CHAPTER EIGHT

Documenting Documentation*

Taking exception to the premises underlying the principles and rules for archival description promulgated by the International Council on Archives, this chapter proposes alternative principles for documenting documentation. They have emerged from the collective activity of many archivists in the U.S. over the last decade. Documentation, which should begin close to the moment of records creation, retains contextual information about the activity that generated the records, the organizations and individuals who used the records, and the purposes to which the records were put. Information systems must be designed to retain sufficient contextual data to support archival management throughout the records' life cycles. User requirements must be considered so that a user can enter the system from knowledge of the world being documented without knowing about the world of documentation.

* Originally published in *Archivaria* 34 (Summer 1992): 33-49. An earlier version of this article, entitled "Description Standards Revisited," was presented at the Australian Society of Archivists annual meeting, Sydney, June 1991. The author wishes to thank Richard Cox, Richard Szary, Vicki Walch, and Lisa Weber for their helpful suggestions.

INTRODUCTION

An Ad Hoc Commission of the International Council on Archives proposed principles and rules for archival description in 1992.[1] Unfortunately the particular principles and rules fall short of what is needed.[2] This chapter elaborates alternative principles for documenting documentation which have been emerging from the collective activity of many U.S. archivists over the past decade but which have not been presented in one place before.

The "Statement of Principles Regarding Archival Description" (referred to as the ICA Principles throughout this text) and "Draft General International Standard Archival Description" rules (referred to throughout as ISAD(G)), circulated for comment by the ICA Ad Hoc Commission on Descriptive Standards in 1992, each consist of statements of differing degrees of generality which might be considered either principles or as rules for archival description. The ICA Principles reflect existing methods of archival description (at least in North America), while those suggested here for documenting documentation have not yet been developed, widely accepted, or even completely elaborated.

They are advanced here in part because the ICA Principles rationalize existing practice which the author believes as a practical matter that we cannot afford, which fail to provide direct access for most archives users, and which do not support the day-to-day information requirements of archivists themselves.[3] They are also advanced because of three more theoretical differences with the ICA Principles:

> (1) in focusing on description rather than documentation they overlook the most salient characteristic of archival records: their status as evidence;
>
> (2) in proposing specific content they are informed by bibliographic tradition rather than by concrete analysis of the way in which information is used in archives; and

(3) in promoting data value standardization without identifying criteria or principles to identify appropriate language or structural links between the objects represented by such terms, they fail to adequately recognize that the data representation rules they propose reflect only one particular, and limiting, implementation.

The principles for documenting documentation derive directly from the relationship of documentation to historical activity. The rules for data content and data representation which flow from them support ancillary principles which state that the purpose of recording information (description) is to support archives administration collections and serve the needs of users.

Before discussing the historical background for the documenting documentation principles and examining in detail their implications for each of these three points, let us briefly examine some distinctions which will be central to the discussion which follows.

First, how does description differ from documentation?

Description is focused on records both as the object being described and as the primary source of information. It seeks to characterize archival materials by constructing a document or collection surrogate. These surrogates, called cataloging records, finding aids, or archival inventories, each represent a "unit of material" or physical records. In archival description systems, these surrogates will be the fundamental record type or central file to which all indexes point.

Documentation is focused on activity in the records-generating institution, or activity of the creator of the records in the case of manuscripts, as the object being documented and as the preferred source of information. It seeks to capture data about the relationship between the activity and the document created or received in that activity which is necessary in order for the document to serve as evidence. Documentation results in the construction of systems with links between databases of activity and databases of documentary materials (archives)

created by, for, or of an activity. In documentary information systems, both the activity and the documentary materials documentation will be physically represented in numerous files; there will be links representing relations among them but no preferred view at the "center" of the data model.

Archives are themselves documentation. Hence I speak here of documenting documentation as a process whose objective is to construct a value-added representation of archives. This is accomplished by means of strategic information capture and recording into carefully structured data and information access systems as a mechanism to satisfy identified information needs of users, including archivists. Documentation principles lead to methods and practices which involve archivists at the point, and often at the time, of records creation. In contrast, archival description, as described in the ICA Principles is "concerned with the formal process of description after the archival material has been arranged and the units or entities to be described have been determined." (1.7) I believe documentation principles will be more effective, more efficient, and provide archivists with a higher stature in their organizations than the post-accessioning description principles proposed by the ICA.

Second, how does the content of an archival description differ from that of documentation?

The data elements of archival descriptions are an amalgam of what archivists have described in the past and those attributes of documents (fields in databases) which are defined by closely allied information professionals such as librarians. The content standards for documentation, on the other hand, are dictated by the principle that the information in documentation systems must support the requirements for archives to be evidence as well as serving internal management and end-user access needs of archives.

These two critiques are independent. The ICA Principles could have been focused on description of documents rather than on documenting context, yet have justified their concrete

content by reference to the value which specific descriptive data has for archival practice or in the support of access to archival materials. After all, archives are not ends in themselves but have as their purpose the preservation and retrieval of evidence of the past which has continuing value to the present. Description standards proposed by archivists might have advanced the principle that information recorded in archival descriptions should support the needs of managing such holdings.

Third, how do the data values in archival descriptions differ from those of documentation?

Although the ICA Principles say that one of their purposes is to "facilitate the retrieval and exchange of information about archival material" (1.3) and that "the structure and content representations of archival material should facilitate information retrieval" (5.1), they in fact advance a set of rules for data content and values, ISAD(G), which make sense only within a particular, if unarticulated, implementation framework. These rules dictate the construction of a specific type of output product (basically a "cataloging record"), probably intended for constructing union catalogs by international data interchange. The more general principles advanced here for documenting documentation recognize that rules for data values in documentation should derive from user needs and that the issue of control over data values is an implementation concern in a local system or an explicit service requirement of a concrete data interchange. Unlike the ISAD(G) rules, however, the principles for documenting documentation do not presume any specific information products or interchange purposes.

During the 1980s, the author and many of his colleagues, hesitantly and incompletely identified many of these distinctions without precisely locating their bearing on archival description because these implications, frankly, were not yet evident. An article on the power of provenance examined the possibility of structuring archival information systems around

documentation and introduced the limitations of the concept of fonds and the reasons for preferring series-level description, but in this respect it only proposed to refocus archival description.[4] Articles on the use of archival descriptions noted that user access begins with the researchers knowledge of the context in which the activity generating records took place, but failed to connect this concretely to principles for data content or to the need to redesign archival information systems.[5] Proposals that the concrete requirements of information interchange between institutions should dictate data content suggested how different content served different requirements within and among institutions. They also advanced the principle that the content of interchanges should follow from what are now known as "service requirements," but did not extend the reasoning to information systems nor to description end-products of individual archives.[6] And a preliminary report of end user "presentation language," undertaken to help define data representation for new archival information systems, did not examine how findings of such studies could or should be reflected in the capabilities of information systems.[7] An analysis of the research literature indicating the limitations of controlled vocabularies and suggesting sources of content that might be appropriate to archives did not explore the data structure of the overall documentation system that might support such access points.[8] This article will not present a comprehensive system design nor provide data to test its efficacy, but it hopes to lay out an integrated theoretical framework for documenting documentation and contrast it as necessary for its understanding with principles advanced for archival description.

DESCRIPTION OF ARCHIVES AND MANUSCRIPTS IN THE 1980S

After several decades of stability in which description meant making inventories, archival description sparked a renewed interest among North American archivists in the 1980s.

In the U.S., the current wave of professional interest in description practice grew out of an interest in building national databases of archival resources,[9] in a specific information interchange. Neither the USMARC Format for Archival and Manuscripts Control (USMARC AMC) data content standards, nor the *APPM* data value standards which are accepted by the U.S. archival community today were created in order to prescribe archival description principles. The National Information Systems Task Force (NISTF) explicitly described its efforts as descriptive (and permissive) as opposed to prescriptive or normative. It sponsored the construction of the USMARC AMC format from a data element dictionary compiled by archivists which was based on data in their existing information systems at that time[10] but NISTF never considered endorsing the data content which its working group mapped to the MARC format. Nor would it have done so both because its Chairman and Director (and probably other members) were keenly aware of the inadequacy of the existing practice which that data dictionary reflected, and because they fervently hoped that the data content standard was a process -- not a product -- and would be extended over time to reflect use requirements.

Likewise, when Steve Hensen first set out to interpret Chapter 4 of the *Anglo-American Cataloguing Rules*, 2nd edition (*AACR 2*), he was not articulating principles but attempting to interpret rules which had been poorly applied to archives and manuscripts. As the Library of Congress manuscripts cataloger, Hensen had to use these newly adopted international rules. His publication, *Archives, Personal Papers and Manuscripts: A Cataloging Manual (APPM)*,[11] made it possible for archivists to follow *AACR 2* rules and ultimately to use the interpretation offered in creating data values in MARC AMC. In the first edition of *APPM* Hensen makes it clear that the effort did not propose description principles, even if it did show that the bibliographic description principles embodied in

AACR 2 could be "interpreted" to support a method of cataloging with which archivists could live.

Using MARC AMC and *APPM*, American archivists have been constructing a national database on the Research Libraries Information Network (RLIN) -- and to a lesser extent on OCLC, WLN, UTLAS, and other bibliographic utilities -- since 1984.[12] Building the RLIN database made them aware of how inconsistent their existing cataloging had been. Task forces within RLG, and informal working groups of the profession worked throughout the late 1980s to build the database and impose greater consistency on it.[13] But they had quite practical aims and did not attempt to define what archival description should be. A few exploratory departures from the existing content standards were attempted in order to share appraisal information and conservation advice and construct a more structured administrative history database, but these were not very successful for a variety of political and economic reasons, and possibly because they lacked adequate theoretical underpinning. In the United States, anyway, there is still no truly theoretical formulation of archival description principles that enjoys widespread adherence, in spite of the acceptance of rules for description in certain concrete application contexts.[14]

In Canada the profession has spent nearly a decade reviewing the entire area of archival description and has aimed since 1985 to build a theoretical foundation for description practice.[15] *Towards Descriptive Standards* defined archival description by reference to three of the four major functions of library description: bibliographic description, the choice of non-subject access points, and subject indexing (leaving classification aside) and the *Rules for Archival Description (RAD)* effort has accepted this framework since. Like the ICA Commission, and the majority of U.S. archivists, the reports of the *RAD* working groups assume that archival description is an activity that takes place in archives, discrete from records creation and

records management, after records have been appraised, acquired, and accessioned archivists.

Throughout these deliberations over the past decade, a number of active participants have felt that all was not well, and certainly not adequate, with existing description standards and standards development efforts. Their concerns arose from at least three independent sources.

First, the MARC AMC format and library bibliographic practices did not adequately reflect the importance of information concerning the people, organizations, and functions that generated records, and the MARC Authority Format did not support appropriate recording of such contexts and relations. Since the mid-1980s, however, efforts had been proposed and undertaken to expand the concept of authority control as it was implemented in MARC-based systems, in order to accommodate a broader vision of the archival information system. This would have consisted of a number of reference files in which the records description file was not privileged.[16] Informal and formal groups also tried to develop vocabularies for indexing records based on their cultural forms rather than their physical formats.[17] A way out of the impasse seemed to be to focus archival attention on the record series, a unit having direct relations to provenancial activity, rather than the fonds or record group, whose relation to provenance was more abstract.[18] However, archival description remained a records-centered activity, and the primary representation was a "unit of material." As a result, none of these proposals truly shifted the focus to a multi-pronged documentation approach which emphasized context of creation and would produce "poly-centric" databases.

Second, archivists found MARC content definitions inadequate to support the operational information needs of the archives, which was to be expected since they were developed to support information interchange in a service defined as a union list. The U.S. National Archives (NARA) -- even though it had contributed all the data it believed in 1983 that it would

want to interchange to the data dictionary which led to the MARC content designation -- rejected the use of MARC two years later because it did not contain elements of information required by NARA for interchange within its own information systems.[19] Others built extensions to MARC records to accommodate local requirements, but did not draw the conclusion that MARC AMC data content must have been designed to satisfy a certain limited, but unarticulated, interchange service requirement, or that other models of what interchange could do, and other formats for description, or an extension of the existing format, would need to be accommodated by any theoretical framework intended to support documentation. This point was made before the MARC AMC format was proposed, but archivists failed to understand then, just as the ISAD(G) standard fails to do now, that rules for content and data representation make sense in the context of the purposes of concrete exchanges or implementations, not in the abstract, and that different rules or standards for end-products may derive from the same principles.[20]

Third, archivists began to have serious doubts about the use of existing descriptions for access to archives. Analysis of cataloging products has revealed widely divergent practices.[21] Some archivists believed that not enough was known about the way in which users sought information in archives to guide in the design of archival information systems.[22] Others felt that the solution to access would be to adopt controlled vocabularies and assign them to indexed fields. After a meeting of the Committee on Archival Information Exchange of the SAA was confronted with proposals to adopt many different vocabularies for a variety of different data elements, a group of archivists who were deeply involved in standards and description efforts within the Society of American Archivists formed an ad hoc Working Group on Standards for Archival Description (WGSAD) to identify what types of standards were needed to promote better description practices. Because WGSAD recently reported on its work in two full issues of the

American Archivist,[23] I need not summarize their actions but shall again make a few observations that will be developed more fully later.

WGSAD employed a matrix of types of standards to help it conceptualize, and then identify, standards relating to archival description which could or might already exist, but whose utility was not known to archivists.[24] It discovered numerous instances of standards that might be helpful to archivists, identified areas in which standards already existed, and located some areas in which standards might potentially be developed. During the course of its deliberations, WGSAD concluded that existing standards are especially inadequate to guide practice in documenting contexts of creation. WGSAD called for additional research in three areas of greatest interest to archivists which were considered the least well-developed: (1) the documentation of the context of creation of records (recommendation 15); (2) the capture and representation of information about records and acquisitions-related activity that is required for management of archives (recommendation 13); and (3) the analysis of user requirements (recommendation 12).[25]

These three research programs were intended to establish the foundation for sound archival documentation theory, although WGSAD did not express it in those terms. Since then, considerable progress has been made in developing frameworks for documentation, archival information systems architecture and user requirement analysis, which have been identified here as the three legs on which the documenting documentation platform rests.

DOCUMENTATION VS. DESCRIPTION

Documentation of the activity which generates archival records, and to a lesser extent of that which generates manuscripts, is a fundamentally different process than description of records which are in hand. Documentation of organizational activity ought to begin long before records are

transferred to archives and may take place even before any records are created -- at the time when new functions are assigned to an organization. Documentation of manuscripts begins with the identification of collecting priorities, research on people, associations, and events, which played a role in history and might have generated records.

When it acquires a function, an organization establishes procedures for activities that will accomplish it and implements information systems to support it. If we understand these activities, procedures, and information systems, it is possible to identify records which will be created and their retention requirements before they are created because their evidential value and informational content is essentially predetermined. Documenting procedures and information systems is fundamental to the management of organizations; thus documentation of organizational missions, functions, and responsibilities and the way they are assigned at various levels of structure and reporting relationships within the organization, will be undertaken by the organizations themselves in their administrative control systems. Archivists can actively intervene through regulation and guidance to ensure that the data content and values depicting activities and functions are represented in a way that will make them useful for subsequent management and retrieval of the records resulting from these activities. This information, together with systems documentation, defines the immediate information system context out of which the records were generated, in which they are stored, and from which they were retrieved during their active life.

The creators of manuscripts do not generate self-documenting information systems nor do they respond to regulation, but the contexts in which they create and use records are nevertheless documentable independent of records description. Historical subjects generate records as a consequence of the relations they maintain during their lives, and these relations exist outside the records in a manner which is useful to

understanding manuscripts as evidence: by recognizing the relations which a person, informal association, or event had, we can identify the records which do and do not exist in a manuscript collection. Documentation thus sheds light on records which are not present, as well as providing independent avenues of associational references by which the remaining records can be accessed and understood.

Documentation of the link between data content and the context of the creation and use of records is essential if records (archives or manuscripts) are to have value as evidence. The importance of this link, and the need for active intervention by archivists in the contexts of record creation to ensure documentation, has become clearer as a consequence of trying to define strategies for documenting electronic records systems. In these environments it is clear that contextual documentation capabilities can be dramatically improved by having records managers actively intervene in systems design and implementation.[26] Recent reports have called for more study on how such documentation objectives can best be achieved and research is now under way.[27] But the benefits of proactive documentation of the contexts of records creation are not limited to electronic records; the National Archives of Canada revised its methods of scheduling in 1990 to ensure that such information about important records systems and contexts of records creation would be documented earlier.[28]

When documentation of the organizational, functional, and systems context of creation of records takes place close to the moment of creation, and is used by people who are intimately acquainted with the organization and its informational processes, the documentation is likely to be intellectually more valid and easier to obtain than a *post hoc* description process. It is also more likely to satisfy the needs of users who are in the first instance staff of the organization seeking documentation associated with activities and responsibilities of the organizations for which they work. Documentation of the context, independent of the records and before the records are actually

created, may be augmented at a later date by archivists analyzing the content of the records themselves and locating in them evidence of the way that the activity was conducted. However, as a principle, the primary source of information about the people and organizations which generate the records, and which have engaged in the transactions which the records document, should be the organizations, activities, and systems themselves.

If this documentation is created in the beginning, and the principles for data content and representation discussed below are followed, it will be useful for administrative control purposes such as assignment of responsibilities, establishment of contacts, determination of records disposition and negotiation of transfers during the pre-archival life-cycle of the records. Both the functions of the organization and the way it established its can or should be known before any records of the function are created. Records managers can schedule such records based on the nature of the activity, its importance to the organization, and the legal, fiscal, and operational need for evidence. Documentation of functions and of information systems can be conducted using information created by the organization in the course of its own activity and it can be used to ensure the transfer of records to archives and/or their destruction at appropriate times. It ensures that data about records which were destroyed as well as those which were preserved will be kept, and it takes advantage of the greater knowledge of records and the purposes and methods of day-to-day activity that exists closer to the events. Most importantly, archivists can actively intervene in systems that will not generate and/or will not retain information of archival significance if they document such functions and systems at the beginning of their active lives rather than long after they have ceased to function.

These principles apply equally, although differently, to manuscript collections. Ultimate end-users of archives and manuscripts are better served through the construction of full-

fledged, "context of creation" reference files, since they cannot know the characteristics of records created by an organization or a person (e.g., description), but they can know the life history of a person or the functions of an organization and seek records that document the relations and transactions which each conducted. In addition, users can know about the generic forms of material or types of cultural documents which they are seeking. In the parts of the information system devoted to recording contextual data, they can locate those organizations and functions which have particular legislated responsibilities associated with search terms relevant to their queries; identify people whose relations with each other, to events, and to organizations are of interest; and explore the forms of material which have data they require; and locate those forms within the systems documentation associated with the information systems metadata in the "context of creation" reference system.

Information systems which do not contain archival description can lead users to such records by documenting the persons and organizations which are affiliated with the contexts of records creation. In-depth study of the process by which queries to archival description systems are formulated has shown that users engage in just this sort of reasoning even if they are seeking to approach a system that does not support access by contextual documentation.[29]

In short, documentation of the three aspects of records creation contexts (activities, organizations and their functions, and information systems), together with representation of their relations, is essential to the concept of archives as evidence and is therefore a fundamental theoretical principle for documenting documentation. Documentation is a process that captures information about an activity which is relevant to locating evidence of that activity, and captures information about records that are useful to their ongoing management by the archival repository. The primary sources of information are the functions and information systems giving rise to the records. The principal activity of the archivist is the manip-

ulation of data for reference files that create richly linked structures among attributes of the records-generating context and which point to the underlying evidence or record.

DETERMINING THE DATA CONTENT OF DOCUMENTATION

When we assert that the focus of documentation should be representation of the characteristics of human activity which result in archives, the functions which these activities are intended to carry out, and the information systems which hold the records, we have not yet provided principles for determining the content of such knowledge representations. The basis for such data content standards is again found in the difference between archives and other documentary materials, in this case a difference in their processing. When we acquire, describe, classify, and catalog library bibliographic materials, our processes do not transform them, but when we accession, transfer, arrange, weed, document, and inventory archival materials, we change their character as well as enhance their evidential and informational value. The fact of processing, exhibiting, citing, publishing, and otherwise managing records becomes significant for their meaning as records, which is not true of library materials.

The location of such principles within the matrix framework adopted by the Working Group on Standards for Archival Description is identified as data content and data values guidelines; no standards were identified in those cells. Unfortunately, WGSAD did not elect to explore these cells further in the papers it commissioned from its members. Had they, a paper on data content and data values guidelines would have stated as a principle that content and data representation requirements ought to be derived from analysis of the uses to which such systems must be put and should satisfy the day-to-day information requirements of archivists who are the primary users of archives, and of researchers using archives for their primary evidential purposes.

The Working Group had covered this ground in its meetings and reached consensus on the potential utility of a logical data and process model of an archival information system as a basis on which specific content rules could be constructed. A prototype of such an data flow model was proposed by the author in 1982 for use by NISTF in developing its data dictionary.[30] When NISTF decided to take a pragmatic approach in using data already present in systems as a method of developing its dictionary, the process and data model was abandoned. As a consequence of discussions which began at the WGSAD meetings, a follow-up effort to define standards for content based on the principle that content and representation standards follow function in the archival information system is now nearing completion.[31]

Building on a model information systems architecture drafted by Richard Szary, Ted Weir, and myself in 1989, fifteen archivists involved in archival description standards efforts received funding from the NHPRC to complete the work. The resulting model defines the activities involved in the administration of an archives and the clusters of data (free text "notes" or groups of data elements describing an aspect of a particular entity and its relations) required as input to or control over each activity as well as the clusters of data produced by each process. As such, the model defines, at the level of data clusters rather than data elements, what the data contents of archival description systems must be in order for them to support each of the various activities involved in archival administration. The data clusters are defined at a level of granularity which does not specify representation of data elements because the model is intended as a logical model not as a physical, or implementation, schema. The principles on which this information architecture standard is constructed are nevertheless quite clear about how one would derive specific rules for actual implementations: the appropriate content and values for the data are derived from the requirements of the archival tasks into which and out of which this

data must flow; these tasks, of course, are specific to the local application or interchange service.

The business processes reflected in the Archival Information Systems Architecture model include those involved in administering the archival repository, establishing its policies, procedures, plans, projects, and actions, as well as activities involving records description, arrangement, shelving, copying, etc. The model also includes the management of information about the creation context, including documentation of activities and of the information systems generating and storing records in organizations that transfer materials to the archives. The Information Systems Architecture working group hopes that one of the benefits of the model will be to demonstrate how information acquired about the function, activity, and/or information system in the records creating organization, such as promises of confidentiality extended to clients, can affect archival management of the evidence of these activities, influencing appraisal, transfer terms, and conditions of access and use. These kinds of relationships make it clear why the representation of data needs to serve subsequent use. By showing clearly the paths information takes and the tasks which it is intended to support, the model can assist archivists to identify how the data should be recorded when they first encounter it.

This approach to the question of what information ought to comprise an archival description does not accidentally differ from that taken by the ICA Principles. It proceeds from the radically contrary principle that the information in an archival description should be what is required by an archives (and its users), and that the way the data is represented should be dictated by the subsequent uses of the data in the system, including requirements for linking the data in the archives with data about entities in the real world contained in other information systems. Both the definition of the data requirements and the concept that this approach should be employed to define standards for archival information systems will be proposed to the Society of American Archivists Committee on

Archival Information Interchange and Standards Board in the winter of 1992-3.

The ICA Commission proposes a principle by which archivists would select data content for archival descriptions, which is that "the structure and content of representations of archival material should facilitate information retrieval" (5.1). Unfortunately, it does not help us to understand how the Commission selected the twenty-five elements of information identified in its standard or how we could apply the principle to selection of additional content. It does, however, serve as a prelude to the question of which principles should guide archivists in choosing data values in their representations.

DOCUMENTING DOCUMENTATION FOR THE USER

Even a consistent model of what contextual documentation requires, and adequate principles for determining data content standards for archival information systems, would not constitute a fully sufficient theoretical framework for principles. The documenting documentation platform rests on three legs: the third is that the language of documentation systems should provide access by users from their point of departure, and that the structure of links made by users should be explicitly represented, so that users will understand the relationship between the records and the context of creation of which they are evidence. The need to ground our principles for data representation in the perspective of the user derives from a fundamental difference between consciously authored materials (books, articles, documentary or fiction films) and archival materials which are records *of* but not *about* activity. Consciously authored materials have a subject matter imposed on them by their authors, and they are rarely appropriate as research material for other topics. Archival records on the other hand shed their light more indirectly, answering not only such factual questions as what took place and who was involved but also more subjective ones such as why partici-

pants acted as they did or how the actions were recorded. Libraries have found that subject access based on titles, tables of contents, abstracts, indexes, and similar formal subject analysis by-products of publishing can support most bibliographic research, but the perspectives brought to materials by archival researchers are both more varied and likely to differ from those of the records creators.

We know too little about what information users of archival information systems are seeking, and how they articulate their requests, to formulate, develop, or select specific vocabularies for representation of the content of archival documentation. We must therefore follow statement of the principle with a call for further study of such language. As a preliminary step, archival repositories throughout the U.S. were invited to participate in a snapshot study of what the author called "user presentation language" in the spring of 1989.[32] This was probably the first systematic, multi-institutional study of what users asked of archives ever conducted. More thorough studies by Paul Conway, completed in 1992, and others should influence archival documentation in the future.[33]

Archivists do know, however, from studies of retrieval using controlled vocabulary, that the benefits of control are not derived from the limitation of terms assigned but from the association between terms in thesauri and headings lists which effectively expand the number of routes by which one can get to the terms used in descriptions.[34] We also know that the effectiveness of controlled vocabulary depends greatly on its implementation and the availability and effectiveness of alternative implementation strategies. Rather than asserting that systems should be implemented in any particular way, we can suggest that user language be accommodated as a means of access into documentation, locating the user in appropriate reference files which employ the terms they use or synonyms of those terms, and providing for search within and among such reference files.

The principle therefore requires that archivists build structures which link the terms suggested by users concerning functions, form of material, subject content, or records creator/recipient by semantic models, to a meaningful documentation framework. One of the purposes of the rules derived from this principle will be to construct representations of archives which will no longer always require archivists to be present as intermediaries in order to translate queries into the structures by which we represent archives. One failure of the standards of description currently employed is that only those with extensive experience in archives understand how to translate a question about information content into the name of the organization or person around whom a fonds would be created. In a study conducted in the mid-1980s of the information retrieval function at the U.S. National Archives, researchers found that archivists pursued a search logic in translating users' subject-based queries into terms reflecting the provenance of records that was in principle replicable by artificial intelligence. Unfortunately their structural representations of the logical relations of the data in the agency history reference files led them to believe that human intermediaries would be required to provide testimony about each specific records-creating context in order for the retrieval to be significantly assisted by artificial intelligence, so the system was not constructed.[35] A better representation of the knowledge which they acquired from question-negotiating reference archivists would have exposed commonalties between types of semantic links that would have permitted them to represent the knowledge of reference archivists about the process rather than the content of searches. Users, they would have found, need to approach an archive from numerous perspectives other than the name of the organization or person responsible for the creation of a fonds. By modeling the relationship between subject terms in organizational histories and personal biographies, functional terms in mission statements and descriptions of activities, and knowledge about forms of

material, they could have demonstrated how best to answer one of the major types of questions which they found. If the object of description at the National Archives had been the record series, as it is at many other archives, a "user interface" in which these relations among creation contexts, forms of material, and content was explicit, would have gone a long way towards enabling the user to query a system without an intermediary. A recent study of the users of the documentary heritage in the United States provides some fascinating data on the differences between the questions being asked by different researchers and the types of materials that would serve as an answer.[36] Further studies along these lines would enable us to model a variety of approaches to archives and develop representations of the documentation system that correspond to the mind set carried to the archives by its users.

The principle here is that the user should not only be able to employ terminology and perspective, which are natural, but should also be able to enter the system from knowledge of the world being documented without knowing about the world of documentation. Gaining access to the names of individuals through the names of groups with which they might have been affiliated, or events in which they might have participated, or transactions with government to which they were parties, requires that an individual's reference files (or knowledge bases) be maintained. Similarly, access by functions (e.g., licensing) or activities (e.g., public hearings) requires the maintenance of reference databases about organizations, their missions, functions, activities, and procedures. Users need to be able to enter the system through the historical context of activity, construct relations in that context, and then seek pointers down into documentation. This frees them from trying to imagine what records might have survived (documentation assists the user to establish the non-existence of records as well as their existence) or to fathom the way archivists might have described records which did survive.

Archival description, or documentation, should make sense to end users not just because the language of documentation corresponds to the terminology of end-users or because the end user is able to search in reference files in order to establish relations between entities that were involved in the creation and use of records. It also involves creating and constructing a model of the archives as an information system which users can maintain as an archetype and employ to navigate through the documentation archivists create.

Given an appropriate model of what an archives is, and how it relates to the society which it documents, the contents of archival documentation can be made accessible to everyday visitors to the reference facility along with description of the contents. This information, moreover, can be used in making judgments about archival appraisal and accessioning prior to the creation of any records by a new function, or their recording, filing, and management by the information system supporting that function.

Instead of asking users who created a document which they are seeking or what institution would have had custody over it, archivists need to be asking them what information they are seeking, so that they might go from the information they want to the forms of material in which such information is represented, and the activities that would have generated such forms or had occasion to capture such information. As Terry Cook has observed about case records, the value of such records to society lies in their ability to provide evidence of discrepancies between the "image" of the transaction promoted by the organization whose function it is, and the experience of the transaction by an individual who, in the case of governmental actions is a citizen.[37] For this we need to have information about the interaction, why it took place, how it was conducted, what information it elicited, how the organization viewed the information, how the client viewed the information, and what purposes the information would ultimately serve. The documentation of documentation, rather

than the name of the creator of the fonds, is the source of the information which we use to appraise such records, and the foundation of the means by which we will ultimately retrieve them.

CONCLUSIONS

While American archivists may seem, from an outsider's perspective, to have recently arrived at a consensus about archival description and its purposes, the illusion disguises a profound confusion. When they departed from the practices of Brooks and Schellenberg in order to develop means for the construction of union catalogs of archival holdings, American archivists were not defining new principles, but inventing a simple expedient. After several years of experience with the new system, serious criticisms of it were being leveled by the very people who had first devised it. These criticisms have been growing in intensity and focus since. In the past several years, a number of efforts to move beyond the consensus on archival cataloging have been launched, including the Working Group on Standards for Archival Description, the Information Systems Architecture Standards initiative, and studies of archival users and the language they use to query reference staff and information systems. Together these initiatives are suggesting three theoretical premises for the documentation of documentation:

> (1) The subject of the documentation is, first and foremost, the activity that generated the records, the organizations and individuals who used the records, and the purposes to which the records were put;
>
> (2) The content of the documentation must support requirements for archival management of records, and the representations of data should support life-cycle management of records; and

(3) The requirements of users of archives, especially their personal methods of inquiry, should determine the data values in documentation systems and guide archivists in presenting abstract models of their systems to users.

NOTES

[1] International Council on Archives, Ad Hoc Commission on Descriptive Standards, "Statement of Principles Regarding Archival Description," First Version Revised (February 1992). Also, International Council on Archives, Ad Hoc Commission on Descriptive Standards, "Draft ISAD(G), General International Standard Archival Description" (January 1992). Both documents were published in *Archivaria* 34 (Summer 1992). Nothing in this chapter is intended to suggest that the proposed standard is not an accurate reflection of archival description principles adhered to by most archivists today or to suggest that those involved in drafting the standard have not been responsive to previous critiques of their earlier draft. The proposed standard has undergone a legitimate development and review process which is, in fact, one of the reasons it conforms so well to what archivists presently believe. These principles uphold record-centered, post-accessioning description activity centered in archives rather than an activity centered documentation and ignore the structuring requirements of data representation dictated by the purposes to which the data will be put precisely because most archivists do.

[2] Because this chapter proposes a set of principles which can be contrasted with those of the ICA, the introduction elaborates on these differences. A detailed critique of the text of the ICA Principles and ISAD(G) rules, which at the time of this writing were still in a draft form, is contained in David Bearman, "ICA Principles for Archival Description," *Archives and Museum Informatics* 6:1 (Spring 1992): 20-21.

[3] This critique of methods on purely practical, rather than philosophical, grounds is developed further in David Bearman, *Archival Methods*, Archives and Museum Informatics Technical Report #9 (Pittsburgh: Archives and Museum Informatics, 1990), 28-38. For analyses of how such systems would be structured, why they will work, and how they can raise the profile of archivists within organizations see David Bearman, *Functional Requirements for Collections Management Systems*, Archival Informatics Technical Report #3 (Pittsburgh: Archives and Museum Informatics, 1987).

[4] David Bearman and Richard Lytle, "The Power of the Principle of Provenance," *Archivaria* 21 (1985): 14-27 was originally drafted and distributed to colleagues during the life of NISTF although not published for several years because we found colleagues so hostile to its ideas.

[5] David Bearman, "'Who about What' or 'From Whence, Why and How': Intellectual Access Approaches to Archives and Their Implications for National Archival Information Systems," in *Archives, Automation and Access*, ed. Peter Baskerville and Chad Gaffield (Victoria, British Columbia: University of Victoria, 1986), 39-47.

[6] David Bearman, *Towards National Information Systems for Archives and Manuscript Repositories: The NISTF Papers* (Chicago: Society of American Archivists, 1987); also Bearman, "Buildings as Structures, as Art, and as Dwellings: Data Exchange Issues in an Architectural Information Network," in *Databases in the Humanities and Social Sciences*, vol. 4, ed. Lawrence McCrank (Medford, New Jersey: Learned Information, 1989), 41-48.

[7] David Bearman, "User Presentation Language in Archives," *Archives and Museum Informatics* 3:4 (Winter 1990): 3-7.

[8] David Bearman, "Authority Control: Issues and Prospects," *American Archivist* 52 (Summer 1989): 286-299.

[9] Kathleen D. Roe, "From Archival Gothic to MARC Modern: Building Common Data Structures," *American Archivist* 53 (Winter 1990): 56-66. Also, Bureau of Canadian Archivists, *Toward Descriptive Standards: Report and Recommendations of the Canadian Working Group on Archival Descriptive Standards* (Ottawa: Bureau of Canadian Archivists, December 1985) cites archival automation as a driver in the search for standards in Canada.

[10] David Bearman, ed., "Data Elements used in Archives, Manuscripts and Record Repository Information Systems: A Dictionary of Standard Terminology," NISTF Report (Washington, D.C.: Society of American Archivists, October 1982) reprinted in Nancy Sahli ed., *MARC for Archives and Manuscripts: The AMC Format* (Chicago: Society of American Archivists, 1985).

[11] Steven L. Hensen, *Archives, Personal Papers, and Manuscripts: A Cataloging Manual for Archival Repositories, Historical Societies, and Manuscript Libraries* (Washington, D.C.: Library of Congress, 1983).

[12] David Bearman, "Archives and Manuscript Control with Bibliographic Utilities: Challenges and Opportunities," *American Archivist* 52 (Winter 1989): 26-39.

[13] Alden Monroe and Kathleen Roe, "What's the Purpose?: Functional Access to Archival Records," in *Beyond the Book: Extending MARC for Subject Access*, ed. Toni Petersen and Pat Molholt (Boston: G.K. Hall, 1990); Marion Matters, "Authority Work for Transitional

Catalogs," in *Describing Archival Materials: The Use of the MARC AMC Format*, ed. Richard P. Smiraglia (New York: Haworth Press, 1990), 91-115 [also published as *Cataloging & Classification Quarterly* 11:3/4 (1990)]. Also see Research Libraries Group Government Records Project, "Online Record Types for Government Records," unpublished draft, July 1990.

[14] Steven L. Hensen, *Archives, Personal Papers, and Manuscripts: A Cataloging Manual for Archival Repositories, Historical Societies, and Manuscript Libraries*, 2nd ed. (Chicago: Society of American Archivists, 1989) is accepted by U.S. archivists as rules for applications involving data interchange of MARC records on national networks.

[15] Bureau of Canadian Archivists, *Toward Descriptive Standards* (Ottawa: BCA, 1992); also, Bureau of Canadian Archivists, Planning Committee on Descriptive Standards, *Rules for Archival Description* (Ottawa: Bureau of Canadian Archivists, 1990-). Not coincidentally, *Towards Descriptive Standards* envisioned an international standard congruent with the ISBD(G) which has now been produced as ISAD(G): General International Standard Archival Description with most of the same data categories identified in Appendix C of the 1985 report. (1st recommendation, p. 57).

[16] David Bearman and Richard Szary, "Beyond Authorized Headings: Authorities as Reference Files in a Multi-Disciplinary Setting," in *Authority Control Symposium*, Karen Muller ed. (Tucson, Arizona: Art Libraries of North America, 1987), 67-78; Lisa Weber, "The 'Other' MARC Formats: Authorities and Holdings, Do We Care To Be Partners in This Dance, Too?," *American Archivist* 53 (Winter 1990): 44-51; David Bearman, "Considerations in the Design of Art Scholarly Databases," *Library Trends* 37:2 (1988): 206-219.

[17] An informal working group was convened at the Smithsonian Institution in 1985 to draft a functions vocabulary. Work on a forms of material vocabulary went forward within the Research Libraries Group and the *Art and Architecture Thesaurus*, sometimes in parallel. A framework for the concept of using a form-of material as an access method appears in David Bearman and Peter Sigmond, "Explorations of Form of Material Authority Files by Dutch Archivists," *American Archivist* 50 (Spring 1987): 249-253. The AAT vocabulary was published as the "Document Types Hierarchy" in *Art and Architecture Thesaurus*, ed. Toni Petersen, (New York: Oxford University Press, 1990).

[18] Max Evans, "Authority Control: An Alternative to the Record Group Concept," *American Archivist* 49 (Summer 1986): 249-261; David Bearman, "Can MARC Accommodate Archives and Museums: Technical and Political Challenges," in *Beyond the Book: Extending MARC for Subject Access*, ed. Toni Petersen and Pat Molholt (Boston: G.K. Hall, 1990), 237-245; for a very early critique, see, Peter Scott, "The Record Group Concept: A Case for Abandonment," *American Archivist* 29 (1966): 493-504.

[19] William M. Holmes, Jr., Edie Hedlin, and Thomas E. Weir, Jr., "MARC and Life-Cycle Tracking at the National Archives: Project Final Report," *American Archivist* 49 (Fall 1986): 305-309; David Bearman, Letter to the Editor, *American Archivist* 49 (Winter 1986): 347-348; and response from Thomas Weir, *American Archivist* 50 (Spring 1987): 172-173.

[20] David Bearman, "Towards National Information Systems for Archives and Manuscript Repositories: I. Alternative Models," (August 1981), reprinted in *Towards National Information Systems for Archives and Manuscript Repositories: The NISTF Papers* (Chicago, Society of American Archivists, 1987); also David Bearman, "Archival and Bibliographic Information Networks" *Journal of Library Administration* 7:2/3 (1986): 99-110 [reprinted in *Archival and Library Administration: Divergent Traditions, Common Concerns*, ed. Lawrence McCrank (New York: Haworth Press, 1986)].

[21] Avra Michelson, "Archival Reference in the Age of Automation," *American Archivist* 50 (Spring 1987): 192-209.

[22] Lawrence Dowler, "The Role of Use in Defining Archival Practice and Principles: A Research Agenda for the Availability and Use of Records," *American Archivist* 51 (Winter/Spring 1988): 74-86, with commentaries by Jacqueline Goggin (pp. 87-90) and Anne Kenney (pp. 91-95).

[23] *American Archivist* 52 (Fall 1989); 53 (Winter 1990).

[24] Victoria Irons Walch, ed., "Report of the Working Group on Standards for Archival Description," *American Archivist* 52 (Fall 1989): 440-461; also David Bearman, "Description Standards: A Framework for Action," *American Archivist* 52 (Fall 1989): 514-519.

[25] Victoria Irons Walch, ed., "Recommendations of the Working Group on Standards for Archival Description," *American Archivist* 52 (Fall 1989): 462-477.

[26] David Bearman, "Management of Electronic Records: Issues and Guidelines," in United Nations Advisory Committee for Co-ordination of Information Systems, *Electronic Records Management Guidelines: A Manual for Policy Development and Implementation* (New York: United Nations, 1990), 17-70, 89-107, 135-189, parts of which are reprinted in this volume as Chapter 3.

[27] National Historical Publications and Records Commission, *Research Issues in Electronic Records: Report of the Working Meeting* (St. Paul: Minnesota Historical Society, 1991) defines the issues. See also: David Bearman, "Archival Principles and the Electronic Office" in *Information Handling in Offices and Archives*, Angelika Menne-Haritz ed. (New York: K.G. Saur, 1993): 177-193, reprinted in this volume as Chapter 5.; David Bearman, "Diplomatics, Weberian Bureaucracy, and the Management of Electronic Records in Europe and America," *American Archivist* 55 (Winter 1992): 168-180, reprinted in this volume as Chapter 9.

[28] National Archives of Canada, Government Records Branch, "Disposition of the Records of the Government of Canada: A Planned Approach," 3 July 1990, typescript.

[29] David Bearman, "Contexts of Creation and Dissemination as Approaches to Documents that Move and Speak," in *Documents that Move and Speak: Audiovisual Archives in the New Information Age*, proceedings of a Symposium held 30 April-3 May 1990 at the National Archives of Canada (New York: K.G. Saur, 1992), 140-149.

[30] David Bearman, "Functional Specifications of an Integrated Information Management System for Administering a Program of Active, Archival, or Manuscript Records," NISTF Report (Washington, D.C.: Society of American Archivists, August 1982). This was the precursor to the Bentley proposal.

[31] Marion Matters, "Building New Directions: The Development of the Archival Information Architecture," unpublished paper delivered at the Society of American Archivists annual conference, 1991.

[32] Bearman, "User Presentation Language in Archives."

[33] Paul Conway's studies of users conducted at the National Archives in 1990-91 have since been published as *Partners in Research: Improving Access to the Nation's Archives* (Pittsburgh: Archives and Museum Informatics, in press).

[34] Bearman, "Authority Control: Issues and Prospects."

[35] Daniel de Salvo and Jay Liebowitz, "The Application of an Expert System for Information Retrieval at the National Archives," *Telematics & Informatics* 3:1 (1986): 25-38; Avra Michelson, *Expert Systems Technology and its Implications for Archives*, NARA Technical Information Paper #9 (Washington, D.C.: National Archives, March 1991). For a critique of the de Salvo study, see David Bearman, "Expert Systems for Archives," unpublished paper delivered at the Mid Atlantic Regional Archives Conference, 8 May 1987; also, David Bearman, "Intelligent Artifices, Structures for Intellectual Control" in Bearman, *Archival Methods*, 49-58.

[36] Ann D. Gordon, *Using the Nation's Documentary Heritage: The Report of the Historical Documents Study* (Washington, D.C.: NHPRC and ACLS, 1992); see especially the multipart question #4 which is analyzed only superficially on pp. 46-48 of the report under the heading "framing research questions."

[37] Terry Cook, *The Archival Appraisal of Records Containing Personal Information: A RAMP Study with Guidelines* (Paris: Unesco General Information Programme, 1991); also review by David Bearman in *Archivaria* 34 (Summer 1992): 217-219.