the near future, however, let's look a bit at the not too distant past. In the past 30 years the archival profession in the United States has developed from all but nonexistence into a respected body of professionals who have produced many innovations meant to make sources more readily available to researchers. The primary researcher in archives has been the historian. Archivists, of course, have always been on hand to help the historian master the sources. Not merely keeper of the record, the archivist is as much the keeper of

The author, Archivist of the United States and a Fellow of the Society of American Archivists, read this paper at the Society's annual luncheon meeting with the Organization of American Historians, on Thursday, Apr. 17, 1969, in Philadelphia, Pa.

Ariadne's thread that leads the historian out of the maze of research material. The archivist spends much of his time charting the sources for historians by producing guides, inventories, special lists, indexes, and other finding aids as reference tools. Over the past three decades the basic finding aid—the archives inventory—was developed to the state that we know it today.

The inventory format was not developed in order to impose system on unsystematic material. The records of concern to the archivist are inherently systematic and need only be described in a manner that reflects their order. Thus, the archivist is not a cataloger or a technician, but he *is* an information specialist in the truest sense—a historian analyzing existing structure and providing information about the content of large bodies of historical material. The archivist is schooled in historical method and historiography and has himself been weaned on Becker, Beard, and Boorstin. His goal is not a compact, universal system but a body of usable information. It is, therefore, one of the tasks of the archivist, working from the base line of his own historical background, to devise and improve ways in which fellow historians can probe the sources of information available to them.

If the traditional inventory approach to records has served us well over the years, the format is beginning to falter. We are approaching a stage in the production of inventories where they can no longer usefully serve the purpose of the researcher as they once did. This condition is being brought about by the growing numbers of inventories, both within an institution and among related institutions. There are a number of things wrong with the traditional inventory that many of you, I am sure, are well aware of. I will enumerate a few of the faults.

The inventory format describes records in a structure that is probably totally irrelevant to the approach to the records the researcher wishes to take. Inventories are arranged according to the hierarchical structure of the agency whose records they describe. There are few researchers around today writing administrative history. Rather, researchers are writing of events and movements and need reference to an agency's activities only as they reflect those events and movements.

The inventory format is based on presentation of whatever information is most obvious to the archivist from an analysis of the records. It does not necessarily reflect the needs of the researcher because the compiler does not regularly study the needs of researchers in historical source materials. We may even assume that the compiler is using terms and descriptions that in many cases are irrelevant to the needs of the researcher. One wonders how many researchers have given up on a group of records after seeing its mass described in that great archival euphemism: General Correspondence.

The inventory format may tend to make moribund a body of living records. Seemingly without recognizing that record holdings grow, are reprocessed, are more deeply analyzed with use, and may be described

in a variety of ways by a variety of people, the published inventory effectively freezes the usable quantity of the records to only those that have been included on the day the inventory goes to press; it inhibits reprocessing of records lest the published inventory be made obsolete (and *then* what does one do with the 200 or so copies still in stock?). The printed inventory does not permit inclusion of new information as such information may emerge from continued use of the records, and it accepts as definitive the description of one archivist, whose own prejudices and historical shortcomings tend to be reflected in the descriptions that he produces.

With all of these objections granted, it is still necessary for the archivist to describe his records in broad generalizations and to put these into some distributable form so that they may be read either at the archives or away from it. This has been the only practical method by which we could deal with great masses of raw materials in an orderly and systematic manner.

Now, however, the archivist is beginning to appreciate the applicability of some new techniques to his problem. He is beginning to see the computer as a boon, not to the advancement of system and order in archival description but rather to the advancement of specific manipulation of large bodies of data to meet the individual needs of the historian or other researcher. The use of the computer has already begun in archives, and its applications will grow as the initial simple routines to which it is being assigned are mastered. As we see it now, the computer will be put to use in three major areas in archives. I should like to speak briefly about them.

First, the computer will be, and, indeed, is being used to perform simple clerical or organizational duties. It is being used to create data banks of the holdings of repositories so that a request to the machine for information about records will reveal their existence, size, location, dates, processing status, and a variety of other related information. This machine approach is much like preparing a machine inventory of the contents of a warehouse or a parts supply depot. If you want to find a part, you feed the machine a description or a name and it comes back with the shelf location, number of units in stock, and perhaps a more complete description. Such programs are very useful in large institutions, and are basically very simple. They do not perform the duties of a trained archivist, because trained archivists are not needed to locate records in the building.

Second, the computer should be used to analyze the research methods of archives' patrons. We propose that an experiment be undertaken in which all finding aids and reference tools for an archives will be put in their present state into a single data bank, and that the data bank be queried every time a question is asked about the records. Such an experiment would serve a number of purposes. All questions about the records, whether simple or complex, would be registered in the system, and a

cumulative record of such *requests* would be maintained by the archives. A cumulative record of all machine *responses* would be maintained. From matching the two over the span of 2 or 3 years, there should emerge a pattern of query and response that would conclusively indicate whether or not the information that archivists provide about their records is pertinent to the research questions posed by historians and other patrons. The results might show that valid questions are not being answered or, conversely, that historians as evidenced by their questions have little knowledge of the nature of the archival materials with which they are working.

The threshold that this computer use leads us to is an understanding of the methodology employed by historians in doing research. It is already fairly evident that, although the archivist describes his records structurally, the researcher most often approaches those same records conceptually. In the past the archivist has attempted to meet the needs of the concept-oriented researcher by producing special guides developed around concepts instead of records structure. The guides might relate to one historical event such as the Civil War, or a geographic area such as Alaska, or a social movement such as civil rights. Such guides are produced by a thorough analysis of standard inventories, special lists, and other reference tools by a synthesis of information from them built around the theme of the special guide, and by possibly limited examination of the records. But the archivist, working sometimes for years on one special guide, cannot possibly keep up with the conceptual variations of the entire range of historical research. Here is where much work has yet to be done and where we think a computer approach might be valid. In an attempt to conceptualize we would not use the computer simply as a filing clerk and typist but rather as a cybernetic extension of the researcher himself.

It is not certain how such a cybernetic approach would work. We can presume that the study of historical methodology that we proposed above would give us some guidelines on how to proceed with developing that methodology. We can also presume that the data base that contains the conceptual framework will need to have in it all the tools that the archivist now uses in manually preparing a special guide: that is, there would have to be in the computer *all* the information currently known about the archives' records. If the archivist preparing a manual guide must have access to finding aids, lists, catalogs, and other reference tools, so must the computer data base. If the archivist must have in his mind certain terms, words, phrases, names and other identifiers when he begins his search, so must the computer data bank contain them, and there must be the capability for modifying and deepening the search as it proceeds and new leads are turned up. In short, the computer must provide a silent rapport between the archivist's inquiring mind and the sources of information at his disposal. The archivist or researcher can then feed his concept into the computer, with perhaps a list of relevant

terms, phrases, names, and so on, and program the computer to conduct a rapid search, to retrieve, and to format the information as a special guide. The data might originally have been entered into the machine system in a hierarchical organization, but once in, it is considered to be a fluid body of data, ready to be retrieved according to the plan of the researcher. We may in fact see the end of printed or published guides to or inventories of records. Instead of the computer forcing the researcher into a strict system of conformity and rigidity according to machine standards, it is obvious that its great storage and processing capabilities are leading to a state of individual research freedom impossible to attain in a manual system. It is the present system that makes all researchers conform to the same pattern of documentation; it will be the computer that will liberate researchers and enhance the role of the individual in his attempt to reconstruct the past by giving each researcher the opportunity to ask for information in the form and to the extent that suits his personal needs best. That is freedom.

Lest you accuse me of hypothesizing too radically for current capability, let me say that members of our information retrieval staff at the National Archives have long been studying the widest application of the computer to the needs of the researcher, and they feel that they have worked out many of the theoretical methods of the cybernetic approach. They are moving ahead, trying to institute research analysis programs, and they hope in the not too distant future to try some automated conceptualization of historical problems. They are operating as an information shop should, in a free atmosphere of inquiry and experimentation on the basic problems facing the archival world today.

We are in the computerized information area for the long pull, and wherever possible we should like to make it a professional effort instead of just an institutional project. From the projects that we have been carrying on in this field so far, we think we are moving in the right direction.

