Documenting Documentation

by **DAVID BEARMAN***

Introduction

An Ad Hoc Commission of the International Council on Archives has recently proposed principles and rules for archival description.[1] Unfortunately the particular principles and rules being proposed fall short of what is needed.[2] This paper 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 together before.

The "Statement of Principles Regarding Archival Description" (referred to as "ICA Principles" throughout this text) and "Draft General International Standard Archival Description" rules (referred to throughout as "ISAD(G)"), which have been circulated for comment by the ICA Ad Hoc Commission on Descriptive Standards, each consist of statements of differing degrees of generality which might be considered either as principles of 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 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] These alternatives 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 the bibliographic tradition rather than by concrete analysis of the way in which information is used in archives.

(3) In promoting data value standardization without identifying criteria or principles by which to identify appropriate language or structural links between the objects represented by such terms, they fail adequately to recognize that the data represen-

tation rules they propose reflect only one particular, and a 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 and serve the needs of users.

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

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 unit surrogate. These surrogates, called cataloguing records, finding aids or archival inventories 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 refer.

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 the objective of which is to construct a value-added representation of archives, by means of strategic information capture and recording into carefully structured data and information access systems, as a mechanism to satisfy the 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.

Secondly, 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 the internal management and end-user access needs of archives.

These critiques are independent of each other. The ICA Principles could have 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 have 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.

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

Although the ICA Principles state 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, framework of implementation. These rules dictate the construction of a specific type of end-product (basically a "cataloguing record"), probably intended for constructing union catalogues through 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 either 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 process.

During the 1980s, the author and many of his colleagues, hesitantly and incompletely identified many of these distinctions without precisely ascertaining their impact 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 adduced the limitations of the concept of fonds and the reasons for prefering 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 researcher's knowledge of the context in which the activity generating records took place, but failed to connect this directly to principles for data content or to the need to redesign archival information systems.[5] Proposals that the concrete requirements of information interchange among institutions should dictate data content suggested how different content served different requirements within and among institutions and advanced the principle that the content of interchange should follow from what are now known as "service requirements," but it did not extend the reasoning to information systems or to the description end-products of individual archives.[6] A preliminary report of end-user "presentation language," undertaken to help define data representation for new archival information systems, did not examine or indicate how findings of such studies could or should be reflected in the capabilities of information systems.[7] An analysis of the research literature exploring the limitations of controlled vocabularies and suggesting sources of content that might be appropriate to archives

did not indicate the data structure of the overall documentation system that might sup-
port such access points.[8] The present article will not provide a comprehensive system
design or data to test its efficacy, but rather aims to lay out an integrated theoretical
framework for documenting documentation, and contrast it as necessary for its under-
standing 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.
Other articles in this issue of *Archivaria* testify to that interest, and present views of
the official history of several description standardization efforts in the United States,
Britain and Canada.

In the United States the current wave of professional interest in description practice
grew out of an interest in building national databases of archival resources, in a specific
information interchange.[9] Neither the MARC-AMC data content standards nor the *APPM*
data value standards, which are accepted by the American archival community today,
were created in order to prescribe archival description principles. The National Infor-
mation Systems Task Force (NISTF) explicitly described its efforts as descriptive (and
permissive) as opposed to prescriptive or normative. It sponsored the construction of
the MARC Archives and Manuscripts Control format from a data element dictionary
proposed by archivists and based on data in their existing information systems at that
time.[10] It did not ever consider endorsing the data content which its working group mapped
to the MARC format. Nor would NISTF have done so, both because its Chairman and
Director (and probably other members) were keenly aware of the inadequacy of the exist-
ing practice which that data dictionary reflected, and because they fervently hoped that
the data content standard was a process—and not a product—and would be extended
over time to reflect user requirements.

Likewise, when Steven Hensen first set out to interpret the rules of *AACR2*, Chapter
4, 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
cataloguer, Hensen had to use these newly-adopted international rules. His publication,
Archives, Personal Papers and Manuscripts: A Cataloging Manual,[11] made it possible
for archivists to follow *AACR2* 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 *AACR22* could be "interpreted" so as
to support a method of cataloguing 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 cataloguing 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] They had
quite practical aims, however, 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 also because they lacked an adequate theoretical foundation. In the United States, in any case, 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 *RAD* effort has since accepted this framework. Like the ICA Commission, and the majority of American archivists, the *RAD* working group reports 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 by 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, corporate bodies 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 records series, a unit having direct relations to provenancial activity, rather than the fonds or record group, the relation of which to provenance was abstract.[18] However, so long as archival description remained a record-centered activity, and the primary representation was a "unit of material," none of these proposals truly shifted the focus to a multi-faceted documentation approach, which emphasized context of creation and would produce "poly-centric" databases.

Secondly, archivists found MARC content definitions inadequate to support the operational information needs of archives—which was to be expected, since they were developed to support information interchange in a service defined as a union list. The United States National Archives, even though it had contributed to the data dictionary which led to the MARC content designation, all the data which it believed in 1983 that it would want to interchange, 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, 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 actual exchanges or implementation, not in the abstract, and that different rules or standards for end-products may derive from the same principles.[20]

Thirdly, archivists began to have serious doubts about the use of existing descriptions for access to archives. Analysis of cataloguing 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 help guide 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 the Committee on Archival Information Exchange of the Society of American Archivists 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 SAA formed an Ad Hoc Working Group on Standards for Archival Description (WGSAD) to identify what types of standards were needed in order 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 the utility of which 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 the 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 the management of archives (recommendation 13), and

(3) the analysis of user requirements (recommendation 12).[25]

These three research agendas 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 requirements analysis, which have been identified 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 from the descrip-

tion 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 which 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 are 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 such 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 respond to regulation, but the contexts in which they create and use records are nevertheless documentable independently 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 fonds. Documentation thus sheds light on records which are not available, 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 creation and use of the 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 context of records creation in order to ensure documentation, has become clearer as a consequence of archivists' 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] The benefits of proactive documentation of the context of records creation, however, are not limited to electronic records; the National Archives of Canada has recently revised its methods of scheduling to ensure that such information about important records systems and contexts of records creation will be documented earlier.[28]

When documentation of the organizational, functional and systemic context of records creation 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 documen-
tation is likely to be intellectually more valid and easier to obtain than a *post hoc* descrip-
tion 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 responsi-
bilities of the organizations for which they work. Documentation of the context, indepen-
dent of the records and before the records are actually created, may be augmented at
a later date by archivists analysing 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 record, 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, determina-
tion 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 infor-
mation systems 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 impor-
tance to the organization, and the legal, fiscal and operational need for evidence.
Documentation of functions and of information systems can be conducted using infor-
mation created by the organization in the course of its own activity, and 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 exist 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 fonds. Ultimate
end-users of archives and manuscripts are better served through the construction of fully-
fledged, 'context of creation' reference files, since they cannot know the characteristics
of records created by an organization or a person (that is the function of description),
but they can know the biography of a person or the administrative history of an organi-
zation and seek records to document the relations and transactions which each conducted.
In addition, users can know about the generic forms of material or types of cultural docu-
ments which they are seeking. In the parts of the information system devoted to record-
ing contextual data, they can locate those organizations and functions which have par-
ticular legislated responsibilities associated with search terms relevant to their queries;
identify people whose relations with each other, events and organizations are of interest;
explore the forms of material which have the 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 con-
texts 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 reason-

ing, 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 source of information is the functions and information systems giving rise to the records, and the principal activity of the archivist is the manipulation 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 produce 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 catalogue library bibliographic materials, our processes do not transform them; 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 facts 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 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 a 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 supplementary 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 prototypal information systems architecture constructed 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 defined the activities involved in the administration of an archives and the clusters of data—free text "notes or groups of data elements which describe 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 not as a physical—or implementation—schema. The principles on which this information architecture standard is based are nevertheless quite clear about how one would derive specific rules for actual implementation: 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 management 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, accommodation, 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 maintaining 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 that it will 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 which 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 which 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 Committee in the winter of 1992-93.

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 the

selection of additional data 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 fictional 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 participants 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 reserach, but the perspectives brought to materials by archival researchers are both more varied and likely to differ from those of the records creators.

Archivists 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 representing the content of archival documentation. We must therefore follow statements of principle with a call for further study of such language. As a preliminary step, archival repositories throughout the United States were invited to participate in a 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 and others are just now being completed, which 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 authored headings 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 providing access to documentation, leading the user to 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, a 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 descriptions 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 of the information retrieval function at the American National Archives, conducted more than five years ago, researchers found that archivists pursued a search logic in translating user 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 data in the agency history reference files led conductors of the study to believe that human intermediaries would be required to provide testimony about each specific records-creating context, in order for 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 commonalities between types of semantic links that would have permitted the conductors of the study 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 the fonds. By comparing the relationships among 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 are found. If the object of description at the National Archives had been the records series, as it is at many other archives, a "user interface," in which these relationships 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 having recourse to an intermediary. A recent study of users of the American documentary heritage provides some fascinating data on the differences between the questions being asked by different researchers and the types of material that would serve as an answer.[36] Further studies along these lines would enable archivists to exhibit a variety of appoaches to archives, and develop representations of the documentation system that corresponds to the mindset carried to the archives by its users.

The principle here is that the user should not only be able to employ a terminology and a perspective which are natural, but also should be able to enter the system with a knowledge of the world being documented, without knowing about the world of documentation. Gaining access to the names of persons 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 person's reference files (or knowledge-bases) be maintained, just as 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 avenues down into the 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 how 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 among 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 which 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 together with description of the contents. This information, moreover, can be used in making judgements 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 people who created the document which they are seeking, or what institution would have had custody over it, archivists need to be asking the users 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 occasionto capture such information. As Terry Cook has recently observed about case files, 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 was elicited, how the organization viewed the information, how the client viewed the information, and the purposes to which information would ultimately be put. 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 shall 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 catalogues 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 levelled by the very people who had first devised it. These criticisms have since been growing in intensity and sharpening in focus. In the past several years a number of efforts to move beyond the consensus on archival cataloguing have been launched, including the Working Group on Standards for Archival Description, the Information Systems Architecture Standards initiative, and studies of archives users and the language they use to query reference staff and information systems. Together, these initiatives are suggesting three theoretical premises for documenting 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 the archival manage-
      ment of records, and the representations of data should support life cycle manage-
      ment of records.

(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

*   An earlier version of this article entitled, "Description Standards Revisited," was presented at the
      Australian Society of Archivists Annual Meeting, Sydney, in June 1991. The author wishes to thank
      Richard Cox, Richard Szary, Vicki Walch and Lisa Weber for their helpful suggestions regarding
      a more recent and much revised draft.

1   International Council on Archives. Ad Hoc Commission on Descriptive Standards, "Statement of
      Principles Regarding Archival Description, First Version Revised" [February 1992]. International
      Council on Archives. Ad Hoc Commission on Descriptive Standards, "Draft ISAD(G), General
      International Standard Archival Description" [January 1992]. (Both documents are published else-
      where in this issue of *Archivaria*.) Nothing in this paper is intended to suggest that the proposed
      standard is not an accurate reflection of archival description principles adhered to by most archivists
      today, or 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 why it conforms so well to what archivists currently
      believe. These principles uphold records-centered, post-accessioning, description activity focused
      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 so.

2   Because this paper 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 writing were still in draft form, is contained in David
      Bearman, "ICA Principles for Archival Description," *Archives and Museum Informatics* 6, no.
      1 (Spring 1992), pp. 20-21.

3   This critique of methods on purely practical, rather than philosophical, grounds is developed fur-
      ther in David Bearman, "Archival Methods," *Archives and Museum Informatics Technical Report*
      9 (Pittsburgh, 1990), pp. 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 Bear-
      man, "Functional Requirement for Collections Management Systems," *Archival Informatics Technical
      Reports* 3 (1987).

4   David Bearman and Richard Lytle, "The Power of the Principle of Provenance," *Archivaria* 21
      (Winter 1985-86), pp. 14-27. This paper was originally drafted and distributed to colleagues dur-
      ing the life of the NISTF, but not published for several years because the authors found their col-
      leagues 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 Peter Basker-
      ville and Chad Gaffield eds., *Archives, Automation and Access* (Victoria, 1986), pp. 39-47.

6   David Bearman, *Towards National Information Systems for Archives and Manuscript Repositories:
      The NISTF Papers* (Chicago, 1987), idem, "Buildings as Structures, as Art and as Dwellings: Data
      Exchange issues in an Architectural Information Network," in Lawrence McCrank, ed., *Data-
      bases in the Humanities and Social Sciences* 4 (Medford, 1989), pp. 41-48.

7   David Bearman, "User Presentation Language in Archives," *Archives and Museum Informatics*
      3, no. 4 (Winter 1990), pp. 3-7.

8   David Bearman, "Authority Control: Issues and Prospects," *American Archivist* 52, no. 3, (Summer 1989), pp. 286-299.

9   Kathleen D. Roe, "From Archival Gothic to MARCModern: Building Common Data Structures," *American Archivist* 53, no. 1 (Winter 1990), pp. 56-66; also, Bureau of Canadian Archivists, *Toward Descriptive Standards: Report and Recommendations of the Canadian Working Group on Archival Descriptive Standards* (Ottawa, 1985) which cites archival automation as driving 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, 1982); repr. in Nancy Sahli, ed., *MARC for Archives and Manuscripts: The AMC Format* (Chicago, 1985).

11   Steven L. Hensen, *Archives, Personal Papers, and Manuscripts: A Cataloging Manual for Archival Repositories, Historical Societies, and Manuscript Libraries* (Washington, 1983).

12   David Bearman, "Archives and Manuscript Control with Bibliographic Utilities: Challenges and Opportunities," *American Archivist* 52, no. 1 (Winter 1989), pp. 26-39.

13   Alden Monroe and Kathleen Roe, "What's the Purpose?: Functional Access to Archival Records," Toni Petersen and Pat Molholt, eds., *Beyond the Book: Extending MARC for Subject Access* (Boston, 1990); Marion Matters, "Authority Work for Transitional Catalogs," Richard P. Smiraglia, ed. "Describing Archival Materials: The Use of the MARC AMC Format," *Cataloging & Classification Quarterly* 11, no. 3/4 (1990), pp. 91-115; see also 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, 1989), is accepted by American archivists as a standard for applications involving data interchange of MARC records on national networks.

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

16   David Bearman and Richard Szary, "Beyond Authority Control: Authorities as Reference Files in a Multi-Disciplinary Setting," in Karen Markey, ed., *Authority Control Symposium* (Tucson, 1986), pp. 69-78; Szary; Lisa Weber, "The 'Other' MARC Formats: Authorities and Holdings. Do we care to be partners in this dance, too?" *American Archivist* 53 (Winter 1990), pp. 44-51; David Bearman, "Considerations in the Design of Art Scholarly Databases," *Library Trends* 37, no. 2 (1988), pp. 206-19.

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 concurrently. A framework for the concept of using 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, no. 2, (Spring, 1987), pp. 249-53. The AAT vocabulary was published as the "Document Types Hierarchy" in Toni Petersen, ed., *Art and Architecture Thesaurus* (Oxford, 1990).

18   Max Evans, "Authority Control: An Alternative to the Record Group Concept," *American Archivist* 50 (1986), pp. 249-61; David Bearman, "Can MARC Accommodate Archives and Museums: Technical and Political Challenges," in Toni Petersen and Pat Molholt, eds., *Beyond the Book: Extending MARC for Subject Access* (Boston, 1990), pp. 237-45; for a very early critique, see, Peter Scott, "The Record Group Concept: A Case for Abandonment," *American Archivist* 29 (1966), pp. 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 (1986) pp. 305-09; David Bearman, "Letter to the Editor," *American Archivist* 49 (1986), pp. 347-48; Thomas Weir's Response *American Archivist* 50 (1987), pp. 172-73.

20   David Bearman, *Towards National Information Systems for Archives and Manuscript Repositories: 1. Alternative Models* (Washington, 1981) [repr. in *Towards National Information Systems for Archives and Manuscript Repositories: The NISTF Papers* (Chicago, 1987)] David Bearman, "Archival and Bibliographic Information Networks," *Journal of Library Administration* 7, no. 2/3 (1986), pp. 99-110 [repr. in Lawrence McCrank, ed., *Archival and Library Administration: Divergent Traditions, Common Concerns* (New York, 1986)].

21   Avra Michelson, "Archival Reference in the Age of Automation," *American Archivist* 50, (1987), pp. 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 (1988), pp. 74-86; also commentaries by Jacquellin Goggin (pp. 87-90) and Anne Kenney (pp. 90-95).

23   *American Archivist* 52, no. 4 (Fall 1989) and 53, no. 1 (Winter 1990).

24   Vicki Irons Walch, ed., "Report of the Working Group on Standards for Archival Description," *American Archivist* 52 no. 4 (Fall 1989), pp. 440-61; also David Bearman, "Description Standards: A Framework for Action," *American Archivist* 52, no. 4 (Fall 1989), pp. 514-19.

25   Vicki Irons Walch, ed., "Recommendations of the Working Group on Standards for Archival Description," *American Archivist* 52, no. 4 (1989), pp. 462-77.

26   David Bearman, "Management of Electronic Records: Issues and Guidelines," United Nations Advisory Committee for Coordination of Information Systems, *Electronic Records Management Guidelines: A Manual for Policy Development and Implementation* (New York, 1990), pp. 17-70, 89-107, 135-89.

27   National Historical Publications and Records Commission, *Research Issues in Electronic Records: Report of the Working Meeting* (St.Paul, 1991) defines the issues. See also, David Bearman, "Archival Principles and the Electronic Office," *Proceedings of a Seminar on the Impact of Information Technology and Information Handling on Offices and Archives*, [in press: (Marburg, 1992)]; David Bearman, "Diplomatics, Weberian Bureaucracy and the Management of Electronic Records in Europe and America," *American Archivist* 55, no. 1 (Winter 1992).

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," *Documents that Move and Speak: Audiovisual Archives in the New Information Age. Proceedings of a Symposium held 30 April to 3 May 1990 at the National Archives of Canada* (New York, 1992), pp. 140-49.

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

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

32   Bearman, "User Presentation Language in Archives," pp. 3-7.

33   Paul Conway's studies of users, conducted at the National Archives in 1990-91, have not yet been published, but are the nearest to complete information available.

34   Bearman, "Authority Control: Issues and Prospects," pp. 286-299.

35   Daniel de Salvo and Jay Liebowitz, "The Application of an Expert System for Information Retrieval at the National Archives," *Telematics & Informatics* 3, no. 1, pp. 25-38; Avra Michelson, *Expert Systems Technology and its Implications for Archives* [NARA Technical Information Paper No. 9]. (Washington, 1991). For a critique of the de Salvo-Liebowitz study, see David Bearman, "Expert Systems for Archives," [MARAC conference paper, May 1987, unpublished]; see also, David Bearman, "Intelligent Artifices, Structures for Intellectual Control" *Archives and Museum Informatics* 6, No. 1 (Spring 1992), pp. 20-21.

36   Ann D. Gordon, *Using the Nation's Documentary Heritage: The Report of the Historical Documents Study* (Washington, 1992); see especially the multipart question 4, which is analysed 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, 1991); also, review by David Bearman elsewhere in this issue of *Archivaria.*