ISSUES IN INTEGRATIVE STUDIES
No. 10, pp.47-59 (1992)

The Archival Information System as a Model for Retrieval of Interdisciplinary Materials

by
Rutherford W. Witthus

The author suggests that the descriptive practices used to retrieve archival information might be a useful model to describe integrative processes and thereby create appropriate intellectual access to interdisciplinary materials.

SCHOLARS INVOLVED IN TRADITIONAL, discipline-based research are familiar with the type of access provided by libraries to collections of books and periodicals. A library catalog, which represents a library's collection of materials, is used to identify and locate items by author, title, topic, and subject classification number. In a similar manner, periodical literature is indexed by author, title, topic, and date. Electronic library information systems often include programs which allow researchers to search for materials by combining names, words, dates, numerical data, specialized codes, and phrases found in library catalogs and indexes.

Intellectual access to library materials is achieved through cataloging, classification, and indexing of each work. These activities proceed according to rules based on standards, which allow descriptive and subject information to be understood and shared within various local, national, and international information systems. The bibliographic community has a long history of creating intellectual access to materials based on description of the physical item and analysis of its contents. Although open to additions and improvements, bibliographic standards are well established and reasonably successful in providing intellectual access to published information because they depend on the vocabulary of the published literature itself to create access points.

Archival information systems as comprehensive as the bibliographic utilities are growing rapidly but are not yet universally available to scholars interested in using archival materials. Local repositories of archives and manuscripts have, for many years, taken the responsibility for creating intellectual access to their unique materials. Archivists arranged and described original source materials using practices which they themselves developed to best serve the clientele in their own repositories: they often created descriptive guides to individual collections for researchers to use on site; and, occasionally, they cataloged the collections to facilitate intellectual access on a local or national database. However, the adoption of recently-developed national and international descriptive standards based on the process of document creation combined with analyses of collection contents is increasing intellectual access to archival collections.

Intellectual access to interdisciplinary information is in its infancy. Fiscella, in one expression of the problems of access to interdisciplinary information, shows that scholars involved with interdisciplinary and integrative studies are able to use bibliographic databases (and, by extension, archival databases) to locate other scholars' work in well-defined interdisciplinary subjects, such as American, ethnic, and women's studies, and, somewhat less easily, in definitional and theoretical issues relating to interdisciplinarity itself. Locating scholarship which actually exhibits the integrative process at work, however, is extremely difficult, not only because the work is diffuse, but also because the work "is too idiosyncratic ... to have its own developed conversation, or literature" (Fiscella, p. 75).

This paper will elaborate on the differences between bibliographic and archival information systems, will endeavor to locate the authority for the creation of access points within the two systems, and will suggest a direction for building intellectual access to scholarly communication which exhibits integrative processes at work.

Intellectual Access to Printed Materials

With some exceptions, intellectual access to printed materials is based on standards of descriptive cataloging, subject cataloging, classification, and controlled vocabularies. A recent book on interdisciplinarity by Julie Klein will serve as an example to illustrate the discussion on the creation of intellectual access to bibliographic items. Descriptive standards permit catalogers to describe an item in a unified and consistent manner, which allows other libraries with the same item to share the work of the original cataloger. Descriptive rules prescribe using the title page of a book as the chief source of information for author, title, place of publication, publisher, and date of publication. Description also includes the number of pages, illustrations, presence of

a bibliography, and other pertinent notes about the physical item. Descriptive information is recorded in a bibliographic record (*Figure 1*), which contains standard elements in a prescribed order.

However, merely *describing* the physical characteristics of a book does not necessarily tell us what the book is *about*.

Figure 1

Klein, Julie Thompson.
Interdisciplinarity: history, theory, and practice / Julie Thompson Klein

Detroit : Wayne State University, 1990.
331 p. : ill. ; 23 cm.
Includes bibliographical references and index.

Interdisciplinary approach to knowledge.

BD255 .K54 1990

Subject catalogers read or browse a book to determine what the book is about. The topics are usually compared to a standard list of subject headings produced by the Library of Congress to determine which terms are authorized for use in a library catalog. In the case of the Klein book, *interdisciplinary approach to knowledge* is the most specific authorized subject heading available to describe what her book is about. Likewise, the form of name used for Julie Klein, i.e., Julie Thompson Klein, distinguishes her from any other author whose name may also be Julie Klein. While descriptive cataloging uncovers what is unique about an item, subject cataloging determines what the item has in common with other items in the bibliographic universe, so that someone seeking information on a topic can find multiple items on that topic. Descriptive and subject cataloging together create intellectual access to bibliographic items. Subject analysis also facilitates physical access: the Library of Congress classification scheme gives the book a subject-based address, BD255, so that it will be near other items on the topic *interdisciplinary approach to knowledge*, while, at the same time, it will be located in the general area of speculative philosophy, and more specifically, in the area of epistemology and theory of knowledge.

Although other controlled vocabularies are in use, the majority of academic libraries continue to use the Library of Congress name and subject headings as authoritative. As new authors emerge and as the body of knowledge grows, new terms are added to the authorized lists. These authorized lists are called authority files. Elements of a bibliographic record or fields in a database are said to be under authority control when a controlled vocabulary is used to standardize what is entered in those records or fields. Elaine Svenonius, who has published widely on controlled vocabularies, explains the purpose of authority control (1987):

> [A]ny data element in a bibliographic record that serves as an access point is a candidate for authority control. ... The reason for establishing unique names of authors and subjects and for making cross-references from variant names is to improve the chances that the user of the catalog will be able to find what he wants.. . . Catalog users should not be expected to guess the terms by which persons or subjects are named in a catalog: rather, the catalog should anticipate the language of its users. (pp. 1-2)

A bibliographic record, then, is a proxy for a book or a periodical, containing descriptive, analytic, and classificatory elements which include terms from controlled vocabularies. One major access mode, the subject heading, parallels the major organizing principle in a library, subject classification, which is based on the idea that books are usually about something.

Library science practitioners created bibliographic information systems based on the subjects of the various disciplines in order to organize and retrieve subject-related information. The locus of authority for the creation of intellectual access points to the literature of the disciplines is the vocabulary of the disciplines themselves.

Intellectual Access to Archival Materials

The search for archival materials proceeds somewhat differently. Archival materials, which include both archives and manuscripts, have traditionally been separated from other materials in libraries because they contain different kinds of information, they are created differently, they look and act differently, and they are organized differently from other library materials. The operating records of the narrow gauge divisions of the Denver & Rio Grande Western Railroad and the civil war letters of Uselma Duncan both have one characteristic in common with other library materials: they contain information. And not only do they contain information, they contain a very special kind of information, for example, first-hand accounts of

contemporary events, individual lives, business transactions, and policy decisions.

Archives and manuscripts differ from each other, however. Manuscripts generally refer to the papers of an individual or a family and may include such items as diaries, letters, speeches, and legal papers. A letter or an entry in a diary describes what is happening or has just happened or is about to happen. Letters, diaries, and other manuscript materials are rich sources of contemporary information about an individual or a family.

Archives, on the other hand, are "the preserved documentary records of a corporate body [or] a governmental agency ... that are the direct result of administrative or organizational activity of the originating body and that are maintained according to their original provenance" (Hensen, p. 9). Archival materials are seldom arranged by subject; rather, archival materials are arranged to reflect their organic source. The principle of provenance dictates that all records from a single source be kept together and that an attempt must be made to maintain or recreate the arrangement of records as they were in the office of origin. This is done for a good reason: "archival documents generally make sense only as part of a group of records. Record groups reflect the many activities which created them and may be useful for many subjects. ... The subject matter of individual documents can only be understood in the context of related documents created by the same activity" (Pugh, pp. 33-34). Respect for the original order of archival materials also "guarantees their essential integrity and historical accuracy and also preserves the evidential value inherent in the original grouping and ordering of materials" (Hensen, p. 4). An example will show the implications of archival arrangement for the researcher.

When a researcher comes to the archives seeking information on the buildings which originally stood on the site of the present Auraria Higher Education Center in Denver, Colorado, the archivist attempts to transform this subject query into terms relating to organizational activity. In other words, the archivist uses a process of inference to determine which organizational unit is most likely to have engaged in an activity that might have generated information on that particular topic. If the buildings were constructed during the first years of settlement, no building permits exist because none were required. However, the constitution of the Auraria Town Company specifies that ail construction in Auraria needed Board approval. The minutes of the Auraria Town Company, which operated under that constitution, contain petitions by various shareholders and property owners requesting approval of various types of structures on certain lots. Hotels, houses, saloons, ice houses, and other improvements are mentioned in the minutes, along with owners' names and building locations.

These records are held together by the fact that they were created in the course of business of the Auraria Town Company. The hundreds of names, the hundreds of locations, the types of buildings—none of these are named in the description of the collection. (*Figure 2*)

Figure 2

Auraria Town Company.
Records of the Auraria Town Company, from October 1856 to March 1860.

[S.l., s.n., 186?]
23 p. 29 cm.
Typewritten copy.

Auraria (Colorado)
Minutes

Yet we are able to locate this information by knowing that certain types of records—town company minutes—contain specific kinds of information—names of inhabitants and types of buildings. By knowing the activities of the company through its constitution, the researcher is able to locate information by knowing what might be included in various types of documents. The principle of provenance requires the archivist to maintain the Auraria Town Company minutes as a unit and in the order in which they were created, not to rearrange them according to shareholder, petitioner, or any other category.

Why don't archivists simply organize their materials by subject so that users can retrieve information from archives and libraries in similar ways? The great archives theorist, T.R. Schellenberg, advises the archivist to "resist any efforts on the part of scholars to induce him to arrange records according to any abstract system of universal subject classification" (p. 188). Because the information researchers want is generated by the activity that created the records, "retaining records in the order generated by the original activity allows access through analysis of function, a powerful mode of access" (Pugh, p. 34).

While cataloger-supplied subject analysis is useful, researchers in an archives also depend on functional analysis of the various forms of archival material: *why* do individuals write diaries; *why* do companies keep ledgers; *why* do organizations develop policy statements; *why* do people correspond with each other? Diaries are intended to record the personal reminiscences of an individual for that individual's own use. Letters are intended to convey an individual's thoughts and feelings to another person. Account books are intended to record and report information about the finances of an individual or organization. And, likewise, all types of materials found in an archives convey a particular activity or function of the creator of the records. Alden

N. Monroe and Kathleen D. Roe, in a chapter of a recent book on subject access to non-book materials, point out that traditional access points in a library or archives answer all but one of the basic journalistic questions: the *why*. "Function terms round out the picture by providing access to *why* records were created. The reason for records creation is too important to be left to guesswork or personal interpretation. The originating organization itself should provide the explanation" (Monroe & Roe, p. 160). Archives and manuscripts are the result of organizational or individual activity, but do not usually contain information about the activity that produced the documents.

How are these collections best represented to archival researchers and what are the appropriate access points? David Bearmau and Richard Lytle suggest that the principle of provenance is the key to information retrieval in an archives. In their paper, "The Power of the Principle of Provenance," they write that

> [p]rovenance information should be thought of as a means for providing access points to records in archival custody. In that respect, provenance information access points are the same in function as other kinds of access points such as chronological or geographical or subject information. . . .The . . . access point . . . is a characteristic which can be used in conjunction with other characteristics to identify a set of objects for examination. This applies equally whether the objects of retrieval are items in a grocery store, books in a library, or records in an archives. What differs is the appropriate characteristics—or, more precisely, which characteristics will prove most discriminating and most useful to searchers, (p. 21)

For example, retrieval of information from the Auraria Town Company records was facilitated by knowing the activities of the organization (i.e.. its functions) and the types of documents generated by those activities (i.e. form-of-material). If function and form-of-material are to act as intellectual access points for archival collections, then the data elements that describe function and form-of-material must be under authority control similar to that found in bibliographic information systems.

A major advancement in archival authority control came in 1990 with the publication of the *Art & Architecture Thesaurus*, which was funded by The Getty Art History Information Program. While devoted to the areas of art and architecture, the thesaurus includes a Document Types hierarchy for form-of-material. a People and Organizations hierarchy for occupations, and a Functions hierarchy "for activities which are conducted in order to accomplish specific purposes" (p. 324).

What data elements should a record in an archival information system contain? An archival record which includes both subject-based and provenance-based access points will serve as an example. (*Figure 3*)

The "title" of this collection, which is a single volume, is simply *Records*. When we refer to the *Thesaurus*, we find that records has a definite meaning, which is: "Recorded information, regardless of medium, created, received, and maintained by an agency, institution, organization, or individual in pursuance of its legal obligations or in the transaction of business" (*Art & Architecture Thesaurus*, p. 804). We find also that the bibliographic record contains a list of documents included in the collection: constitution, bylaws, regulations, and membership lists—all designations for *form-of-material*. The record informs us further that the Association organized the first lending library and reading room in Denver—an indication of the *function* of the Association. Imagine the kinds of questions this one small collection might help to answer: demographics of the reading public in Denver in 1860, small business and entrepreneurship in early Denver, and biographical information about Denver settlers. We gain access to the information contained in the collection not only through standard subject headings, but also by knowing the kinds of records generated by the activities of the organization. David Bearman (1986) contends that

> ... 'function' [is] a valuable proxy for accessing the information content of... materials. Moreover, types of records which are generated by any activity are culturally dependent and these 'forms of material' are also strong indicators of intellectual content. ... In addition, because archivists do not describe each item in the large bodies of records they accession, form of material and function continue to serve in place of content description of the materials for retrieval purposes." (p. 106)

Figure 3

Denver City and Auraria Reading-Room and Library Association
Records, 1860 [manuscript]
1 v.
Constitution and bylaws, regulations, list of officers and directors, and membership records of 99 subscribers. Includes typed index to the list of members and biographical notes about them, compiled by Rena Reese.

Source unknown.

First circulating library and reading room in Denver, Colo. Organized Feb. 10, 1860, in connection with the book, stationery, and news business of Arthur E. Pierce. Business later bought by David H. Moffat.

Presented to the library of the Chamber of Commerce and Board of Trade, May 22, 1890, by A.E. Pierce.

Moffat, David H.
Pierce, Arthur E.
Denver City and Auraria Reading-Room and Library Association.
Libraries, Subscription — Colorado — Denver.
Constitution.
Bylaws.
Membership lists.
Lending books.

While archival information systems may use library-type subject headings, they depend for effective retrieval on an understanding of the activities and functions of individuals and organizations and the forms of material generated by those activities. Archivists created archival information systems based on the order and context of archival records in order to organize and retrieve information generated by the activity that created the records. The locus of authority for the creation of intellectual access-points to the records of an organization or the papers of an individual is the activity of that organization or that individual.

Intellectual Access to Interdisciplinary Materials

Bibliographic information retrieval systems contain data elements representing what is unique about particular library materials (author, title, data) and what those materials have in common with other items in the library (subject headings). Archival information retrieval systems contain data elements representing what is unique about particular archival collections (provenance) and what those collections have in common with other collections in the archives (function, form-of-material, and subject headings). Could either of these information systems, or a combination of the two, serve as a model for more efficient retrieval of scholarly activity involving integrative processes?

Julie Klein describes integrative studies as a process:

> ... the dominant metaphor of a system is an organism. The organic metaphor enjoys great favor in the discourse because it establishes interdisciplinarity as a natural, generative process. It stresses evolution and fluctuation of knowledge rather than rigid architectural taxonomies and states of equilibrium. The image of an organism puts knowledge into 'live relationships,' emphasizing a fecundity that spawns new disciplines, (p. 80)
>
> The organic image assumes there are linkages which have been obscured and even damaged by *arbitrary* divisions. The belief that natural connecting forces will reestablish existing relationships is the dominant ideal of interdisciplinary discourse. (p. 84)

Archivists describe a similar process:

> A *practical* understanding must be gained of organizations as living cultures or organisms which create and use information, ... (Bearman & Lytle, p. 14)

> Archival materials are created as the natural byproduct, or record, of the activities or functions of persons or corporate bodies. Such materials are often said to be generated organically. (Hensen, p. 4)

The desire of those involved in integralive studies to maintain an organic unity of knowledge and of archivists to maintain an organic unity in their records is too similar to be ignored. Information retrieval in archives suffers when only a subject-based approach is imposed on the records; likewise, retrieval of interdisciplinary information suffers when only discipline-based vocabularies are imposed on the search. Archivists have discovered that retrieval based on access points which spring naturally from the activities which produced the records is the most fruitful. Are there access points which "spring naturally" from the results of integrative activity and which will facilitate access to interdisciplinary research?

We must not think that we will ever find the *subject* of interdisciplinary research, for as Klein points out: "Interdisciplinarity is neither a subject matter nor a body of content. It is a process for achieving an integrative synthesis . . ." (p. 188). Similarly, "archival material is of the activity that generates it, but seldom is it consciously authored to be about something. ... Archival materials are used to understand the contexts of their creation ..." (Bearman, 1989, p. 289). Is it possible to attach a process word or an integrative idea to a record in an information system that would allow an interdisciplinary researcher to locate not only the results of integrative studies, but also the integrative process itself? Archivists have discovered an integrated vocabulary based on the principle of provenance. Do interdisciplinary scholars have such a vocabulary or could an integrative vocabulary be developed? Svenonius (1986, p. 336) and others have concluded that the vocabulary of a discipline should *precede* vocabulary control for information retrieval. The emerging vocabulary of interdisciplinarity and integrative studies must be made explicit by its practitioners.

Klein (pp. 188-189), in a discussion of the integrative process, enumerates a series of steps, or activities, for achieving an integrative synthesis. The terms she uses, while general in nature, are similar to the functional terms in archival information systems: *defining, developing, specifying, resolving, collating, integrating, confirming,* and *deciding*. Add to this list such terms as *exhibiting, applying, illustrating*, and *exemplifying*, and we see the beginning of a vocabulary which might prove useful when attempting to retrieve the results of integrative processes. However, in order to be useful, these terms must be defined by those involved in the process itself, accepted as standard within the community of integrative scholars, and combined with other access points appropriate to the form of material.

The information retrieval process for interdisciplinary materials might also be improved by exploiting the power of inference. For example, the question *In what type of studies by what type of authors writing in what context would I most likely find interdisciplinary information?* might be more fruitful than attempting to find established subject headings reflecting interdisciplinary topics. Researchers might look at groups of studies, groups of authors, and various contexts rather than at individual items. The search for the needle in the haystack would become, instead, a careful look at the haystacks.

What kinds of information systems will be created to support integrative studies? Bearman and others are exploring the use of archival principles in the advent of scholarly or cultural databases, such as the prototype Smithsonian Institution Bibliographic Information System (SIBIS). Bearman (1986) says that

> [s]cholarly databases ... must provide for different disciplines to index information according to their own approaches—art historians being more concerned with geo-cultural entities like Flemish painters than geo-political entities like Dutch painters . . . [An] approach ... which does not assert the privileged position of any given set of standards or authorities, but requires the 'cataloging' organization to identify its approach, permits the construction of scholarly databases rather than just bibliographic databases. Improved capabilities are required actually to support such distinctive viewpoints. (pp. 107-108)

Archivists have assumed an important role in the creation of scholarly databases by insisting that archival authority files, which contain descriptions of the functions and activities of organizations and individuals, should occupy an expanded role in the search for information resources. In scholarly databases, provenance-based authority files become reference files, to be consulted for their information independently of their use to resolve problems of usage. No longer do authority files merely list terms for comparison. Rather, the authority file as reference file is used to answer and enhance user queries.

Perhaps with the building of scholarly databases, interdisciplinarians will have found a fertile ground to plant the results of their research. Archivists have defined their unique contribution as the organic principle of provenance and have based their organizational, access, and retrieval systems on that principle. The task facing those engaged in interdisciplinary studies is to help information specialists define the data elements which best represent the organic *process* of integration. When the process is defined in authoritative terms and the terms are linked to an information retrieval system, intellectual access to the unique contributions of those involved in integrative studies will be increased.

Biographical Note: Rutherford W. Witthus is Assistant Professor and Coordinator of Library Development and Special Collections at Auraria Lihrary—which serves the University of Colorado at Denver, Metropolitan State College of Denver, and Community College of Denver. He holds master's degrees in both Philosophy and Librarianship and Information Management. He has published articles on access to and instruction in original source materials.

References

Art & architecture thesaurus. (1990). New York: Oxford University Press.

Bearman. D.A. (1986). Archival and bibliographic information networks. *Journal of Library Administration, 7*(2/3), 99-110.

Bearman, D.A. (1989). Authority control issues and prospects. *American Archivist, 52*(3), 286-299.

Bearman, D.A., & Lytle, R.H. (1985-86). The power of the principle of provenance. *Archivaria, 21,*14-27.

Fiscella, Joan B. (1989). Access to interdisciplinary information: Setting the problem. *Issues in Integrative Studies, 7*, 73-92.

Hensen. S.L. (1989). *Archives, personal papers, and manuscripts: A cataloging manual for archival repositories, historical societies, and manuscript libraries. 2nd edition*. Chicago: Society of American Archivists.

Klein, J.T. (1990). *Interdisciplinarity: History, theory, and practice*. Detroit: Wayne State University Press.

Monroe, A.N., & Roe, K.D. (1990). What's the purpose? Functional access to archival records. In T. Petersen & P. Molholt (Eds.). *Beyond the book: Extending MARC for subject access* (pp. 157-170). Boston: G.K. Hall.

Pugh. M.J. (1982). The illusion of omniscience: Subject access and the reference archivist. *American Archivist, 45*, 33-44.

Schellenberg, T.R. (1956). *Modern archives: Principles and techniques*. Chicago: University of Chicago Press.

Svenonius, E. (1986). Unanswered questions in the design of controlled vocabularies. *Journal of the American Society for Information Science, 37*(5), 331-340.

Svenonius, E. (1987). Authority control: What and for whom. In K. Muller (Ed.), *Authority control symposium*. Occasional Papers of the Art Libraries Society of North America, No. 6, (pp. 1-8). Tucson, AZ: Art Libraries Society of North America.