Developing Intellectual Access and Control Mechanisms for Discipline-Based Virtual Libraries that Feature Media Integration

Lei Zeng
School of Library and Information Science
Kent State University
P.O. Box 5190
Kent OH 44242-0001

D. Kaye Gapen
Sharon Schmitt
Case Western Reserve University
10900 Euclid Avenue
Cleveland, OH 44106-7151

I. BACKGROUND, ASSUMPTIONS, AND DEFINITIONS

Introduction

On December 9, 1992, Case Western Reserve University Library began the work of creating its first virtual library in the geological sciences. Ten more discipline-based virtual libraries will be developed by the end of 1994. One of the most vital components of a virtual library is the intellectual access to information and knowledge resources in all formats, thus the control mechanisms enabling access become very important issues which challenge traditional approaches in the construction of classification schemes, thesauri, and other types of controlled vocabularies. This paper presents some basic and important considerations in the process of developing intellectual access and control mechanisms for discipline-based virtual libraries.

This paper explores the idea that a virtual library environment represents the future of information management, leading eventually to knowledge management. The paper discusses the possibilities for developing user-framed access and control mechanisms to enable the virtual library environment. The authors represent three distinct perspectives: a Director who has spent a career conceptualizing the future of information use and management including the articulation of what will be required to realize such a concept, a project manager trying to clarify the component parts of a new, media-intensive information environment, and a classification researcher and faculty member exploring applications of access and control mechanisms in this new environment. In a spirit of mutual respect, the authors have combined their questions, concerns, and comments in this paper. Our hope is that in articulating some of the issues surrounding virtual libraries, classification researchers whose work addresses the essential issues of access will be able to apply their expertise to these issues.

For this discussion, the term virtual library refers to an environment, an environment in which the client services aspect is the most commonly referred to and the most immediately relevant to the user. In our discussion, a virtual library environment is not access to some local or remote OPAC, nor is it access to the Internet or some specific listserver on the Internet. The client server component of a virtual library environment may offer all of the latter as part of client services, but

as a concept, a virtual library environment goes far beyond those notions. A virtual library environment is one in which component parts combine to provide intellectual and real access to information, the value of which is framed entirely from the users point of view, meeting the individuals unique information needs. A more detailed definition of the virtual library environment appears in the section: Definition of terms.

In a virtual library environment, we have no special interest in only recreating an electronic version of a documents warehouse. We are interested in intellectual access to materials on several levels. We are equally interested in control mechanisms to facilitate use of virtual libraries, also on several levels. These levels include: The initial query for information of some kind; a mechanism for accessing the query results and for continuing iterations of the query until the user is satisfied that the net cast out is sufficient in both depth and relevance; a mechanism for marking the location of relevant information; a mechanism for revealing the subject content or aboutness of the material; a mechanism for marking the relevant sections that will enable return to them by the user when necessary; a mechanism for creating additional or more complex links between chunks of located information; a mechanism for manipulating and modeling the located information; a mechanism that maps, for the user, a given set of linkages; a mechanism that permits navigation among the links; a mechanism to store and remember the linkages and locations of linked information, as well as information newly created by the user; and a mechanism for making user-created information available to queries from other users, if so desired.

The conceptualization of this environment includes the answers to questions such as the following. How will these access and control mechanisms be formed? How can they be integrated together? What principles, projects, methodologies, etc. from classification research can be used as guides to forming these mechanisms? What new paradigms and metaphors must be drawn to capture the ideal environment? What elements from the users information environment are sufficiently universal to be applicable to this development? What elements of classification theory will be critical in a virtual library? How will these elements apply when some of the entities important to the user will include full motion video, three dimensional modeling, musical scores, etc.? How many questions can be framed through research and analysis, using existing concepts, and how many can be framed only through vision and insight into possibilities that do not yet exist? These are the kinds of issues we wish to bring to the attention of classification researchers in search of their comments, questions, expertise, and solutions. The authors believe they have formed the right questions and they will discuss in this paper the issues relevant to the questions formulation.

Background

What follows is an abbreviated description of how one academic research library made a choice to pursue the development of discipline-based virtual libraries. A detailed account is included in Appendix A.

Much has been written about the changing nature of libraries, about their changing missions, about the impact of transforming informational technologies, and about shifting paradigms of service. No libraries feel change more strongly than academic research libraries whose clientele are members of the scholarly community and whose traditional needs are now matched by a growing set of new

demands occasioned by the infiltration of new and powerful computing tools that have changed and will continue to change the face of scholarly work, scholarly communication and information access.

The paradigm shifts are related to changes in the nature and quantity of information, especially information generated in an electronic format. The fact is that now all formats of recorded information (voice, video, graphics, sound, text, etc.) can be transmitted by means of one common process: digitization. The power of this concept is exponentially increased when it is coupled with the concept of powerful infrastructures like fiber optic networks that can deliver these digitized formats in real time across the barriers of time and space. Each new generation of computer hardware and software increases the capability of producing models that are more and more life-like, that make the computer as a tool more and more invisible. The promise of virtual reality, where the model is indistinguishable by the human senses from reality, hovers on the horizon, promising the ultimate in modeling capabilities.

One of the immediate ripple effects of these new digitized formats and their accompanying modeling capabilities is that the nature of access to knowledge must be redefined. Access, in a traditional academic research library, was once clearly defined as ownership. Access equaled ownership. If a library wished to make materials accessible to its scholarly community, it purchased, cataloged, shelved, and generally maintained the physical item. However, it is clear that with digitized information capable of transmission across vast computer networks, the nature of access must be redefined. By definition, for many academic research libraries, access is now defined as "intellectual access," meaning access to the information, wherever it resides, locally or remotely, whenever the user desires it. Thus, physical access and intellectual access, as constructs, must be transformed beyond ownership and the public catalog. In addition, access may also be extended to include the ability to manipulate parts of the accessible data in ways that will enhance the synthesis of knowledge, and by way of that synthesis, create new knowledge, that must in turn be managed and made accessible.

As the concept of access is redefined, the entire notion of how universities view their libraries, what services models are necessary, what scholars will require from the libraries in the future, also shift. Access must now include an idea not just of how information may be retrieved but how it must be structured, formatted, tagged, etc. so that it will have a successful growth path, remaining accessible to a variety of search engines designed to retrieve it.

In addition to the conceptual leaps required by shifting definitions of access, there is the sweeping scope and power of the informational technologies that have already and will continue to impact on the professions concerned with information and knowledge management, the nature of the skills required of those professions, the focus of those skills, and the organizational structure of any business, corporation, library, or research center in the twenty first century.

When one academic research institution set the stage to plan its new library facility, an entire process was set in motion that sought to capture and respond to shifting paradigms and forces of change. Beginning with a new Director, charged with envisioning the Library of the Future, the process moved through a series of analyses, both internal and external, that resulted in a decision to develop discipline-based virtual libraries for the campus scholarly community. This process

included the following key decision points: (see Appendix A for a detailed accounting of the steps and decisions involved.)

In establishing a new foundation for framing the library of the future, a decision was made to adopt the theoretical model proposed by Robert S. Taylor in his book, *Value-Added Processes in Information Systems* — a user-framed, value-added model[1] — and to use the Taylor model as an evaluative tool.[2] The Library staff adopted the position that the value added by any process, system, service, personnel, etc. would be determined from the standpoint of the user. The concept of developing information use environments for aggregate groups of users was adopted and a decision made to use detailed versions of those information use environments to drive virtual library design. A commitment, supported by the academic Deans, was made to conceptualize the components of the virtual library environment and move toward operationalizing those components.

A decision was made that the value-added model, which is non-predictive, would not be a tool for providing the vision of how people would learn, think, research, teach, and model differently. That vision would require individual investigation, insight, intuition, and risk and the responsibility would fall on the leadership of the campus, charged with envisioning the future and shaping planning efforts accordingly.

The Library Directors response to that challenge constituted an extension of two previous decades of thinking and writing: The Library of the Future will have a new mission and that mission will be knowledge management. Knowledge management is a phrase that attempts to describe an environment where an individual can be brought into contact with information in a way that is exponentially more productive and engaging than any currently available environment, an environment where clarity and understanding can be brought to data. The task of anyone associated with bringing this environment to fruition will be to form the right questions and create the right tools that will ultimately lead to the creation of an environment in which knowledge management is possible.

The conclusion of the Director of University Library is that the Virtual Library will be the enabling environment that will pave the way for knowledge management. This environment must, in theory, have all the mechanisms for intellectual access and control mechanisms that need to accompany any information-intensive environment, but because it is certain to include computer-based media technologies, it must reflect the transforming capabilities of information technologies.

---

1. Taylor, Robert S. *Value-added processes in information systems.* Norwood: Ablex Publishing Corporation, 1986. Taylor refers to his model as rather an early presentation of a way of thinking about the field of information science and also as a frame of reference for ordering what we know about information use environments It is a very complex, powerful, and sophisticated model. To mention some of its principle components only briefly, as we do in this paper, is to do the model and the book an injustice. Interested readers should examine the book for a true idea of the range of Taylor's thinking.

2. Taylor's model is a theoretical one and is not predictive. The University Library adapted it as a working tool and in turn, adopted the concepts of value-adding and the importance of the information use environments as critical guiding principles in the construction of the Library of the Future, and consequently, of the virtual library environment. Any misinterpretations or departures from Taylor's original definitions or descriptions of components can be attributed to University Library of CWRU.

Assumptions

- All future designs for intellectual access to data and control mechanisms for data should be user-framed and the evaluation of information services should primarily consider what value those services add for the client.
- A virtual library is an environment. It is dynamic, fluid, and is made up of component parts that will be assembled uniquely by groups and individuals according to their specific needs and preferences.
- The virtual library environment hinges on the principles of making information, in all its forms, optimally available to users who are characterized by unique information environments.
- When fully realized, the virtual library environment exists wherever it is needed by the user, whenever it is needed. The questions that will dictate its structure are applicable to any kind of data, including data traditionally found in scholarly communications and data rarely associated with libraries, academic or otherwise.
- Tools will be created to manipulate information in ways that add value on a person-by-person basis. Tools will be created to allow an individual to match their information gathering, structuring, storing, and manipulating capabilities to their information environments. These tools their development, evaluation, and subsequent re-design, will be an integral part of the virtual library.
- Tools will be created that will allow information to be manipulated and modeled in ways that dramatically increase or reveal the levels of meaning in the data, i.e., visually modeling data using time, motion, sound, sensation all in three dimensions.
- The effort to establish true virtual libraries will help point the way toward achieving knowledge management.
- While theories of access and control mechanisms may remain constant as they apply to a virtual library environment, the nature of access and control mechanisms will change. The most difficult part of developing these new mechanisms for an environment that does not yet exist will not be finding the solutions, but correctly framing and determining what questions need to be asked.
- We do not need to ask how informational technologies can improve and enhance what we are doing now in terms of access and control mechanisms. We need to ask how informational technologies can allow us to do what has not been readily available or reported before in terms of access and control mechanisms, perhaps allowing us to realize these mechanisms in entirely new ways.

Definition of Terms

Knowledge Management (KM): In the broadest sense, knowledge management is the processing of any data in such as way as to impart significant meaning or clarity beyond what is present at the time of the data's retrieval. A knowledge management environment is an environment which would facilitate such

processing. For example: A user retrieves a set of salinity readings for the Chesapeake Bay from a data stream accessible on the Internet. In a knowledge management environment, the user might chose to make the data more meaningful by using the data to visually model the impact of the salient readings on the Bay in a three dimensional picture that would accurately depict time, space, and provide color coding keyed to the salinity readings. The environment might also allow the user to link the data to other elements previously retrieved by the user, or bring up related elements, such as a map of the Bay and a recent report from a commission studying how to save the Chesapeake. In academic environments, (KM) is often used to refer to a model for scientific and scholarly communication in which faculty and research librarians share responsibility for the collection, structuring, representation, dissemination, and use of knowledge, using information technologies that feature media integration. In the paper, (KM) is usually referred to in the broader sense.

Virtual Library Environment: An environment in which component parts combine to provide access to information, the value of which is framed entirely from the users point of view and enables the individuals unique information needs. Virtual libraries are not a single entity, but a host of component parts brought together in a dynamic environment. Frequently, virtual libraries are also defined as the act of remote access to the contents and services of libraries and other information resources, combining an on-site collection of current and heavily used materials in print, microformats, and electronic form, with an electronic network which provides access to, and delivery from, external library and commercial information and knowledge sources worldwide. In essence, the faculty member and student are provided the effect of a library which is a synergy created by bringing together technologically the resources of many libraries, information services, and knowledge stores. (See Appendix D for a graphic depicting the components of the virtual library environment.)

Intellectual Access: All the tools and means to provide access to the attributes that describe an entity desired for retrieval. This means going beyond the quest for the best way to provide access to subject contents or the aboutness of an entity. Entities whose purposes or origins of creation and mode of construction differ widely from traditional academic materials may have significant attributes that go beyond subject access or include subject as only one of many significant attributes.

Control Mechanisms: All the tools and means provided to bring data into some form which can allow it to be maximally useful to users. Transforming technologies have, throughout history, initially outstripped the control mechanisms for dealing with them. Excellent discussions of these control phenomenons are available in James Beniger's book, *The Control Revolution.*[1] An example of the development

---

1. Beniger, James R. *The Control Revolution: Technological and Economic Origins of the Information Society.* Cambridge, MA: The Harvard University Press, 1986.

of control mechanisms from Beniger's book include those developed around the steam engines that created the national railroad system, e.g., the creation of time zones to allow for the creation and regulation of train schedules.

Information Use Environment: A critical concept within Robert S. Taylor's user-framed, value-added model, information use environments are essentially pictures of how individuals or aggregate user groups gather and use information. Integral to Information Use Environments is the acknowledgment that information has many traits and that those traits exist across a continuum of value. The approach also incorporates the view that users have problems to solve (not merely questions that require an answer) and that those problems also have dimension across a continuum of values. How closely any system design comes to matching the users information environment will determine how valuable that system is to the user. As a basic foundation, for the development of the virtual library environment, therefore, information use environments are the most difficult and most essential element in the user-framed, value-added model and the basis for development of the virtual library environment.

II. CONSIDERING THE VIRTUAL LIBRARY ENVIRONMENT (see Appendix D for a graphic depicting the component parts of the environment.)

Goals: We are paraphrasing Jesse Shera's goal for librarianship[1], in saying that we are striving to bring information (in all its forms) and users (with all their individual preferences and needs) together in the most significant, meaningful, and powerful way possible. The current strategy for reaching this goal is developing the virtual library environment and the mechanisms to support such an environment.

Elements: When we consider the development of the virtual library environment, we are concerned primarily with two elements: the materials in the information universe, and, the users needing to make use of those materials.

The materials may be in any format currently in existence and in some formats not yet known to exist. The materials will include existing materials that are relevant to the user, but materials will also include entities newly created by the user. In the virtual library environment, the user is not only a Gatherer and Consumer, but also a Modifier and a Creator.

---

1. The actual quote is: I am convinced that the role of the librarian in society…is to maximize the utilization of graphic records for the benefit of society. In other words, his function is to serve as the mediator between man and graphic records; not only books, but sound recordings, pictures, audiotapes, charts, whatever contributes to the advance of human knowledge…. The object of the library is to bring together human beings and recorded knowledge in as fruitful a relationship as it is humanly possible to be." Shera, Jesse H. *Sociological Foundations of Librarianship.* 1st ed., New York: Asia Publishing House, 1970.

The users need to facilitate their attempt to conceptually map their world. To do so, they periodically supplement their own knowledge with what is available in the information universe. This creates a dynamic subset of materials that are relevant to the user, and therein lies the value of that subset. In the virtual library environment value is added to every element related to the users interaction with their subsets, beginning with the creation of the subsets.

Perspective: In a moment, well look at the component parts that a virtual library environment would require in order to bring materials and users together to meet the above goal. Every project has a perspective, and ours is that the user-framed perspective frames the design of the virtual library environment. The means that the most important element is the user and the value of every other element in the design is evaluated by what value that element adds for the user. This does not mean that any of the supporting elements are unimportant. It means that all of the supporting elements, including, but not limited to the structuring of information resources, the design of retrieval mechanisms, the interface elements, and the tools for managing, manipulating, and modeling the data, must ultimately reflect the needs of the user. The designs will often contain complex, even convoluted solutions to the vast problems of access and control, but those solutions must appear seamless to the user.

If the most powerful contribution of personal computers (or any device possessing machine intelligence) is to enable individuals beyond anything previously possible, then the most powerful contribution of the virtual library environment will be to employ information technologies to facilitate the individual user within the framework of their information use environment.

Elements Making Up the Virtual Library Environment: As currently envisioned, a virtual library environment will need component parts that address: Planning and Design, Management and Administration, and Client Services. These three designated areas represent concepts only. The functionality of the three areas may actually be realized through any combination of human expertise, machine intelligence, and hardware and software components all of which will integrate and mature over time as more powerful technologies present more powerful solutions.

While the Client Services area is the most widely modeled component and the most directly relevant to our discussion, the three areas will not exist in separate strata, but will be bound together by the interdependent nature of their elements. They will be fluid and dynamic. Integration and flexibility are watchwords of the virtual library environment. When we discuss them separately, it is only one way to try and convey their functionality. By necessity such representation fails to demonstrate their dynamic nature, but in a text-based medium we have only static representations to use as models.

Details of the Three Areas:

The **Planning and Design** area would be the space where the data gathered to describe user environments would be used to build the user environments. The assembled user environments could reside with each individual, within a network, within a global

system, wherever a given configuration of hardware and software would use it most efficiently. In addition to the Information Use Environment, the new Budget Models, and the models for access and control mechanisms would reside in the Planning and Design area, giving whoever or whatever is responsible for this function the necessary perspective and information from which to construct design specifications.

Management and Administration: The Management and Administration area would implement the designs constructed by the Planning and Design function, using a projected value-added analysis of the design as a benchmark to monitor the performance of the design. Close monitoring, analyses of the design, and fine tuning of the implementation would occur in this area along with whatever standard system maintenance and upkeep would be required. The Management and Administration area would receive that data gathered from the Client Services area, analyze the data, and pass its interpretation of the data back to the Planning and Design area.

Client Services: The Client Services contains the resources of the virtual library environment most directly impacting and contacting the user. Client Services area would contain the mechanisms to access the information universe and create, manipulate, and maintain relevant subsets. Client Services would also contain the tools that enable users to utilize the available sets of client services. The mechanisms to evaluate the system design based on what value they add to the user, how the design fits the users information environment, etc. would reside here, providing the data gathering element in the system.

The data gathered in the Client Services area would be routed to the Management and Administrative area where it would be analyzed and passed on to the Planning and Design area. The Planing and Design area would incorporate the results into the design, resulting in a cycle of redesign, reevaluation, and implementation of new designs.

What Is Likely to Change:

In the next section, we will discuss some issues integral to classification research, issues whose theoretical underpinnings will doubtless contribute to designing the virtual library environment. No one can say with great accuracy what will be most transformed in this new environment, what will have the most impact, or what new needs will arise from new possibilities, but based on the way current informational technologies can alter the way we work, we could make some observations about what is likely to change:

Future information entities will not be linear, not in the way document-oriented data sets are now. Future materials may be a set of components, grouped and ungrouped as required, without a permanence of order, existing in a plethora of iterations, each one valued by someone and consequently stored and available to them on demand. Perhaps, in its original format, newly

created entities will have something analogous to a permanent die or watermark, that will travel with it forever, like a pedigree. These marks, whatever form they take, could document the path an entity has traveled as it evolves over time in various versions of its original self. Such a scheme would acknowledge the lineage without stunting new growth.

Future information entities may not be created once and then, theoretically, protected by law in an attempt to keep the entity static, frozen always in its original form. Entities may be assembled and disassembled, with the accompanying data taking whatever form is most clarifying to the user.

Entities will not exist as isolates, but may be linked together and stored from several perspectives, including perspectives relevant only to the user. Similarly, an entity might belong to a systematic classification scheme that places it within a widely accepted system, but also carrying many designations that carried only individual relevance.

Entities will be represented more by media than by text. Modeling or visualization techniques may replace lengthy lines of textual description. We have strong bodies of research addressing issues of text-based models. How will we address media-based models that represent concepts in three dimensions?

What Is Likely to Not Change:

While change is certainly the order of the day, based on our current work with constructing information use environment, we could make some projections about what is likely, at least in principle to persist:

The presence of masses of data, even in newly engaging media formats, does not make the data accurate or relevant or clear. Users will continue to value guides that provide filtering mechanisms for masses of data. Users may place progressively higher values on mechanisms that accurately help them locate highly relevant material quickly.

Browsing or the serendipitous discovery of information may continue to be one of the activities that delight users and mechanisms to facilitate browsing across a much broader variety of materials will be welcomed.

Just because information use environments are of primary importance to the design of virtual library environment does not mean that users will be any more aware than they ever have been regarding their information gathering and use behaviors. There is no reason to believe they will alter their desire or ability to analyze their options, describe and define the traits of their information use environments, consider the traits of the information they manage, or actively participate in the design of the information systems. Such things will remain the concern of designers who must find the mechanisms to answer needs that may not be clearly articulated, or do not yet exist.

Users will continue to need clues about where they are in the larger scheme of things. Classification schemes and control mechanisms that help place a concept in relation to something already known

will be welcome orientation devices. Similarly, navigation devices will continue to be vital to helping users pick their way through the information universe. Anything that enables conceptual mapping will continue to be of high value.

III. CRITICAL ISSUES FOR ACCESS

It becomes very clear that there is a real need to organize information in all formats and resources. However, while connecting with existing remote information resources, one ongoing problem of access is how to work with multi-databases which use various controlled vocabularies and apply different indexing and searching policies.

It is also obvious that there is a requirement for a transparent interface, i.e., one and only one indexing and retrieval language must be learned in order for users to search information and knowledge in a variety of resources and formats. This language will function as: (1) a command language which administrates the process of connecting, searching, transforming, etc.; (2) an access language which controls both bibliographic descriptive access and subject access; and (3) users' language which an user/author will apply in searching and organizing their work for synthesized retrieval results or producing/distributing creative works. Here the third function, the users' language requirements, presents the most unique demands and raises questions that have not yet been answered anywhere in today's online environment. What path will development of this third requirement follow? What methodologies can be used to develop the intellectual access and control mechanisms for virtual libraries?

In the process of developing intellectual access and control mechanisms for the virtual libraries, we have examined the established tools for knowledge organization and discussed several key decision-making issues. We would like to post our topics here in order to seek help from the workshop participants:

1. **Controlled language vs. natural language.** This is an issue being discussed by many published research papers. However, the previous discussions have not addressed the questions posed by the virtual library environment. Previous discussions also resulted in contradictory findings because: 1) they usually focused on the comparison of advantages and disadvantages of controlled vocabulary and natural language in terms of system development, search results, and the user-friendliness; 2) the discussions about the complementarity of natural and indexing languages are more or less limited to the use of non-controlled vocabularies (more precisely, on the use of any form which is not a thesaurus, a subject heading list, or a classification scheme) such as code systems, keyword indexes, citation indexes, etc.

   There is no evidence which provides clear guidance for either approach. We could use controlled vocabulary inside the system, which could be a virtual or invisible system to the user unlike traditional approaches. Natural language should be used in search while the responding results should came out with the support of highly-

controlled, fully semantic-networked indexing and retrieval language. A key research question is how to link controlled and natural languages or to merge them in one integrated language.

2. **Pre-controlled vs. post-controlled languages**. Most existing controlled vocabularies are pre-controlled, i.e., created prior to, or at the same time of, the establishment of a database. The frequency of updating a vocabulary by a good database producer is usually at least three months, while in more cases can be a year or more. At this point, pre-controlled vocabularies always have the problem of being out-of-date as soon as they are published. On the other hand, the general search tends to be easy in the pre-controlled vocabulary situation, though certain highly specific searches may be virtually impossible.

We are seeking an approach to include relatively more dynamic and up-to-date vocabulary. A post-controlled vocabulary, with the control applied at output rather than input, and increased and maintained with the use of the database by users, seems to be promising.

Theoretically speaking, this kind of vocabulary system can offer all the advantages of natural language with many of the attributes of the pre-controlled vocabulary.[1] The problems with this approach in the virtual library environment is that the post-controlled vocabulary could be an extremely useful searching-aid, but it may not be a good indexing language (which usually requires a highly-structured form). In addition, research on the theory and practice of the post-controlled vocabulary are less reported, though recently some approaches in achieving compatibility of indexing languages demonstrated similar functions of a post-controlled vocabulary. In practice, a number of significant gaps still remain to be explored in establishing post-controlled vocabulary.

3. **Systematic access vs. alphabetical access**. While examining various enumerative, analytic-synthetic, and hybrid vocabularies, one thing has been very clear: a good tool for knowledge organization should provide both systematic and alphabetical approaches to any sources. We note that a in the United States, most online bibliographic databases and online public access catalogs provide alphabetical access as the primary access, with or without systematic access (classification or categorization) as auxiliary access; b) to date, information retrieval systems including OPACs have primarily concentrated on querying based on the Boolean model for retrieval and provided little support for any browsing requirements; c) there is no sufficient evidence to show the power of classification in an online environment.

From the point of view of disciplined-based virtual library users, it is important to access and organize knowledge systematically. The basic question here is how we

---

1. Lancaster, F. W. *Vocabulary Control for Information Retrieval*. 2nd ed., Champaign: Information Resources Press, 1986, p. 168-173.

can provide such a tool for knowledge organization, a tool which combines both systematic and alphabetical approaches, a tool which facilitates both general and specific searches, and a tool which supports both querying and browsing. Here, querying assumes an element of pre-knowledge, purpose, precision and certainty on the part of the searcher, and browsing implies that the searcher is exploring, seeking to define or clarifying an imprecise information need. At a more sophisticated level, the question is how we can provide such a tool which should assist the user (also the author) in the whole process of search, retrieval, storage, and information synthesis gathered in ways that are individually meaningful and relevant.

4. On knowledge-based systems approach. ConQuest's[1] approach is one of the close approaches we are considering. By using published electronic dictionaries combined with sophisticated linguistic and statistical processing, ConQuest meets two expectations: first, it allows a user to use plain English in the search queries; and, second, it provides a capability for managing the entire research process including index, search, retrieve, browse, organize and route information. Other attractive features include a) delivering a semantic network of words, word meanings, and word relationships in searching, and b) displaying search results in ranked order.

The major difference between the virtual library environment and ConQuests capabilities is that for each disciplined-based virtual library, a core component of the knowledge-base would be a controlled (mostly post-controlled) indexing and retrieval vocabulary of particular subjects, not just published e-dictionaries, so that better user warrant and document warrant can be ensured. This highly discipline-oriented macrovocabulary is supported by the analysis and processing of existing controlled vocabularies of related subjects, user queries, full-text of documents, expert experience, special subject classification systems, and other reference tools such as e-dictionaries. Further, the virtual library vocabularies will not be merely alphabetical lists of terms, since discipline-based classification systems would be also important. One advantage of using knowledge-based system in virtual libraries is that our libraries and information sources will be highly disciplined-based, thus the domain is very clear and the terminology used in related sources would be less-ambiguous. The question we are facing is how will the knowledge-bases be established. Experience with ConQuest provides very useful information.

5. Other highly relevant issues include vocabulary visualization, online display, automatic thesaurus construction and indexing issues related to search engines. Staff training is also a concern related to producing library staff capable of participating in the design of knowledge systems and able to act, in addition, as educators, researchers, navigators, and problem-solvers and synthesizers in the

---

1. ConQuest Software offers ConQuest, a text search and retrieval system that employs published dictionaries to index data automatically.

library of the future. User training issues are also relevant since users will need to be taught how to access, use, evaluate, communicate, transfer, and disseminate media-integrated materials.

6. Finally, this section on critical issues for access conveys, in part, the long standing concern focused on the subject content of information entities the aboutness of the thing in question. In traditional academic environments, querying mechanisms and authority structures share a set of commonly recognizable attributes that have evolved along highly structured lines with limited options. The shared goal of most intellectual products produced in the academic environment is that of researching and writing for scholarly publication in a print environment. But what of entities that will require entirely different or markedly different attributes? How will the relevant attributes of entities be devised, represented, and made available in a way that is significant to users from a variety of disciplines and information use environments? When the information universe becomes extremely varied, will the concept of subject be expansive enough to contain all the significant attributes required to describe and define such a world? In terms of research, what new attributes will join subject as the focus of attention and resources?

IV. SUMMARY OF THE PROBLEM AND FUTURE RESEARCH DIRECTIONS

It will be very important to address the need for developing intellectual access and control mechanisms for future libraries in this classification research workshop. The approaches used in the project and the questions generated as the project goes on, will provide the audience with a preview of the challenges that classification researchers are or will be facing. We would like to get feedback from the researchers as well as their reactions and suggestions to our approaches.

Why is it so important to try to form these questions about an environment that does not yet exist? One reason is because it will obviously arrive in some form, with or without our input, and whether or not we have given the matter sufficient thought. Another reason is that it is important to focus some resources, research, and thought on how to create environments in which we will think and learn, problem-solve and dream differently than we do now or ever have before. We need to theorize, examine, test, and eventually design for an environment that truly facilitates what happens when any human being reaches out from within themselves, looks up from what already exists within their private worlds of knowledge and reaches out, in whatever way is possible, to the larger universe, seeking something outside of themselves. We have designs and mechanisms right now that provide access and control, but they don't begin to go nearly far enough.

The promise of greatest consequence regarding informational technologies is not their capabilities to replace books or libraries or librarians or scholars or researchers or universities. Their power is simply that they can create an environment that would never exist without them. They will redefine the parameters of what can be done in terms of bringing people into meaningful contact with their world: the world that lives within them, and the world that surrounds them and remains to be discovered. Discovered, in our context, means accessed in a meaningful way, at the moment when such access is most appropriate for the person seeking a connection.

As with any design process, classification researchers proceed from a knowledge base formed of their predecessors work, past problems, projects, and investigations. It is foolhardy to even suggest ignoring the wisdom gained from previous work. One might contend however, that research centered on print technologies and the entire environment driven and shaped by print technologies will not ultimately result in access and control mechanisms for the environment that will permit the design of true virtual libraries. Information technologies, media technologies, all the intelligent technologies that place power in the hands of the user, have had enormous impact on the print environment, but that will not create an environment designed using the fresh parameters drawn by transforming information technologies. Efforts to improve what we do now will not provide answers to future designs. What such efforts will create are more effective, electronic versions of the print parameters. The mechanisms that made the print environment so powerful will not produce a powerful virtual environment; the possibilities will change too dramatically.

We must discover the touchstone concepts and principles, the ones whose vitality endures and will continue to endure. We must then explore how to realize them in new environments. It is a large challenge, but only another one in a long series of such tasks.

When transforming technologies focused on transportation, for example, the resulting chaos seemed to nearly overwhelm society, that initially had no controls to master the sweeping changes that came one after another, among them railroads, paddle boats, and the mass-produced automobile. Even a superficial look at some of the activity that went on around the emergence of the auto shows that some intellectual activities proved more fruitful than others. Market studies to see who would buy a horseless carriage would have predicted utter failure. Asking if horseless carriages would replace the horse, the blacksmith, and the buggy whip maker wouldn't have proven very helpful either. But acknowledging the essential need for people to get from one place to another and asking, What could the nature of transportation be if the automobile becomes fully realized? would keep the questioner on a reasonable track toward planning for the future. Of course, it helps to have a thorough knowledge of the environs about to be transformed.

Classification researchers have something special to offer to this search for the right questions for developing access and control mechanisms for the virtual library environment. Classification researchers understand an essential concept: When you create an entity and store it away, how you created it, how it is structured, what the entity understands about itself and its place in the world (its intelligence or lack thereof), how it is described and defined to others — all have an enormous amount to do with ever being able to revisit that entity should the moment arise when such a rediscovery would be appreciated. This essential understanding is not common. It might even be said to be dangerously scarce. In the rush to make new information sources available, such vital concerns often get lost, resulting in serious consequences for future access.

We must not ask how we can provide better access and control. We must ask how we can provide access and control in a way never before possible, with our goals and perspectives clearly in mind. We cannot expect users to articulate their need for this transformed environment. Users will not express a need for something that cannot be done easily or conveniently now, something that is not even within their realm of experience. We can examine critical theory for new ways to fulfill its potential. We can examine our users information environments for ongoing needs that can be met

by completely different approaches. We can attempt to set aside responding to current perceptions of what is possible and concentrate instead on transforming our notion of what will be possible. The reward of so difficult and demanding a task may be a small glimpse into the future.

The authors have told you that they seek the advice and insight of classification researchers, but we have not provided classification researchers or anyone else with any specifications to fulfill, no conceptual Requests For Quote to respond to, except in the broadest and most abstract sense. We have not and we cannot say, Here is the problem you must help us find a solution to. We are, in essence, alerting you that we would like to transform information management, every aspect of it. Were asking this of the people who know a great deal about the mechanisms that do the job now. We are alerting those people so that when a transforming technology occurs, we can all work together to figure out how we can apply it to our own transforming project.

BIBLIOGRAPHY

Addison, E.R., Wilson, H.D., & Feder, J. (1993) The impact of plain English searching on end users. *Proceedings of the Fourteenth National Online Meeting*. Medford, NJ.: Learned Information, Inc. 1993: 5-10.

Allen, Bryce. Technological Capabilities and Human Constraints: Automation and Organizational Change in Academic Libraries. In *Illinois Library Association/College and Research Libraries Forum Fall Conference in Urbana, IL.*, College and Research Libraries, 1990.

ASLIB Proceedings. In *ASLIB Proceedings* edited by Susan Grant, ASLIB, 165-181, 1988.

Associates, Robert Ubell. *Draft Preliminary Findings of the Rights for Electronic Access to and Delivery of Information (READI) Project*. Robert Ubell Associates for the Coalition for Networked Information, 1992. draft report

Beniger, James R. *The Control Revolution: Technological and Economic Origins of the Information Society*. Cambridge, MA: The Harvard University Press, 1986.

Besser, H. Visual access to visual images: the UC Berkeley Image Database Project. *Library Trends* 38 (4 1990): 798-798.

Bower, J. M. Vocabulary control and the virtual database. *Knowledge Organization* 20 (1 1993): 4-8.

Campus strategies for libraries and electronic information. EDUCOM strategies series on information technology, ed. Caroline R. Arms. Bedford: Digital Press, 1990.

ConQuest software handouts distributed at the 1993 National Online Meeting, New York, NY USA May 4-6, 1993.

Craven, T. Concept relation structures and their graphic display. Paper presented at the *5th International Study Conference on Classification Research (ISCCR '91)* June 24-28, 1991, Toronto. 1991.

Gapen, D. Kaye. The Virtual Library: Knowledge, Society, and the Librarian. In *The Virtual Library: visions and realities*, ed. Laverna M. Saunders. 1-14. Westport: Meckler, 1993.

Harter, Stephen P. *Online Information Retrieval: Concepts, Principles, and Techniques*. Orlando: Academic Press, Inc., 1986.

Hill, Linda L. and Edie M. Rasmussen. Studies in Multimedia: State-of-the-Art Solutions in Multimedia and Hypertext. In *Mid-Year Meeting of the American Society for Information Science in San Jose, CA,* edited by Susan Stone and Michael Buckland, Learned Information, Inc., 1991.

Kahin, Brian. Scholarly communication in the network environment: issues of principle, policy and practice. *The Electronic Library* 10 (October 1992): 275-286.

Lancaster, F. W. *Vocabulary Control for Information Retrieval.* 2nd ed., Champaign: Information Resources Press, 1986.

Lancaster, F. W. *Indexing and Abstracting in Theory and Practice.* Champaign: Information Resources Press, 1991.

Lucier, Richard E. Knowledge Management: Refining Roles in Scientific Communication. In *New Technologies and New Directions,* ed. G. R. Boynton and Sheila D. Creth. 118. Westport: Meckler Publishing, 1993.

Meincke, P. P. M and P. Atherton. Knowledge space: A conceptual basis for the organization of knowledge. *JASIS* 27 (1 1976): 18-24.

Molholt, P. The role of the 'Art and Architecture Thesaurus' in communicating about visual art. *Knowledge Organization* 20 (1 1993): 30-34.

Nelson, Theodor Holm. *Literary Machines.* 87.1 ed., Bellevue, WA: Owl International, Inc., 1987.

Organizing and managing information resources on campus. ed. Brian L. Hawkins. McKinney: Academic Computing Publications, 1989.

Pao, Miranda Lee. *Concepts of Information Retrieval.* Englewood: Libraries Unlimited, 1989.

Rheingold, Howard. *Virtual Reality.* New York: Simon & Schuster, 1991.

Shera, Jesse H. *Libraries and the Organization of Knowledge.* Hamden, CT: Archon Books, 1965.

Shera, Jesse H. *Sociological Foundations of Librarianship.* 1st ed., New York: Asia Publishing House, 1970.

Svenonius, E. (1986) "Unanswered questions in the design of controlled vocabularies." *JASIS* 37(5): 331-340.

Symposium on Scholarly Communication (1991: University of Iowa). *New technologies and new directions: proceedings from the Symposium on Scholarly Communication.* ed. G. R. Boynton and Sheila D. Creth. Westport: Meckler Publishing, 1991.

Taylor, Robert S. *Value-added processes in information systems.* Communication and information science, Norwood: Ablex Publishing Corporation, 1986.

Tufte, Edward. *Envisioning information.* Cheshire: Graphics Press, 1990.

User Interfaces for Online Public Access Catalogs: A Research Workshop. (Workshop notes distributed at the SIGIR '93, the Sixteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Pittsburgh, PA USA June 27-July 1, 1993).

Veltman, K. Electronic media and visual knowledge. *Knowledge Organization* 20 (1 1993): 47-53.