

Manuscripts of Science— Analysis and Description

By NATHAN REINGOLD *

Yale University

LET US make several assumptions at the outset. (1) Scientists will create and maintain bodies of manuscripts. (2) These manuscripts will eventually find their way to suitable depositories. (3) The depositories in turn will provide the National Union Catalog of Manuscript Collections with entries for these collections to enable scholars to locate promising bodies of source materials and obtain a general idea of their contents. (4) The scholar interested in pursuing research at a depository will have available a suitable array of guides, catalogs, inventories, registers, lists, and indexes to collections that will themselves be properly arranged.

Of the four assumptions, the first two are the least likely to be true now or in the near future. Otherwise, the Conference on Science Manuscripts would never have come into being. The third assumption, that concerning the National Union Catalog of Manuscript Collections, is subject to several reservations. First, not all institutions with manuscript collections are committed to furnish entries to the catalog. As the catalog grows, these organizations should come into the fold. Secondly, many manuscript sources for the history of science will be in the possession of scientific institutions whose managers are not likely to think of sending in entries or to receive invitations to do so unless a special effort is made to reach them. Furthermore, the Union Catalog operates at present under a practical limitation. Only personal papers of individuals and collections of records of institutions or other organizations in the possession of manuscript depositories are included. Specifically excluded at this time are the holdings of archival agencies, that is to say, the records of institutions retained and serviced by them or in

* Paper read at the Conference on Science Manuscripts, Washington, D. C., May 6, 1960. The author was formerly on the staff of the National Archives, where his work in analysis and description resulted notably in the compilation of the *Preliminary Inventory of the Records of the Coast and Geodetic Survey* (Washington, 1958). In 1959 he joined the staff of the Science and Technology Division of the Library of Congress, and in the fall of 1960 he took a leave of absence to become visiting senior research fellow in the department of history of science and medicine, Yale University.

a specific archival agency established or designated for the purpose. The Union Catalog does encourage the submission of entries for private papers and archival bodies in archival institutions where these records are not explicitly part of the main archival groups one would expect to find there. The complexity of private manuscript collections is so great that some such limitation is necessary at the start. But the permanent exclusion, for example, of such great bodies of source materials as the records of the Federal Government in the National Archives, the archives of the States, and the archives of the great industrial concerns would seriously hamper the development of the catalog as the central file of information on the major research materials in the United States. For my part I hope that the National Union Catalog of Manuscript Collections will eventually expand its scope to include archival collections. When the vast quantities of personal-paper collections are substantially covered by the catalog, policies concerning archival institutions will undoubtedly be reappraised.

The final assumption, which largely determines the scope of this paper, is of great importance but is frequently taken for granted or passed over with a few generalities. Obviously this conference wishes to promote not the preservation of collections as relics piled up on shelves but the preservation of manuscripts in ways facilitating use. The success of the National Union Catalog will depend to a great degree on the extent to which the manuscript custodians have successfully arranged, described, and analyzed their collections.

Except for passing references later on, I shall not speak about the arrangement of manuscripts in collections. The importance of proper arrangement is self-evident. What I propose to talk about is the existing policies and practices governing the preparation of analyses and descriptions of manuscripts and what the ideal practices and policies might be.

Discussion of present policies and practices is somewhat difficult. There is no clearcut policy governing the treatment of manuscripts in all depositories, but rather a confused pattern of old viewpoints cheek by jowl with new approaches. In actual practice there is an appalling lack of uniformity. Here the standards of the National Union Catalog and the procedures of the archival profession give hope of greater standardization. What follows, then, is a broad and impressionistic generalization rather than a statistically accurate picture of the present situation.

When the modern historical profession emerged in the nineteenth century, there developed a great emphasis on the use of

documentary sources. Since the focus of interest was then on the medieval and early modern period, the profession was concerned with bodies of manuscripts that are small compared to modern collections. Whenever possible, these records, being relatively few and therefore relatively more valuable, were published in full. If publication in full was not possible, the preferred treatment was calendaring—abstracts arranged in chronological order. Next in terms of desirability was the preparation of item-by-item lists, catalogs, and indexes, sometimes with attempts at abstracting the contents of each item. Common to all three approaches is that each manuscript is somehow treated individually. Professional librarianship not only led to the individual document approach but added the paraphernalia of modern cataloging with cross-references, “added entries,” etc.

But as more recent bodies of manuscripts came into the hands of custodians, attempts at item-by-item treatment became increasingly difficult and were more and more set aside in favor of simply keeping on top of the influx of voluminous collections. The making of cards or lists with indications of subject content was dropped in favor of name indexes. Name indexes, in turn, became increasingly uncommon and the larger institutions began to rely on the general entry for an entire collection, modeled on the item entry.

Here the practices of the archival profession, especially in the National Archives, were influential; for archival bodies are usually very voluminous and complex. The National Archives, consequently, developed a form of treatment of records modeled not on the library-oriented entry but on the practices of European archivists. At the present time, where manuscript custodians face large and complex collections, they tend to analyze and describe them in archival or quasi-archival terms. That is to say, a collection is divided into its component series, and each series is described in physical terms (kind of documents, inclusive dates, quantity, etc.), with varying degrees of attempted analysis of contents. Most frequently, historical societies and libraries analyze collections in terms of names of correspondents with little or no reference to other kinds of subjects. In the National Archives, in contrast, entries rarely use personal names but as a rule show details of organization and of function.

In American institutions having manuscript collections one is likely to find one or more of the following situations. (1) Part of the collection was treated in the past on an item-by-item basis. (2) Individual items coming in as such are still separately treated. (3) Special prized collections are treated on a paper-by-paper basis, as

in the presidential papers project in the Library of Congress. (4) Collections are described in general entries influenced by library practices with a minimum of subject treatment. (5) Collections, especially the larger ones, are archivally described in inventories, registers, and the like, with emphasis on their physical features.

An interesting question is the optimum and maximum numbers of items permitting item-by-item treatment. Mere quantity is not a sufficient criterion since the complexity of the collection is also involved. A simple, large collection (simple in terms of kinds of manuscripts and contents) is easier to handle than a small, variegated body of manuscripts. Since no careful operational studies exist, I shall rely on my impressions and observations of various collections and some crude calculations based upon present practices, under the assumption that no technological revolution will alter the situation radically. Disregarding the matter of complexity, the question has two aspects: the size of the total holdings of an institution and the size of a single collection. As to the former, the experience of the American Philosophical Society, which specializes in manuscripts relating to science, is illuminating. The society is one of the last institutions committed to having a catalog card for each item. But not all of the nearly 200,000 manuscripts in the library are so treated. The society's own archives are so treated only through 1888; several large, modern collections are represented in the catalog by one general card each; and for some collections the catalog cards give little or no subject information and are closer to the index card than the catalog entry. Total manuscript holdings of 100,000 to 200,000 pieces probably represent the upper limiting range for item-by-item treatment. Only by tolerance of huge backlogs or by the use of crash programs could a depository embark on item-by-item treatment of larger quantities.

As to single collections, if we assume that no crash programs are started, the situation is something like this. Within the normal budgetary restraints a collection of up to a thousand items could be given item-by-item treatment. For collections between 1,000 and 10,000 items, the depository could embark on a cataloging program, but only with great difficulty and with a tendency to limit the treatment to names and dates only. Over the 10,000 mark, cataloging or quasi-cataloging is impractical unless funds for a crash program are raised.

But many institutions with less than 100,000 manuscripts and with collections of fewer than 1,000 pieces each have long since given up attempts at describing each item, except where an item

stands alone and not as part of a collection. The norm is the description of a collection as a unit for at least three reasons: (1) manuscript depositories lack the funds needed for item-by-item treatment; (2) the size of the total holdings and the number of large, complex collections are increasing as more modern papers are accessioned; and (3) since not every manuscript in a collection is equally valuable, it would be unwise to expend limited funds on all of them. Here the crucial observation is that some documents are important not in themselves but in their cumulative contribution to the whole.

We can state the problem of describing and analyzing modern bodies of manuscripts as follows: Since it is not practical or always desirable to have item-by-item treatment, how can we describe large bodies of manuscripts in the mass and still give adequate information or sense of the particular? As collections become bigger and more complex, the intellectual squeezing of an enormous body of manuscripts onto a 3"x5" card or the summation of a series comprising several hundred linear feet into a series entry of a few lines is less and less useful to scholars. As a practical matter, they either turn to unpublished, more detailed lists and similar aids that may exist for some collections or they are forced to plow through a mass of manuscripts to find that few or none are pertinent.

Any general policy governing our approach to manuscripts and their analysis must start from a conception of the nature of the source materials and of the intellectual operations performed upon them. At any given moment in history, such as the present, we can postulate an ideal distribution of information. Some information is in print, some in manuscript, still other in the form of artifacts; and a considerable body of information is solely within the minds of individuals. In time much of this information is lost. The surviving printed works are in libraries, the artifacts go to museums, and, if we are lucky, some oral traditions transmit information in human minds. Manuscript material is important because it is in character midway between the unrecorded mental knowledge and the more public knowledge of print and artifacts. Manuscript material is important because, at best, it is created in the heat of the event or very soon afterwards.

A historian facing his sources (books, artifacts, manuscripts, and oral tradition) does not recreate the past. That is an impossibility. What he *faces* is an abstract of the past and often an abstract of only a fragment of the past. What he *does* is to deduce the relationship of his abstract to the entire past he is investigating. To some

extent what occurs is similar to the work of the archeologist or the historical geologist. Each bit of source material and each collection for a given period or topic must be identified, their relationships clarified, and their place in the particular layer or stratum carefully noted and analyzed. From the surviving fragments one deduces the nature of a past culture or geological era. If manuscripts (or pre-historic artifacts) are removed from their place of discovery and split up without careful identifying marks, the deduction of the whole is much more difficult.

Collections are therefore ideally maintained as entities, their place of discovery and how they reached that location are noted if possible, and their general features are carefully determined. Of these the most important are names and dates. By names I mean two things: the name of the collection as a whole and the names that occur within the collection, let us say as correspondents. But names of collections need not concern us, though styles vary considerably. Most libraries and historical societies emphasize names in descriptions, apparently under the assumption that thus they are largely discharging their obligations to provide subject information. Dates given are usually the inclusive dates.

The importance of names and dates to the analysis of the historian is that they provide fixed points, as it were, from which to start investigations. (Here I refer to names of corporate entities as well as natural persons.) They are the "statics" of the historian's system. They are also the easiest information to provide, especially in the age of the typewriter. Provided that funds are forthcoming, I see no reason why the names and dates cannot be handled by nonprofessionals and perhaps coded on punchcards to produce lists and indexes. If a more general treatment of dates is required, as in a general catalog entry or a series entry, the problem is to indicate the inclusive dates *and* the especially rich periods or the gaps. A general treatment of names in a collection or series involves indicating *all* important correspondents whose documents appear frequently. We can adopt a convention of listing all individual names if a minimum number of items is present and any other *known* items for significant persons. Obviously, the larger the body of manuscripts, the less the probability of spotting and mentioning any one important document.

Once the statics of the collection are taken care of, we can turn to the "dynamics," the subject other than names that determine why and how a particular manuscript came into existence at a given time and place and what role that manuscript played in time. I am well aware that many librarians and archivists will regard my attitude

towards names and dates as a cavalier dismissal. Yet all too often a description of manuscripts will simply list a few names and engage in such begging descriptions as, say, that papers of an organic chemist relate to organic chemistry or that papers of the head of a university science department relate to the administration of that department. What is needed is concrete content information, extracted and presented systematically. Unlike the listing of names and dates, this is not a task to be delegated to the subprofessional, nor is it at present amenable to any significant mechanization. Until we can have machines to read and critically summarize manuscripts, we must pick up papers in our hands, master their contents, and write descriptions.

What I envision as a possible procedure is this. A particular body of records is first described in terms of its physical attributes and in generalities as to contents, according to our present practice. Depending upon the nature of the collection, we next select a portion or portions as samples and describe in great detail the contents of all documents in the sample. The samples, obviously, must be representative of the whole and their nature must be clearly specified. In a chronologically arranged body of manuscripts we might select three years (early, middle, and late) ; in an alphabetically arranged collection we could take any letter or letters in the alphabet; or in either of these we might simply take a random sample such as every hundredth document or folder.

The possible merit of this procedure is that the description will not be of those subjects that are of interest to the describer or that strike his eye during the examination of the records. But this merit is lost if our describer uses contemporary terms that slant his entries. So far as possible the detailed descriptions of the samples should use the terms employed in the documents themselves. To complete our description of the series we can next note the names of correspondents present at or above the minimum frequency and any known unique items or features of the body of records. The result can be described as a beefed-up series entry or an attenuated calendar.

The procedure outlined above does not obviate the need for more general descriptions, nor does it preclude the preparation of item-by-item descriptions for particular collections suitable for that expensive treatment. An institution that prepares inventories or registers may elect to use the proposed procedure only on large, complex series of special importance.

To return to our original assumptions: If all collections are in suitable institutions and are so described, we can arrange all the entries in such a manner as to approximate an abstract of the entire

period covered. This, while valuable, would not eliminate the need for two further steps. First, similar terms in the sample descriptions could be gathered and subsumed under general terms current today to produce something like a thesaurus. Someone interested in organic chemistry, for example, could check all the terms appearing under that or related designation to get leads to possible source collections. Similarly, all the entries for collections with descriptive terms falling under organic chemistry can be published separately as a guide to the sources for the history of that field. Or since they are abstracts of abstracts, these entries will also serve as the starting point for a more detailed guide. One great virtue of any system of analysis and description is the opportunity afforded for the preparation of subject guides. This conference might very well consider the possibility of sponsoring a guide or guides to all known collections of manuscripts relating to the history of science.

All of this is quite theoretical in view of the general low level of support for libraries, archives, and historical societies in the light of the great quantities of manuscript material awaiting description. Where possible, I hope that graduate students and scholars using manuscripts will write descriptions of these in their "Notes on Sources" as an aid to the manuscript custodians. Nor is my proposed procedure advanced as a panacea. It is, however, one possible approach to the solution of a most difficult problem. What is needed badly in the manuscripts field is some bold but hardheaded experimentation in new techniques to supplement and replace traditional ways that derive from a far different situation.

Since the item-by-item treatment has died without issue the need for such experimentation has become acute. For records of the past, relative scarcity made the item-by-item treatment appear worth while. The great mass of modern records precludes item-by-item treatment but does not obviate the need for careful, successful mass descriptions. Modern records, if not properly identified and described, will simply become a bulky albatross around the historian's neck. Research in the history of modern science requires that the scholar have a fair chance to find all possible sources and to select from the many bodies of records those that are most pertinent to a well conceived and designed research task. Beyond the needs of the historian there is another compelling reason for extensive and intensive analysis and description of manuscript sources for the history of science. Our period is one of the great eras in the history of science and in the significance of science for other aspects of our culture. Our achievements put us under obligation to preserve in usable form the primary sources of this notable chapter of man's quest for knowledge and control of his environment.