

Youngjoo Moon. A study on the hypercatalog applying bibliographic relationships: a gateway to information in a networked environment. A Master's paper for the M.S. in L.S. degree. August, 2001. 73 pages. Advisor: Jerry D. Saye

Current library catalogs do not fulfill users' needs for navigating freely from one record to related records. They also fail to meet the second objective of the library catalog either because of the generic defect of current linking devices and omission of virtual copies. The application of hypertext system to the traditional library catalog is suggested for a solution, because this type of OPAC system is able to present links in a more user-friendly way. The new system is named hypercatalog. An expanded USMARC format is suggested for data structure of the hypercatalog.

The hypercatalog is expected to be able to offer an appropriate interface to present links for related bibliographic records for both of Web resources and traditional library resources. The end users activate these links after a query using the hit as the starting point merely clicking one of links on the full bibliographic record screen.

Headings:

Library catalogs

Hypercatalog

Bibliographic relationships

A STUDY ON THE HYPERCATALOG APPLYING BIBLIOGRAPHIC
RELATIONSHIPS : A GATEWAY TO INFORMATION IN A
NETWORKED ENVIRONMENT

by
Youngjoo Moon

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Library Science.

Chapel Hill, North Carolina

August, 2001

Approved by:

Advisor

Table of Contents

Introduction.....	1
Literature review.....	3
Research methodology.....	15
Identified problems of the traditional catalog.....	16
Proposed solutions.....	47
Design of hypercatalog applying bibliographic relationships.....	52
Conclusions.....	68
Bibliography.....	69

1. Introduction

Since Charles Ammi Cutter published 'Rules for a dictionary catalog' in 1876, the objectives of the library catalog have been summarized to finding function and collocating function. Through development from early book-form catalogs to card catalogs and online catalogs, the library catalog has been successful in meeting the needs for finding function. However, it has been said that the collocating function, even in the most recent online catalogs and web catalogs, has not been successfully implemented to meet user's needs. With the development of information technology, various information media have emerged including CD-ROMs and the Internet. Patrick Wilson (1989) thought that the changing text of an electronic documents and the concept of a work, together with the capabilities of computer networks led to redesigning the classical library catalog. He pleaded for Cutter's second objective that is also called as collocating function and wanted "the work" as the unit to be described in the catalog instead of the physical unit that contains it. To improve the weak collocating function of current online catalogs, this paper explores applicability of bibliographic relationships in constructing hypercatalogs that use hyperlinks to relate bibliographic records that share common bibliographic relationships.

Humans think associatively. This means they navigate from an idea to related ideas freely and directly. However, current paper-based media have a linear structure unlike the nature of human thinking. Whenever we want to read related parts or references during reading a book, we have to search again the related resources through the table of content, the index, or the bibliographical references of the book. To overcome

these limitations, hypertext technology has been developed. Hypertext technology enables us to find related parts directly through hyper-links that link related nodes through predetermined paths.

As the current library catalog has the same limitation of printed media, it can be hoped that the limitation can be overcome with adoption of hypertext technology in the library catalog. In this paper it is attempted that using hypertext technology to link related catalog records through bibliographic relationships. Tillett summarized bibliographic relationships in total seven, which are used in this paper.

This paper also raises issues about the collocation function and the linear structure of current library catalogs. As an alternative, a catalog with hypertext functionality by applying bibliographic relationships is suggested, and the record structure and the interface design of the catalog are explained.

2. Literature review

2.1 Bibliographic relationships

2.1.1 Tillett's bibliographic relationships

Probably the most comprehensive and significant work on bibliographical relationships to date is Tillett's doctoral dissertation, "Bibliographic relationships: toward a conceptual structure of bibliographic information in cataloging" (1987). She investigated the general nature of bibliographic relationships and developed a taxonomy of relationships. She defined seven relationships in total (Tillett, 1987, p. 24-25):

- a. *Equivalent relationship : a relationship between exact copies of the same manifestation of a work, or between an original item and its reproductions*
- b. *Derivative relationship : a relationship between a work and a modification based on that same work. "These include: (a) variations or versions of another work, such as editions, revisions, translations, summaries, abstracts, digests; (b) adaptations or modifications that become new works but are based on earlier works; (c) changes of genre, as with dramatizations and novelizations; and (d) new works based on the style or thematic*

content of another work, as with free translations paraphrases imitations, and parodies

- c. *Descriptive relationship : a relationship between a bibliographic work or item and a criticism, evaluation, or review of that work*
- d. *Whole-part relationship : a relationship between a component part of a bibliographic item or work and its whole, as with an individual selection from [the work] and the whole anthology, collection or series*
- e. *Accompanying relationship : a relationship between a bibliographic item and the bibliographic item it accompanies*
- f. *Sequential relationship : a relationship between bibliographic items that continue or precede one another, as between successive titles of a serial*
- g. *Shared characteristics relationship : a relationship between a bibliographic item and other bibliographic items that are not otherwise related but coincidentally have a common author, title, subject, or other characteristic used as an access point in a catalog, such as a shared language, date of publication, or country of publication*

Tillett conducted both analytical and empirical studies on bibliographical relationships. In the analytical studies (Tillett, 1987, 1991, 1991b), she investigated the history of cataloging codes and compared individual rules to observe similarities and difference in the handling of relationships. She selected

24 rules for review that were identified major cataloging rules and sets of rules used in the United States. An exception was the Panizzi's rules that have been acknowledged as the basis for cataloging codes used in the United States.

Through the analysis she derived seven bibliographic relationships as mentioned above and identified various links and linking devices adopted in major 24 cataloging rules.

In her empirical study Tillett (1987, 1992) measured the incidence of her categories of bibliographic relationships in Library of Congress machine-readable bibliographic records from 1968 to 1986. She analyzed the frequency of occurrence by counting the incidences of codes associated with each relationship type within the machine-readable records. She also examined the characteristics of bibliographic items exhibiting particular relationships by language, place of publication, publication date, subject, and bibliographic format. She confirmed that bibliographic relationships are prevalent and an important feature in a research library catalog.

Although her study on bibliographic relationships was thorough and comprehensive, it had some limitations. She only examined 4 factors as compared to each MARC format for each relationship type identified due to the limitations of available LC programs. She also only dealt with explicit relationships expressed by catalog records in her empirical study. According to Smiraglia's (1992) observation 63.2% of all derivative bibliographic relationships were not expressed by catalog records.

2.1.2 UNIMARC

UNIMARC (1980) is the universal MARC format for the communication of bibliographic information. It was created in the 1970's by the International Federation of Library Association (IFLA), and was based on the MARC format created at the Library of Congress in the late 1960's. UNIMARC defines three types of bibliographical relationships (UNIMARC, 1980, p. 22-23):

- a. *Vertical – the hierarchical relationship of the whole to its part, and the parts to whole, e.g., downward link: a serial to its sub-series or to individual volumes of the series: upward link: the individual volume to its sub-series and/or series*
- b. *Horizontal – the relationship between versions of an item in different languages, formats, media, etc.*
- c. *“Chronological – the relationship in time between issues of an item, e.g., the relation of a serial to its predecessors and successors...*

Overall, UNIMARC's three bibliographical relationships are not mutually exclusive and comprehensive. According to Tillett (1987), the distinction between horizontal and chronological relationships was vague, because in the case of a handbook issued annually in updated versions, with slight title changes the relationship between ensuing editions was both horizontal and chronological. Compared to Tillett's bibliographical relationships, UNIMARC's three relationships omit some relationships such as relationships between a casebook

and the work it analyzes, between review and the book under review (descriptive relationship), and between a supplement and the work it accompanies (accompanying relationship).

2.1.3 CCF (Common Communication Format)

The Common Communication Format (1988) was created by the Special Committee for CCF sponsored by UNESCO to ensure data exchange between bibliographic control related organizations (e.g., libraries and index and abstract services). It defines two kinds of bibliographic relationships: vertical and horizontal. Vertical relationship means the relationship between a book and its chapters of the book or an issue of a journal and the articles of the issue. A horizontal relationship means the relationship between editions of a book or different translations of a book. The bibliographic relationships of CCF also lack comprehensiveness and exclusiveness in comparing to Tillett's relationships.

2.1.4 Gossens and Mazur-Rzesos' hierarchical relationships

Gossen and Mazur-Rzesos (1981) introduced a schematic representation for hierarchical relationships to express simple and superimposed, complex tree structures. They provided the theoretical basis for manual and computerized solution to expressing such relationships. Although their definition for bibliographic relationships was not original, and according to Tillett (1987, p. 10)

it must have been derived from the UNIMARC definitions, they successively demonstrated the application hierarchical tree structures to certain types of bibliographic materials.

2.1.5 Smiraglia's derivative bibliographic relationships

Smiraglia (1992, 1999) further developed Tillett's definition of the derivative relationship. Tillett stated that "derivative bibliographic relationships exist between any new conception of a work and its original source, or its successor, or both" (Tillett, 1987, p. 26). Smiraglia developed a taxonomy of the derivative relationship, expanding its definition to include eight different categories of derivation (1999, p. 495).

- a. *Simultaneous derivations – works that are published in two editions simultaneously, or nearly simultaneously...*
- b. *Successive derivations – works that are derived one or more times, and issued with statements such as "second ... edition," [and] works that are issued successively with new authors...*
- c. *Translations*
- d. *Amplifications, including illustrated texts, musical settings, and criticisms, concordances and commentaries...*
- e. *Extractions, including abridgements, condensations, and excerpts.*
- f. *Adaptations, including simplifications, screenplays, librettos, arrangements of musical works, and other modifications.*

g. Performances, including sound or visual ... recordings.

Through two empirical studies, Smiraglia (1992, 1994) measured the amount and extent of derivation in the Georgetown University library catalog (GEORGE), the OCLC Union Online Catalog (OLUC, now known as WorldCat), and the Research Libraries Information Network (RLIN) catalog of the Research Libraries Group. His results indicate that 49.9% of all works in a local research library catalog exhibited derivative bibliographic relationships. He confirmed Tillett's finding about the control of bibliographic relationships in general and reassure that the amount of bibliographic relationship in local and union catalog is large enough be concerned about. He also found that less than half of all derivative relationships in the bibliographic records were explicitly stated in the catalog.

2.1.6 Leazer's model for linkages in catalog using bibliographic relationships

Leazer (1993) described a conceptual model for the control of all types of relationships. This model contained a structure similar to the structure proposed by Gorman (1982) for modeling bibliographic relationships. He moved beyond Gorman by taking advantage of the more recent work done by Tillett and Smiraglia. Leazer's model specified the types of relationships in a catalog that contain a linking structure. Leazer also emphasized that the linkages among descriptive records should be used to control entire bibliographic families. He suggested that the development of direct and explicit control of bibliographic families would greatly enhance the user's ability to navigate the bibliographic universe, and one enhancement would be the development

of a dictionary of works that could stand alongside traditional bibliographic control retrieval system.

2.1.7 Functional Requirements for Bibliographic Records (FRBR)

Functional Requirements for Bibliographic Records was created by the IFLA Study Group on the Functional Requirements for Bibliographic Records in 1997 as a conceptual framework to guarantee the quality and relevance of bibliographic records. The study group (1998, p. 2) stated that “the aim of the functional requirements for bibliographic records is to produce a framework that would provide a clear, precisely stated, and commonly shared understanding of what it is that the bibliographic record aims to provide information about, and what it is that we expect the record to achieve in terms of answering user needs.” It differentiates *work*, *expression*, *manifestation*, and *item*. According to the study group (1998), *work* is a distinct intellectual or artistic creation, *expression* is the intellectual or artistic realization of a *work* reflects intellectual or artistic content, *manifestation* is the physical embodiment of an *expression* of a *work*, and *item* is a single exemplar of a manifestation reflects physical form.

Bibliographic relationships are also defined in the functional requirements for bibliographic records. It defines relationships between *work*, *expression*, *manifestation*, and *item*, relationships to persons and corporate Bodies, subject relationships, and other Relationships. Other Relationships includes *work-to-work* relationships, *expression-to-expression* relationships, *expression-to-work*

relationships, *manifestation-to-manifestation* relationships, *manifestation-to-item* relationships, and *item-to-item* relationships.

Functional Requirements for Bibliographic Records clearly defines the entities dealt in bibliographic records and bibliographic relationships between them. The definitions and relationships made in the Functional Requirements for Bibliographic Records will be helpful for future discussion on how current cataloging codes and MARC have to be changed

2.2 Hypercatalog

2.2.1 Hypercatalog project at Linköping University

The term “hypercatalog” was first introduced in a project started in Sweden in the 1980s at Linköping University. The conventional catalog was extended to a system with “...links and relations between fields, records and files...”. Such a new catalog also held records not available in the local collection (Hjerppe 1986, 1989). According to Hjerppe (1989), the hypercatalog was designed to be an extension and enhancement of the traditional catalog. Thus, its design incorporated the following features:

- a. *It supports browsing and navigation as the primary modes of using the catalogs but also provides traditional searching, in which the user has to specify what they want*
- b. *It is much more structured internally than traditional catalogues*

- c. It has alternative means for presenting and illustrating structures, information and relations*
- d. It has tools for establishing relations and following 'trails'*
- e. It is dynamic in its nature, not only by increasing the number of records*
- f. It provides each user with the ability to specify, and save, their views of the library*
- g. It has models of the users available*
- h. It contains more and different information than present catalogs, especially in terms of links and relations between fields, records and files, and facilities for utilizing these*
- i. In addition to providing information on individual items, it also offers information on collections.*

However, this ambitious project was never finished.

2.2.2 The Hypercatalog Graz - Budapest (HyperKGB)

The hypercatalog was a prototype hypertext-based library catalog developed jointly by Austria and Hungary in 1996. In this hypercatalog, a user had three general possibilities for searching:

- a. matching
- b. browsing
- c. user-defined collections.

According to Baptist (1996), the catalog had following search feature:

While matching, the user can search by any combination of fields in the database using Boolean operators 'and', 'or' as well as 'andnot'. After a search was executed, the retrieved books are presented graphically on the screen. On the spine of each book, its author and title can be seen. If a user wants more information about a specific book, they have to double-click it. The corresponding book is 'opened' and all bibliographic data will be displayed. In the opened book it is not only possible to get information about it, but also to make personal annotations. These remarks can be read by all other users. After a book was selected (single mouse click) or opened, it is possible to browse through the catalog.

Baptist (1996) also mentioned that the prototype proved to be easy to learn and easy to use. Nearly all students preferred the catalog to the existing OPACs at the University of Graz though there was criticism on some parts. Most criticism focused on the detailed version of querying which was found to be very complicated.

2.2.3 Bertha's concept for a hypercatalog

In Bertha's paper, "Inter- and intra-bibliographical relationships: a concept for a hypercatalog," she (1993) thought that concepts and experiences of handling the structure of a thesaurus could be applied in setting up an extended catalog with linked

bibliographic units. She titled such an expanded catalog, a hypercatalog, because the relationships between descriptors were analogous to the relationships between bibliographic units. She suggested to use international bibliographic numbers such as ISBN, or ISSN as the unique identification of bibliographic units, and to generate all links by the system itself by defining the relationships between bibliographic units. She devised these links: additional copies, variant editions, translations, predecessor or successor publications, parent or part publications, and secondary literature (Bertha, 1993). Finally, she suggested using a hypertext system to offer an appropriate interface between links.

Bertha's article was the very first in suggesting a concept for a hypercatalog. She suggested using bibliographic relationships to connect bibliographic units. However, in comparison to Tillett's relationships, Bertha's were not comprehensive. She lacked accompanying relationship and shared characteristic relationship. She didn't present specific methodology in complementing the hypercatalog. She vaguely mentioned that the structure of thesaurus can be applied to construct the hypercatalog.

3. Methodology

This paper adopts the system evaluation methodology. A typical system evaluation study consists of five stages:

- 1) Definition of the problem
- 2) Description and analysis of the existing system
- 3) Design of the modified system
- 4) Implementation of modification, and
- 5) Testing and feedback.

A system evaluation study begins with the identification of problems or shortcomings in current system. The third stage of a system evaluation study is the description of a modified system created specifically to address the shortcomings of the current system. This research will be limited to the problem definition, description of the existing system, and design phases of a system evaluation study. Neither implementation or testing nor testing and feedback are included in the paper.

4. Issues of the traditional library catalog

4.1 The objectives of the library catalog

Cutter's 'Rules for a dictionary catalog' (1904) governed the creation and maintenance of catalogs, and became a major influence on twentieth-century Anglo-American codes. Cutter's objectives are credited as the first widely accepted statement of catalog functions (Taylor, 2000, p. 6). Cutter states that the catalog functions in order (Cutter, 1904, p. 12):

- a. *To enable a person to find a book about which any of the following is known*
 - the author*
 - the title*
 - the subject*
- b. *To show what the library has*
 - by a given author*
 - on a given subject or*
 - in a given kind of literature*
- c. *To assist in the choice of a book*
 - as to its edition (bibliographically)*
 - as to its character (literally or topically)*

Kumar (1981, p. 5) stated that the first objective emphasizes that the library catalog should serve as a finding list for specific documents. This would require provision for individual entries for each book, providing approach through author, title, and subject. The second objective emphasizes that the library catalog should serve as a finding list for groups of documents. This would require provision for a uniform entry for each group. The third objective deals with document description in the catalog. According to it, the description should enable a user to distinguish between different editions of a given documents. Additionally, it should assist in the choice of a document by providing enough information about its distinctive features.

According to Ranganathan (1964, p. 77), a catalog should be designed so as to:

- a. *Disclose to every reader his or her document*
- b. *Secure for every document its reader*
- c. *Save the time of the reader; and for this purpose*
- d. *Save the time of the staff*

The above objectives are simple and deeply enough expressed in the light of his five laws of library science. These, however, are general in nature. On the other hand, the objectives proposed by Cutter are more specific.

Lubetzky (1953, p. 36) gave a slightly different emphasis to the functions of the catalog.

The first objective is to enable the user of the catalog to determine whether or not the library has the book he wants... The second objective is to reveal to the user of the catalog, under one form of the author's name,

what works the library has by a given author and what editions or translations of a given work.

His first objective stipulates the control of individual books; the second objective specifies the control of bibliographic works. Lubetzky himself pointed out that “the objectives are inherently in conflict.” (Lubetzky, 1953, p. 37-38)

The International Conference on Cataloging Principles was held at Paris in 1961.

The Paris conference resolved that:

Functions of the catalog

The catalog should be an efficient instrument for ascertaining

1) *Whether the library contains a particular book specified by*

a. its author and title, or

b. if the author is not named in the book, its title alone, or

c. if author and title are inappropriate or insufficient for identification, a substitute for the title; and

2) *a. which works by a particular author and*

b. which editions of a particular work in the library

The above functions are essentially restatement of Cutter’s objectives.

Smiraglia identified three general functions for the library catalog: identification, collocation, and evaluation (1987, p. 17-18). Weintraub (1979, p. 392) and Taylor (2000, p. 5) have similar theoretical formulations of catalog functions. Furthermore, there is a close correspondence between the functions of the catalog as defined by Smiraglia and the objectives of Cutter, Lubetzky, Langanathan, and ICCP.

Finding function enables a user to find a particular bibliographic object. Finding is directly related to first objectives of Cutter, Lubetzky, and ICCP. Collocating function is premised upon the finding function, and is closely related to Cutter's, Lubetzky's, and ICCP's second objectives, and partially fulfills Cutter's third objectives. Taylor (2000, p. 5) defined;

Collocation is a means for bringing together in one place in a bibliographic tool all entries for like and closely related materials... In many cases a particular work is shown in its relationship to a larger group of works – e.g., the bibliographic record for a play based upon Huckleberry Finn should be found with records for [book] editions of Huckleberry Finn, which are in turn found with records for other works of Mark Twain.

Leazer (1993, p. 16-17) stated that;

The placement of bibliographic records in physical proximity to each other is accomplished by entering all records under one form of the author's name in conjunction with the item's title. The objective of authority control is to maintain consistency in the selection and form of an author's name and the item's title. The library catalog functions to control individual items and works. Works are controlled by collocating items into groups that reflect works. Collocation is made possible manually in card catalogs by placing bibliographic records front to back, or side by side in a book-form catalog. In the computer environment the actual physical

order of the descriptive records is not important, but the index entries for the records must be the same or in alphabetical sequence.

4.2 Bibliographic unit versus literary unit

The distinction between a work and a book that physically contain the work has long been an important concept to cataloging. Even after a long discussion between Lubezky and Verona during the 1960s, the problem generated from the distinction has not been solved. The problem was aggravated by many cataloging codes that use the terms work and item loosely. Lubezky (1969, p. 11) identified the problem;

The book, it should be noted, comes into being as a dichotomic product – as a material object or medium used to convey the intellectual work of an author. Because the material book embodies and represents the intellectual work, the two have come to be confused, and the terms are synonymously used not only by the layman but also by the cataloger himself.

Lubezky and Verona generally agreed on what the objectives of a cataloging should be and the advantages and disadvantages of using either the book or the work as the basis for the main entry. Verona stated that;

“A catalog designed for the first objective and neglecting the other two will have as its basic element the individual book...” (1959, p. 79) and

“Hence particular books will not be considered as single items but as representatives of a whole group of similar items, all belongings to the same [work]” (1959, p. 80)

According to O'Neill and Vizine-Goetz (1989, p. 168), Lubetzky and Verona implicitly agreed on followings in addition to the objectives of the library catalog:

- 1. The first two objectives are in many cases conflicting so that it is not possible for any catalog to weight them equally;*
- 2. The book should be the primary bibliographic entity for cataloging purposes;*
- 3. The main entry will be selected to represent the book if the emphasis is on the first objectives or the work if the emphasis is on the second.*

However, they differ on the function of the main entry. Verona's position is that the main entry should be used to identify bibliographic units, Lubetzky's that it should be used to assemble literary units. It is said that "a catalog based on the primary of the bibliographic units is seen as favoring the first objectives of the catalog, whereas one based on the literary unit is seen as favoring the second." (Carpenter and Svenonius, 1985, p. 152)

With the development of the MARC and computerized library catalogs the situation has been greatly changed, since Lubetzky-Verona's debates. Kilgour (1979, p. 34) recognized that;

The online computerized library catalog is a wholly new type of catalog having a drastically different design from the seventeenth century book-form catalog and the nineteenth century card catalog. The book-form catalog and the card catalog are linear arrangements of bibliographic entries, sometimes of enormous length. Online catalogs consist of a large series of miniature catalogs...

From this position, it could be said that (O'Neill and Vizine-Goetz, 1989, p. 169);

The creation and dissemination of machine-readable catalog records has resulted in new ways of storing and accessing bibliographic data.

Computer-based catalogs can support far more complex data structures than book or card catalogs. Unlike the card catalog, whose structure is evident, the physical organization of the online catalog is hidden from the user. Entries which would burden the card catalog pose no similar problems in computer-based catalogs.

Functional Requirements for Bibliographic Records divides bibliographic units and literary units into four: *work*, *expression*, *manifestation*, and *unit*. In terms of Lubezky and Verona's distinction, *work* and *expression* are literary units, and *manifestation* and *unit* are bibliographic units. The IFLA Study Group on the Functional Requirements for Bibliographic Records defined the relationships of them. (1998, p. 23)

A work may be realized through one or more than one expression (hence the double arrow on the line that links work to expression). An expression, on the other hand, is the realization of one and only one work (hence the single arrow on the reverse direction of that line linking expression to work). An expression may be embodied in one or more than one manifestation; likewise a manifestation may embody one or more than one expression. A manifestation, in turn, may be exemplified by one or more than one item; but an item may exemplify one and only one manifestation.

The functional requirements have been planned for the changing environment of bibliographic data with the emergence of electronic resources and the Internet. The question has been arisen: what kind of bibliographic records are required nationally and to an increasing degree also internationally to meet the different uses and user needs. Although the functional requirements attempted to answer to the question, more discussions are needed for them to be reflected in cataloging codes and MARC format.

4.3 Second objective of the library catalog in a networked environment

According to Bertha (1993, p. 212), “Most all current library catalogs enable users to locate a well-defined ‘bibliographic unit’ easily, but they do not allow looking for a ‘literary unit’ without any problem.” The library catalog has two kinds of objects to control: bibliographic unit (item) and literary unit (works). Lubetzky recognized the tension inherent in the library catalog and other bibliographic tools. He said (1969, p. 248)

The essence of the modern concept of cataloging ... has gradually emerged from a growing realization that the book (i.e., the material record) and the work (i.e., the intellectual product embodied in it) are not coterminous; ... that the book is actually only one representation of a certain work which may be found in a given library or system of libraries in different media.

As stated above, a function of collocation is to identify and control works. The description of works is accomplished implicitly by grouping related items and works under one individual entry. Works are controlled implicitly by grouping together descriptions of items. This is not really the description or the control of bibliographic works *per se*. Rather, it is similar to an enumerative definition. The description of a work is limited to listing of descriptions of the item that contain a manifestation of the work. The current library catalog that is item-based catalog requires the system user to recognize what the item-description have in common and what their relationship is to each other.

However, the “finding function” has been considered to be more important than the “collocation function”. Since current library catalogs only deal with *manifestation* and *item* among four entities (*work*, *expression*, *manifestation*, and *item*) defined in the Functional Requirements for Bibliographic Records, different manifestations of a work as variant editions, translations, or versions in other media or successor titles are not necessarily placed together in sequence in present systems. References to other manifestations are not recorded systematically, and they always refer in one direction such as from a later edition to an earlier one or from translation to the original publication. Furthermore, such references are not standardized, and, therefore, they cannot be used to generate links automatically.

According to Wilson (1989), the “collocation function” is not less important than the “finding function” in the current network environment. Wilson (1989, p. 9-10) said that;

The unit of record, the basic unit of cataloging, had been the book or more generally the publisher's unit, not the work itself, which was often regarded as major concerns of library users. What we describe in a basic unit record is a particular publisher's item; we may describe it as an appearance of a particular book, but still the basic fact is that we are describing a particular book or serial or other publication unit ... The practical dominance of the first objective has, along with the limiting technology of the card catalog, prevented us from even considering the form of record appropriate to a bibliographical system in which the second objective had first priority. But, now we can think of new and more appropriate ways of showing the works we have in a collection.

4.4 Hypertext system and problems with linear information and traditional linkages

The concept of hypertext was first suggested by Vannevar Bush (1945). As the Director of the Office of Scientific Research and Development during the World War II, Bush oversaw thousands of American scientists as they concentrated their combined skills on weapons research. As the war moved on to its end, Bush believed science had to turn in another direction. He decided to consolidate all that hard-won war knowledge, and all the knowledge before and after that remained virtually inaccessible. Bush advocated the widely dispersed pieces of scientific knowledge be drawn together into a single entity - a whole much greater than the sum of its collective parts. He proposed a machine called 'memex', a device in which an

individual stored all books, records, and communications, and which was mechanized so that it might be consulted with exceeding speed and flexibility. According to his description, the memex resembled a desk with two pen-ready touch screen monitors and a scanner surface. It was supposed to enable us to search and retrieve vast amount of information stored in it. However, the essential feature of the memex laid not only in its capacities for retrieval and annotation but also in those involving ‘associative indexing’ - what presented hypertext systems term a link – ‘the basic idea of which is a provision whereby any item may be caused at will to select immediately and automatically another’ Bush's description of how a memex user created and then followed links joined his major recognition that trailed of such links themselves constituted a new form of textuality and new form of writing. As Bush (1945, p. 103) explained;

“When numerous items have been thus joined together to form a trail . . . It is exactly as though the physical items had been gathered together from widely separated sources and bound together to form a new book.”

Bush (1945, p. 104) also said;

“It is more than this for any item can be joined into numerous trails’, and thereby any block of text, image, or other information can participate in numerous books.”

Although the concept of hypertext was first proposed by Bush, it was Ted Nelson who coined the term ‘hypertext’. Nelson (1965), one of Bush's most prominent disciples, defined hypertext as non-sequential writing. More accurately, hypertexts are documents (or, collections of documents) that provide explicit support for non-sequential reading. These documents have a data structure that organizes the content

material, and a user interface that mediates users' access to the contents. Nelson was the leader of Xanadu project, which tried to connect all literature in the world into a huge online hypertext system.

Hypertext is a technology of non-linear writing and reading. In other word, hypertext is a text composed of blocks of words (or images) linked electronically by multiple paths, chains, or trails in an open-ended, perpetually unfinished textuality described by the terms link and node. Hypertext construct information in network structure, and enables users to search related information directly. Hypertext consists of nodes and links. Nodes store information, and links link related nodes (information) each other. Therefore, the pieces of information are called nodes; the connections or cross-references between nodes are called links. Together they form a hypertext document. The nodes and links can be viewed as forming a graph, which may be arbitrarily complex. Below is a simplified view of an extremely small hyperdocument, having only five nodes and seven links. This figure also shows that links are tied to a specific point (or word or region) within a node, called an anchor.

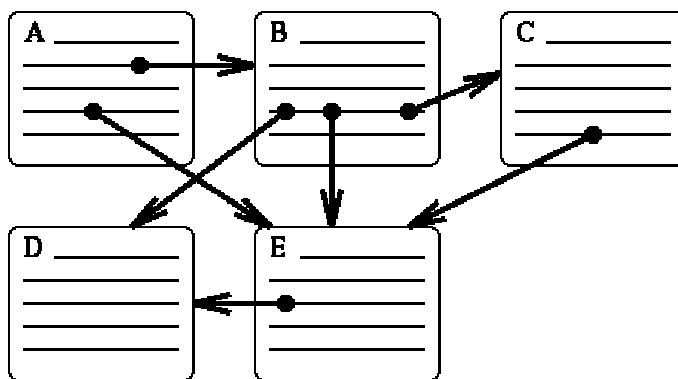


Figure 1. The concept of hypertext (Definition of hypertext and hypermedia)

Each node usually stores information that can be shown on one screen. The kinds of information can be stored in nodes are various. They include text, video, animation, and graphics. A link is a device that link related nodes. Links may link words and phrases within a hypertext document, or link a piece of information in a hypertext document to other hypertext document. In paper documents there are a few limited forms of links as well. The index is a source of links, but it is not possible to go directly from a word in the book to one of the pages indicated in the index, without first jumping to the index and then to the desired page. Examples of direct links are references to the bibliography, and, more importantly, footnotes. Hypertext is sometimes called the “generalized” footnote, because footnotes indeed have a visible element indicating the presence of a link, and the destination of this link is a separate information node. Because of the way footnotes are consistently placed in all books, following a footnote-link is something all readers of books seem to master. Remde et al. (1987, p. 181) enumerated several deficiencies of traditional, linear text that motivated the creation of hypertext interfaces. Among other reasons, they suggest that it is “too hard to find information in ordinary text, and too hard to acquire information in a sequence other than that determined by the author...extremely difficult to integrate and update large bodies of frequently changing information from many different sources.”

As briefly mentioned above, traditional linking methods within text are usually one-way streets. They only point to the next piece of information and rarely enable user to retrace the path to where a particular reference was indicated in a document. Also, traditional references could only point the user in one of two directions: they

were strictly limited to referencing information either earlier or later in the text. On that point, the most precious advantage of hypertext systems is the ease of traversing reference trails of information. The use of a computer to follow links makes retrieval effort necessary for almost all information chunks equivalent and provides an easy means to return to a specific citation. Hypertext systems permit the imposition of various types of structure on chunks of unstructured information. The system makes the organization of information into hierarchical, relational, or network structures relatively straightforward.

Hypertext systems are of course not applicable to all areas. Schneiderman (1989, p. 82) suggested the terms of relevant areas of hypertext system:

- 1) *The whole set of information consists of lots of small information fractions.*
- 2) *The small information fractions are related to each other.*
- 3) *Users need just small number of information fractions at one time.*

From the above areas, it could be said that the library catalog is one of the possible areas where hypertext systems are applicable among various library functions.

Isikawa (1991, p. 33) explained the applicability of hypertext system to library catalogs, and suggested new functions that may be added to current library catalogs:

- 1) *Browsing from one record to other records through hyperlinks*
- 2) *Faster search of various access points*
- 3) *Links to authority file*
- 4) *Provision of multi-media information*
- 5) *Checking the correctness of database*

The weaknesses of traditional text mentioned above also exist in traditional library catalogs. Current catalog systems, whether online or web catalogs, do not fulfill the users' need for free navigation from one catalog record to related records. As mentioned above, they also do not match with the nature of human thinking either. We navigate freely from an idea to associated ideas, when we think about something. However, the linear structure of traditional texts and library catalogs prevents us from free navigation.

4.5 Linking devices adopted in metadata formats

4.5.1 Library catalogs

Various devices have been used to link related bibliographic records in the library catalog. Related bibliographic records mean the records which share one or more bibliographic relationships. Linking devices encompass catalog entries, cross-references, uniform titles, authority control, and others.

1) Catalog entries

Leazer (1993, p. 40) stated that

Contemporary library catalogs, both manual and computerized, are linear files that comprise a long sequence of individual bibliographic descriptions. Most manual catalog descriptions are arranged in a single long alphabetical arrangement. Each record includes a basic description

of the item so that the described item can be uniquely identified. A basic description is reproduced so that it can be placed into several locations in the catalog. Each copy of the description receives a heading, which becomes that copy's point of entry in the alphabetical sequence of the file. Any individual item described in the catalog may receive a number of entries in the catalog; entries are commonly provided for the author's name, the item's title, and one or more subject headings derived from a list of subject headings. The structure of the catalog in the computer environment is basically the same as in the manual environment. However, the requirement that descriptive records be placed into physical sequence is not necessary, nor is a copy of the record required for each entry point in the catalog. Instead, a single descriptive record contains all access points for that record.

a. Main entries

One method for linking separate bibliographic records in the use of a common main entry in order for two or more related bibliographic records to display together in a catalog. This is particularly useful for showing descriptive relationships, whole-part relationships, and accompanying relationships. (Tillett, 1991)

b. Added entries

A device to provide supplementary access to bibliographic records from names, titles, and subjects closely associated with the bibliographic items being described. (Tillett, 1991)

c. Analytical entries

An entry for a part of some whole work. In other words, an entry for a bibliographic record of a part of a bibliographic item for which a comprehensive record may be made. (Tillett, 1991)

d. Multilevel description

An entry based on multi-level description. It is used for whole-part relationships and accompanying relationships. (Tillett, 1991)

2) Cross-references

Reference is a direction from one heading or entry to another. Reference plays a very important role in the success of a dictionary catalog. A *see reference* directs a user from one heading which the user might be reasonably expected look, to the form that has been chosen by the cataloger as a heading. A *see also reference* directs a user from one heading to a related heading.

3) Uniform titles

According to AACR2 (1998, p. 624), uniform title is

1. *The particular title by which a work is to be identified for cataloging purposes.*

2. *The particular title used to distinguish the heading for a work from the heading for a different work.*
3. *A conventional collective title used to collocate publications of an author, composer, or corporate body containing several works or extracts, etc., from several works (e.g., complete works, several works in a particular literary or musical form).*

In other words, a uniform title used for following purposes.

- A device to collocate materials by their general form
- A device to differentiate among identical or similar titles, and
- A device to link manifestations of a work

4) Authority control

Leazer (1993) defined that authority control was a technique for controlling and expressing data elements. It is needed, because two or more records may share the same value for a common data element, and the data must be in the same form for all records. He also stated that (1993, p. 41)

A variety of techniques are used to standardize data values between records. One technique is to provide a set of instructions for the formulation of the data element. Another technique is to keep a secondary list or file of all headings used for that particular kind of element. The Library of Congress Name Authority File (LCNAF) primarily for the control of names, and the Library Congress Subject Headings (LCSH) for the control of subject headings are examples of the technique. These

authority files list all authorized headings in their acceptable forms, and occasionally include instructions for the application of the heading. Most authority lists also refer from variant forms of the heading to the proper form through a network of “see” and “see also” references. This network of references between unauthorized headings and their authorized forms constitutes one portion of the library catalog’s synthetic structure.”

5) Other linking devices

Tillet (1991, p. 15) identified following additional linking devices.

Notes

In bibliographic records, notes are used to provide further information about the bibliographic descriptions and bibliographic history of the items described. Notes pertaining to bibliographic relationships give information about title variations, series, full or partial contents, specific library holdings, or related works.

Edition statement

It alerts the user to the other editions of the item described.

Series statement

It informs the user that the item in hand is a part of a series.

These linking devices have evolved along with the type of catalog available. Which means that these linking devices were designed to take advantage of the book and card formats of the catalogs. Without thought of improvements that might be made in the computer environment, many of the old devices for expressing relationships were

embodied in online and web catalogs. New linking devices that take advantage of the computerized and networked environment have been made possible partly with the advent of new metadata schemes such as MARC and Dublin Core.

4.5.2 MARC21

MARC21 is the standard machine-readable catalog format of bibliographic records in the United States and Canada. Libraries have used MARC21 as their machine-readable data format of catalog records, and almost all computerized catalog systems have adopted MARC21 format as their basic data structure. Although the local implementation of MARC21 records varies from setting to setting, MARC21 exists as a common record format at the core of almost all major library catalog systems in the United States. The MARC21 formats describe the structure of a single MARC21 record. Each MARC21 record for bibliographic data is a representation of an individual bibliographic entity. A collection of a number of MARC21 records makes up a catalog. A MARC21 records itself consists of a lot of fields and sub-fields.

Tags 760-787 of the MARC21 contain information that identifies and links other bibliographic items. They are called linking entry fields. Each of the linking entry fields specifies a different relationship between the target item described by the record and a related item. These relationships fall into three categories (MARC21):

- 1) *Related items that assist the user in continuing to search but are not physically required to obtain the target item. (e.g., former entries for serials, translations of the target item)*

2) *Related items that have to be obtained physically in order to use the target item. (e.g., the host item for a component part: a journal issue containing a specific article)*

3) *Related items that are constituent units of a larger whole (e.g., the individual photographs contained in a visual material collection).*

The linking entry fields are designed to generate a note in the record in which they appear. They can also provide machine linkage between the bibliographic record for the target item and the bibliographic record for the related item, if the related item is covered by a separate record. The name and content of each field is in the following (MARC21):

1) *760 Main series entry*

Information concerning the related main series when the target item is a sub-series

2) *762 Sub-series entry*

Information concerning a related sub-series when the target item is a main series or a parent sub-series

3) *765 Original language entry*

Information concerning the publication in its original language when the target item is a translation

4) *767 Translation entry*

Information concerning the publication in some other language other than the original when the target item is in the original language or is another translation

5) 770 *Supplement/special issue entry*

Information concerning the supplement or special issue associated with the target item but cataloged and/or input as a separate record

6) 772 *Supplement parent record entry*

Information concerning the related parent record when the target item is a single issue, supplement or special issue of the parent item

7) 773 *Host item entry*

Information concerning the host item for the constituent unit described in the record

8) 774 *Constituent unit entry*

Information concerning a constituent unit associated with a larger bibliographic Unit

9) 775 *Other edition entry*

The entry for another available edition of the target item

10) 776 *Additional physical form entry*

Information concerning another available physical form of the target item

11) 777 *Issued with entry*

Information concerning the publication that is separately cataloged but that is issued with or included in the target item

12) 780 *Preceding entry*

Information concerning the immediate predecessor of the target item

13) 785 *Succeeding entry*

Information concerning the immediate successor to the target item

14) 786 Data source entry

Information pertaining to a data source to which the described item is related. It may contain information about other files, printed sources, or collection procedures.

15) 787 Non-specific relationship entry

Information concerning the work related to the target item when the relationship does not fit any of those defined in fields 760-785.

Tag 856 of MARC 21 contains electronic location and access information for the item cataloged. It is for (MARC 21)

The information needed to locate and access an electronic resource. The field may be used in a bibliographic record for a resource when that resource or a subset of it is available electronically. In addition, it may be used to locate and access an electronic version of a non-electronic resource described in the bibliographic record or a related electronic resource.

Since a lot of information resources are available through the Internet, 856 tag can be used usefully to link the electronic version of the target item cataloged and the catalog record of the item.

4.5.3 Dublin Core

Metadata is structured data about information resources. MARC 21 is the metadata format for the library community. The most popular metadata format is Dublin Core. The Dublin Core has been developed by an informal group of librarians, networking and content specialists. Since its creation in March 1995 in what is now called the first Dublin Core Metadata workshop, the Dublin Core has gained popularity as the format for describing document-like objects in the Web. Dublin Core metadata is specifically intended to support resource discovery. The elements represent a broad, interdisciplinary consensus about the core set of elements that are likely to be widely useful to support resource discovery. Anyone can use Dublin core elements as a convenient basis for descriptive systems. Web pages are one of the most common types of resources to have Dublin Core descriptions, often using HTML meta-tags.

Dublin Core has 15 data elements that have been defined for identification of the Web resources. 15 data elements are specified with minimal stipulation of content rules in keeping with the original intent for simplicity and flexibility of use. In those elements, relation elements and source elements are used to link related resources (Dublin Core Metadata Initiative).

1) Relation

The relation element is an identifier of a second resource and its relationship to the present resource. This element permits links between related resources and resource descriptions to be indicated. Examples include an edition of a work (IsVersionOf), a

translation of a work (IsBasedOn), a chapter of a book (IsPartOf), and a mechanical transformation of a dataset into an image (IsFormatOf). (Dublin Core Metadata Initiative website)

A list of types which accommodates most expected relationships is:

IsPartOf	HasPart
IsVersionOf	HasVersion
IsFormatOf	HasFormat
References	IsReferencedBy
IsBasedOn	IsBasisFor
Requires	IsRequiredBy

(Dublin Core Metadata Initiative website)

Usage examples:

Title="Reading Turgenev"

Relation = "IsPartOf Two Lives"

Title="Candle in the Wind"

Subject="Diana, Princess of Wales"

Date="1997"

Creator="John, Elton"

Type="sound"

Description="Tribute to a dead princess"

Relation="IsVersionOf Elton John's 1976 song Candle in the Wind"

Title="Gombrich's Story of Art"

Relation="HasVersion 13th Edition, 1972"

(Dublin Core Metadata Initiative website)

2) Source

The source element contains information about a second resource from which the present resource is derived. While it is generally recommended that elements contain information about the present resource only, this element may contain a date, creator, format, identifier, or other metadata for the second resource when it is considered important for the discovery of the present resource. Source is not applicable if the present resource is in its original form. In general, the source element includes information which does not fit easily into the relation element. (Dublin Core Metadata Initiative website)

Usage example:

`<META NAME="DC.Source" CONTENT="RC607.A26W574 1996">`

(Dublin Core Metadata Initiative website)

4.5.4 VRA Core (Visual Resources Association's Core Categories for Visual Resources)

VRA Core Categories are intended as guidelines for describing visual documents depicting works of art, architecture, and artifacts or structures from

material, popular, and folk culture. They consist of a single element set that can be applied as many times as necessary to create records to describe works of visual culture as well as the images that document them. The Data Standards Committee followed the "1:1 principle," developed by the Dublin Core community, i.e., only one object or resource may be described within a single metadata set. How the element sets are linked to form a single record is a local database implementation issue. The order of the categories in the VRA Core is arbitrary, and local implementations are encouraged to determine their own field sequence that will appropriately describe their data.

The VRA Core 3.0 (VRA Core Categories website), which is the most current version, is intended as a point of departure - not a completed application. The elements that comprise the Core are designed to facilitate the sharing of information among visual resources collections about works and images. These elements may not be sufficient to fully describe a local collection and additional fields can be added for that purpose. The Committee recommends the use of qualifiers with certain elements in the VRA Core 3.0 so that the data values contained in the element may be more precisely identified. For instance, a "Notes" qualifier to clarify the data may be an appropriate addition to many of the current elements. Furthermore, every element may be repeated as many times as necessary within a given set to describe the work or image.

1) Relation

Terms or phrases describing the relationship between the Work or Image being cataloged and other Works or Images. Relationships can be whole/part (which occurs when one or more parts are dependent upon the whole, e.g., a series) or they might be associative (when two or more Works or images share a relationship through association). (VRA Core Categories website)

Proposed list for relationship types

part of	larger context for
larger entity	sketch for
based on	cartoon for
model for	study for
plan for	document for
document of	prototype for
copy after	copy of
original of	facsimile of
version of	format of
references	referenced by
derived from	source for

(VRA Core Categories website)

Usage examples:

Relation.Part of = Part of Disasters of war

*Relation.derived from = Drawing by Georg Pencz in the Staatsarchiv,
Nuremberg, Germany*

*Relation.source for = Triumphal Arch of Maximilian II (1570)
(VRA Core Categories website)*

2) Source

A reference to the source of the information recorded about the work or the image. For a work record, this may be a citation to the authority for the information provided. For an image, it can be used to provide information about the supplying Agency, Vendor or Individual; or, in the case of copy photography, a bibliographic citation or other description of the image source. In both cases, names, locations, and source identification numbers can be included. (VRA Core Categories website)

Usage examples:

Source = University of Michigan Museum of Art

*Source = Indian bronze masterpieces: the great tradition: specially
published for the Festival of India*

(VRA Core Categories website)

4.6 MARC or metadata?

Library community has long been the center of bibliographic control of many different formats of information resources. To fulfill its mission of bibliographic control center, library community has adopted MARC format as the standard bibliographic data format. For almost 30 years, the MARC format has been successfully used without great change. However, with the advent of the Web, a variety of metadata formats have been emerged to be alternatives of the MARC format. It is partly because the amount of Web resources are too huge and are growing too rapidly, and partly because the nature of the Web itself is constantly changing and unstable. Due to these characteristics of the Web, the library community seems to have failed to meet the needs of bibliographic control of the Web resources. The other reason is because the MARC format is too extensive and esoteric to be used outside the library community.

The Dublin Core is one of the most popular and common metadata formats suggested as an alternative of MARC for Internet resources. The Dublin Core is becoming the standard of metadata formats. With the broad acceptance of metadata formats represented by Dublin Core outside the library community, the library community has been concerned about interoperability between MARC and Dublin Core.

Currently, co-existence of MARC and the Dublin Core seems to be inevitable. Michael Gorman (1999, p. 22) mentioned that

We now have four approaches to the bibliographic control of electronic resource – full cataloging by MARC; enriched Dublin Core records; minimal Dublin Core records; and relying on unstructured full-text

keyword searching ... relatively few electronic resource that are seen as having continuing value could be cataloged fully in accordance with national and international standards resulting in full MARC records

The Cooperative Online Resource Catalog (CORC), a project undertaken by OCLC, is a major example of efforts for MARC/Dublin Core co-existence. CORC includes automated record creation, descriptive cataloging, subject headings assignment, classification, and dynamic page building that libraries can integrate with their gateways. CORC calls itself as "a metadata creation system for bibliographic records and pathfinders describing electronic resources," the goal of which is to use "both MARC and Dublin Core records to create a searchable database of quality Internet resources" (OCLC CORC). CORC's interoperability is one of its major features. Contributors can enter records in either MARC or Dublin Core format. These records are then stored in Extensible Markup Language (XML), and delivered or exported to the end-user in either MARC or Dublin Core format.

As seen from CORC, it is now needed for us to consider providing bibliographic records for Web resources through the library catalog. Dublin Core records can be included in the library catalog by converting them to MARC records. If we do so, library OPACs can be the gateway to electronic and Web resource as well as traditional library resources.

5. Proposed solution

5.1 Bibliographic relationships

Bibliographic relationships are central to the objectives of the library catalog. Without them the catalog's collocating functions could not be met. Bibliographic units, whether it is bibliographic or literary, exist in a network of relationships. In other words, bibliographic units have relationships with each other. These bibliographic relationships can be defined as associations between two or more bibliographic units (Tillett 1991). Monographic as well as serial publications were considered as bibliographic units including publications at the analytical level. Tillett (1987) analyzed 24 cataloging codes to determine their capabilities to express bibliographic relationships. A test of machine-readable records of the Library of Congress resulted in a comprehensive list of bibliographic relationships. She defined seven relationships in total. The library catalog needs to link these bibliographic relationships, in order to enable users to search related bibliographic records from one record.

5.2 Merge of bibliographic records for Web resources and traditional library resources

The library catalog needs to be a gateway to both of Web resources and traditional library resources. For the library catalog to be a gateway, all different information resources including Web resources that have continuing value have to be cataloged in MARC format by professional catalogers.

The cataloging of Internet resources differs from print resources primarily in two areas: the description of the changing characteristics of Internet resources and the provision of access to those resources. The basic rules used for cataloging in the United States are the *Anglo-American Cataloguing Rules* Second Edition (AACR2). International Standard Bibliographic Description for Electronic Resources (ISBD(ER)) is generally accepted as the equivalent AACR2 for describing Internet resources.

A persistent problem with cataloging Internet resources is that the creators of those resources often do not provide the standard bibliographic information that AACR2 requires. The terms 'author', 'title', and 'publication information' frequently do not have the same meaning when applied to Internet resources. Therefore, some changes to the cataloging rules are needed for if Internet resources are to be handled effectively. One problem that needs to be addressed concerns the multiple versions that must be created if AACR2, specifically Rule 0.24, is followed literally. That rule states that a new record must be created if there is a variation in the physical carrier between two documents, even if the documents have the same intellectual or artistic content. (Anglo-American Cataloging Rules, 1998, p. 8) This means, for example, that both the print and online versions of a serial must be cataloged as separate

records, even if the content is identical. Cataloging identical items separately based only on differences in physical carrier could result in hours of wasted time for the cataloger and frustration for the catalog user who must sort through many retrieved records that contain the same basic content.

The ways of combining MARC and metadata to improve access to materials has been continuing to be explored. As a result, tag 856, Electronic Location and Access in the MARC was created in 1994. This allowed the URL to be placed directly in the MARC record.

Integration of bibliographic records for Web resources and traditional library resources is now inevitable. Most of all it is called in the name of serving the users by providing access to information resources through library OPACs. McCallum (2000) said,

The discussions in the library community have thus focused on description and retrieval of electronic resources along with other physical information media, rather than separately. Distinguishing characteristics of electronically presented material are identified so that established cataloging principles for these other media can take electronic documents into account. It is clear that the MARC format can provide a vehicle for the description of Web and networked resources, and has kept up-to-date with developments in the medium. This is very positive and reassuring given the large community investment in MARC-based control – the vast body of important non-electronic resources to which MARC is the key to

interchange and the cost saving bibliographic services and tools that have been built on the standard format.

5.3 Concept of the hypercatalog

O'Neill and Vizine-Goetz (1989, p. 179) expected that

The widespread use and acceptance of online catalogs presents us with an opportunity to make enhancements for the library catalog. However, current online library catalogs do not really utilize the opportunity, because they inherited the data structure of traditional library catalogs. The real enhancement to computer-based catalogs can be made only by improving their data structures to fit into computerized and networked environment. Computers can support complex data structures that would be impossible to use in card catalogs. However, this underlying complexity can be transparent to the user since display mechanisms can create display records customized to meet the needs of individual users. With well-designed data structure and interface, online catalogs can meet both the first and the second objectives of the catalog without compromise.

Unlike the linking devices adopted in current library catalogs, the mechanism to collocate various manifestations of a work can be based on the bibliographic relationships for establishing physical links between the units concerned. The core elements of hypertext system are nodes and links. In the library catalog, the node is individual bibliographic record, and the links are bibliographic relationships between

bibliographic records. As reviewed above, the library catalog has adopted various linking devices to link related bibliographic records. However, they are not systematically organized, and often out-dated. Moreover, the linking structure of current library catalogs is one-way. For example, catalog records only refer users from translated versions to original version, do not refer from the original version to translated versions. Users also have to search the catalog again after they get information on related items from the catalog, because current library catalogs do not directly link related bibliographic records.

If hypertext functionality is added to current library catalogs, users would be able to navigate from one bibliographic record to related records through hyper-links. The hyperlinks would connect bibliographic records that share one or more bibliographic relationships. It is expected that the catalog with hyper-links which will be called as hypercatalog hereafter would greatly enhance “collocation function” which is weak in traditional catalogs, because the hypercatalog physically and directly link related bibliographic records through hyperlinks.

6. Design of hypercatalog applying bibliographic relationships

In previous chapters, limitations of traditional library catalogs were reviewed and a possible method for overcoming the limitations was suggested. The suggestion was that the weak collocating function in the traditional library catalogs would be greatly improved, if we apply hypertext technology to them. The core concepts of hypertext are nodes and links. In the library catalogs, nodes are individual catalog records and links are bibliographic relationships between them. Therefore, in the hypercatalog a hyperlink will be made through each bibliographic relationship. As mention in the literature review, Bertha suggested a similar model in 1993. Her model provided following links for related publications. (Bertha, 1993, p. 222)

- *Additional copies*
- *Variant editions*
- *Translations*
- *Predecessor or successor titles*
- *Parent or analytical publications*
- *Secondary publications*

She also speculated the relationships between bibliographic units were similar to the relationships between descriptors in a thesaurus. She suggested, therefore, concepts

and experiences of thesaurus could be applied to create a system for linking bibliographic units. (Bertha, 1993)

Even though Bertha's model was the first to apply bibliographic relationships for the hypertext-based library catalog, it didn't cover the entire bibliographic relationships of Tillett. Among Tillett's seven relationships, Bertha's model lacked the accompanying relationships. She didn't suggest detailed methods for realization of hypercatalog either. This paper proposes a model of hypercatalog through suggestion of data structure and interface design.

6.1 Data structure

Various MARC formats have fields for linking records that share one or more bibliographic relation. The name is not same among various formats. For example, MARC21 calls them linking entry fields and UNIMARC calls them linking entry block. However, these fields are used to display relations in bibliographic units and link related records directly. It is thought that these fields can be used for constructing a hypercatalog.

6.1.1 UNIMARC linking entry block

Each field in linking entry block contains sub-fields that identify the bibliographic item to which the link is being made. The data in this field made to be sufficient to

identify the record for the item being linked to, or, if there is no record, to identify the item itself.

Following table is the linking entry block of UNIMARC with Tillett's bibliographic relationship for each field.

Table 1. Linking entry block of UNIMARC

Field	Name	Bibliographic relation
410	Series	Whole-part
411	Sub-series	Whole-part
421	Supplement	Accompanying
422	Parent of supplement	Accompanying
423	Issued with	Equivalence
430-437	Preceding entries	Sequential
440-448	Succeeding entries	Sequential
451	Other edition in same medium	Derivative
452	Other edition in another medium	Derivative
453	Translated as	Derivative
Field	Name	Bibliographic relation
454	Translation of	Derivative
455	Reproduction of	Equivalence
456	Reproduced as	Equivalence
461	Set level	Whole-part
462	Sub-set level	Whole-part

463	Piece level	Whole-part
464	Piece-analytic level	Whole-part
470	Item reviewed	Descriptive
481	Also bound with this volume	N/A
482	Bound with	N/A
Field	Name	Bibliographic relation
488	Other related works	N/A

6.1.2 CCF (Common Communication Format)

As mentioned in the literature review, CCF was developed to provide a bridge between the major international exchange formats. It was also intended to be a communication format between different information-related agencies including libraries and index/abstract services. Record linking has always been more important for the secondary services than for national bibliography production because they wish to provide records at different bibliographic levels in a common database. Nevertheless, UNIMARC has had from the outset a record linking technique and many national MARC formats also include record linking. The real difference is that secondary services cannot begin to function without their databases taking into account bibliographic levels and their relationships; the vast majority of records from national libraries, in percentage terms, do not contain any explicit links. Indeed, some national MARC formats still do not include a facility for record linking. It is interesting to look at the complex record the developers of the CCF had in mind when

the CCF was being devised. Below figure is an example of relationships in a complex CCF record.

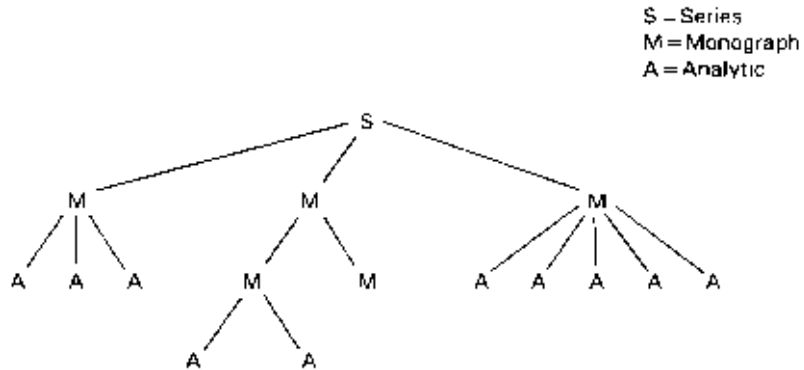


Figure 2. Example of relationships in a complex CCF record (CCF)

CCF has following segment-linking fields to provide links between related records.

Each field is displayed with Tillett's bibliographic relationship for each field.

Table 2. Segment linking fields of CCF

Field	Name	Bibliographic relation
080	General hierarchical relation	Whole-part
081	Hierarchical relation of single-volume set	Whole-part
082	Hierarchical relation of multi-volume set	Whole-part
083	Hierarchical relation of serials	Whole-part

085	Horizontal relation	Equivalence, derivative, accompanying, descriptive, sequential
-----	---------------------	--

6.1.3 MARC21 linking entry fields

Fields 760-787 of the MARC21 contain information that identifies other bibliographic items. Each of the linking entry fields specifies a different relationship between the target item described by the record and a related item. These relationships fall into three categories (LC MARC21):

- 1) *related items that assist the user in continuing to search but are not physically required to obtain the target item (e.g., former entries for serials, translations of the target item)*
- 2) *related items that have to be obtained physically in order to use the target item (e.g., the host item for a component part: a journal issue containing a specific article)*
- 3) *related items that are constituent units of a larger whole (e.g., the individual photographs contained in a visual material collection).*

The linking entry fields are designed to generate a note in the record in which they appear. They can also provide machine linkage between the bibliographic record for the target item and the bibliographic record for the related item, if the related item is covered by a separate record.

When we relate each linking entry field with Tillet's bibliographic relationships, we get the relationships in the following table.

Table 3. Linking entry fields of MARC21

Field	Name	Bibliographic relation
760	Main series entry	Whole-part
762	Sub-series entry	Whole-part
765	Original language entry	Derivative
767	Translation entry	Derivative
770	Supplement/special issue Entry	Accompanying
772	Parent record entry	Accompanying
773	Host item	Whole-part
774	Constituent unit entry	Whole-part
775	Other edition entry	Derivative
776	Additional physical form Entry	Equivalent
777	Issued with entry	Accompanying
780	Preceding entry	Sequential
785	Succeeding entry	Sequential
786	Data source entry	Derivative

787	Non-specific relationship Entry	N/A
-----	------------------------------------	-----

6.1.4 Conversion of Dublin Core records to MARC21 records

Library Congress and OCLC sponsored several research efforts on the conversion between MARC and Dublin Core formats. Library of Congress's MARC documentation website (<http://lcweb.loc.gov/marc/marcdocz.html>) provides specifications of conversion between them. OCLC CORC (Cooperative Online Resource Catalog) project has been successful in the conversion. Through the OCLC CORC website (<http://www.oclc.org/corc/>), conversion between MARC21 records and Dublin Core records can be made automatically. With the success of CORC project, OCLC WorldCat now provides cataloging records for electronic and web resources along with records for traditional library resources. By the efforts of those two organizations, interoperability between MARC21 and Dublin Core has been ensured.

6.1.5 Data structure of hypercatalog

MARC21 is the standard format of bibliographic records. Therefore, it is needed to use the MARC21 format as the record structure of the hypercatalog. Linking entry

fields of MARC21 can be used to directly link related bibliographic records.

However, the linking entry fields lack fields for linking bibliographic records that share descriptive relationships among Tillett's seven relationships. The fields also lack fields for linking bibliographic records that share derivative relationships except translations and editions. In order to encompass all seven bibliographic relationships of Tillett, we need to expand the linking entry fields of MARC21 to link all of them.

To link records which share descriptive relationships, 778 (mother record entry) and 779 (descriptive relation entry) are added to current linking entry fields. To link records which share derivative relationships, 768 (original record entry) and 769 (derivative relation entry) are also added.

The specification of each linking entry fields is in the following:

760 same

761 same

765 same

767 same

768 same with 765 except second indicators

second indicators: # summary of

1 revision of

2 condensation of

3 adaptation/modification of

4 genre change of

5 paraphrase of

6 other version of

769 same with 767 except second indicators

second indicators: # summarized as

1 revised as

2 condensed as

3 adapted/modified as

4 genre changed as

5 paraphrased as

6 other version of

770 same

772 same

773 same

774 same

775 same

776 same

777 same

778 same with 765 except second indicators

second indicators : # criticism of

1 evaluation of

2 review of

3 annotation of

779 same with 767 except second indicators

second indicators : # criticized as

1 evaluated as

2 reviewed as

3 annotated as

780 same

785 same

786 same

787 same

Also Sub-field w of each field is changed to not repeatable (NR).

6.2 Interface design of hypercatalog

On the full bibliographic record screen of OPAC, the hypercatalog provides “hot” links to search related bibliographic records from one record. In current OPACs, users have to retype author name or title to search related records from one record, or browse again the search result screen. Unlike current OPACs, the hypercatalog would enable users to search related records from one record through hyperlinks by just clicking the links on the screen.

Sompel (1999) mentioned that

In the context of networked library services, the necessity to integrate secondary data, catalogues and primary information has been expressed quite some time ago. More specifically, librarians have brought to the fore the need to link abstracting databases with library catalogues; catalogues with primary information abstracting databases with full-text primary

information. These specific linking notions have evolved towards a concept of connecting all the available information, in order to come to a fully interlinked information environment.

Therefore, it is needed in a networked environment the library catalog may be linked to the Internet-based bibliographic information providers. For example, if a user searches for “Iliad” by Homer, it can be imagined the following full bibliographic record screen.

Figure 3. Example of full bibliographic information screen of hypercatalog

Title	: Homeri Ilias / recensuit Arthurus Ludwich.
Uniform title	: <u>Iliad</u>
Author	: <u>Homer</u> .
Published	: New York : <u>Academic Press</u> , 1995.
Edition	: Editio stereotypa, editionis primae (1902)
Subject	: <u>Achilles (Greek mythology) --Poetry</u> . <u>Trojan War --Poetry</u> . <u>Epic poetry, Greek</u> .
Series	: <u>Bibliotheca scriptorum Graecorum et Romanorum Teubneriana</u>
Other Author	: <u>Ludwich, Arthur, 1840-1920</u> .
Other titles	: Illiad.
Material	: 232 p. ; 22 cm.
Notes	: Text in Greek, with prefatory notes and commentary in Latin. Includes bibliographical references and index.
OCLC no	: <u>33001061</u>
ISBN	: <u>3519042835</u>
<u>Related records</u>	

The underlined parts are hot links. Links are provided for author names, uniform titles, series titles, publishers, subject headings, OCLC number, ISBN/ISSN, electronic location and access information (provided through 856 field) and related records. If a user clicks the 9 links, each link activates followings:

a. Uniform title

Search the OPAC by the title

b. Author name

Search OPAC by the author name, optionally it also could be linked to directory websites to provide additional information on the author

c. Publisher

Links to *Books in Print* website (<http://www.booksinprint.com/bip>) to show information on the publisher

d. Subject heading

Search the OPAC by the subject heading

e. Series

Display other items belong to the series

f. OCLC number

Search OCLC WorldCat by the number (Users can get holding lists though it)

g. ISBN

Links to *Books in Print* website to show information on the book, optionally it also could be linked to Amazon website (<http://www.amazon.com>) to show additional information on the book.

h. ISSN

Links to *Ulrich's Periodical Directory* website

(<http://www.ulrichsweb.com/ulrichsweb/>) to show information on the serial

i. Electronic location and access

Links to the electronic location

j. Related records

When a user clicks related links, the screen displays abridged bibliographic records (title, author, publisher, and publication year) of related records that share one or more bibliographic relationship for browsing under each bibliographic relation name as shown in the following.

Figure 4. Example of abridged bibliographic information screen of hypercatalog

Criticism			
Title	Author	Publisher	Year
<u>Tradition and design in the Iliad</u>	Bowra, C.M.	Clarendon Press	1930
<u>Athetized lines of the Iliad</u>	Bolling, George	Harper & Bros.	1956
<u>Essays on the Iliad : selected ...</u>	Wright, John	Oxford Univ. Press	1978
Annotation			
Title	Author	Publisher	Year
<u>The Iliad, a commentary</u>	Kirks, K.S.	Cambridge Press	1992

Other version			
Title	Author	Publisher	Year
<u>Homer : the Iliad</u>	Collins, W. Lucas	Blackwood	1961
Homer's the Iliad		Chelsea House	1987
The story of Iliad	Church, Alfred John	Macmillan	1972
The Iliad [Web resource]	Homer		

Users can also select each item to see full bibliographic record from the screen.

6.3 Linking between related bibliographic records

To enable above described hypercatalog, library catalogers need to search possible related resources and insert links manually in their library catalog records.

Links between related records can be made directly through record ID (Sub-field w: Record control number). Those links are not necessarily limited to the physically possessed resources by the library. Links can be made between locally available resources and electronically accessible resources or resources available through inter-library loan.

For a future research topic, we can think of methods for automatic generation of links between related bibliographic records. If the automatic generation is made possible in the future, it will save a lot of time and efforts of library catalogers.

6.4 Comparison of linking structure between traditional catalogs and hypercatalog

Table 4. Comparison of linking structure between traditional catalogs and hypercatalog

Functions		Hypercatalog	Card catalog	Online/Web catalog
Various access points		Possible	Impossible	Possible
Collocation Function	Equivalent relationship	2-way	One way	One way
	Derivative relationship	2-way	One way	One way
	Descriptive relationship	2-way	One way	One way
	Whole-part relationship	2-way	2-way	2-way
	Accompanying relationship	2-way	2-way	2-way
	Sequential relationship	2-way	2-way	2-way

Web resource	Included	Not included	Partly (through 856 field)
Method of collocation	Direct link	Indirect link	Indirect link

7. Conclusion

Current library catalogs do not fulfill users' needs for navigating freely from one record to related records. They also fail to meet the second objective of the library catalog either because of the generic defect of current linking devices and omission of virtual copies. The application of hypertext system to the traditional library catalog is suggested for a solution, because this type of OPAC system is able to present links in a more user-friendly way. The new system is named hypercatalog. The classical retrieval function of matching formal characteristics like the name of an author, publisher, uniform title, or subject headings is embedded in the system of relationship as well as Tillett's seven bibliographic relationships. An expanded USMARC format is suggested for data structure of the hypercatalog.

The hypercatalog is expected to be able to offer an appropriate interface to present links for related bibliographic records for both of Web resources and traditional library resources. The end users activate these links after a query using the hit as the starting point merely clicking one of links on the full bibliographic record screen.

References

Ayres, F.H. (1990). Duplicates and other manifestations – a new approach to the presentation of bibliographic information. *Journal of Librarianship*, 22(4), 236-251

Baptist, et al. (1996). *The Hypercatalog Graz-Budapest (HyperKGB)*
<http://www.asis.org/annual-97/hyperkgb.htm>

Bertha, E. (1993). Inter- and intra-bibliographical relationships : a concept for a hypercatalog. In A. H. Helal & J. W. Weiss (Eds.), *Opportunity 2000 : understanding and serving users in an electronic library* (pp. 211-224). Essen: Essen University Library.

Bush, Vannevar (1945). As we may think. *Atlantic Monthly*, 176 (7), 101-108

Carpenter, M. and Svenonius, E. (1985). *Foundations of cataloging : a sourcebook*. Littleton, Colo.: Libraries Unlimited

CCF : the common communication format (1988). – 2nd ed. – Paris: UNESCO

Cutter, C.A. (1904). *Rules for a dictionary catalog*. 4th ed. London: Library Association

Definition of hypertext and hypermedia [online]

Available: <http://www.wis.win.tue.nl/2L670/static/definition.html>

Dublin Core Metadata Initiative [online]

Available: <http://dublincore.org/>

Goossens, P. and Mazur-Rzesos, E. (1981). Hierarchical relationships in bibliographic description : problem analysis. In *Library systems seminar* (pp. 58-63). Essen: Essen University Library.

Gorman, M. (1982). Authority control in the prospective catalog. In M. W. Ghikas (Ed.) *Authority control : The key to tomorrow's catalog : Proceedings of the 1979 Library and Information Technology Association Institutes* (pp. 166-80). Phoenix: Oryx Press.

Gorman, M. (1999). Metadata or cataloging? A false choice. *Journal of Internet Cataloging*, 2 (1), 5-22.

Hjerppe, Reland (1986). Hypercatalog and three meta-schemata for database views: knowledge organizing, collection derived, and user established structures. In Kinsella, Janet, ed. *Online public access to library files: second national conference*. Oxford: Elsevier International Bulletins, 101-110

Hjerppe, Roland (1989). Hypercat at LIBLAB in Sewen: a progress report. In Hildreth, Charles R. ed. *The online catalogue. Development and directions*. London: The Library Association, 177-209

IFLA Study Group on the Functional Requirements for Bibliographic Records. (1998). *Functional requirements for bibliographic records : final report*. München : K.G. Saur

Isikawa, Eigaku. (1991). The possibility of hypertext system in library catalogs. In *The future model of library systems*. Tokyo : Toeisha

Kumar, Girja and Kumar, Krishan. (1981). *Theory of cataloging*. 3th ed. New Delhi: Vikas Publishing House

Leazer, G. H. (1993). *A conceptual plan for the description and control of bibliographic works*. D.L.S. dissertation, School of Library Service, Columbia University.

Library of Congress (2000). *MARC 21*.

<http://lcweb.loc.gov/marc/bibliographic/ecbdhome.html>

Lubetzky, Seymour (1953). *Cataloging rules and principles*, Washington: Library of Congress

Lubetzky, Seymour (1969). Bibliographic dimensions in information control. *American Documentation*, 20, 247-252

MacCullum, Sally. (2000). Extending MARC for bibliographic control in the Web environment : challenges and alternatives. In *the LC Bicentennial Conference on Bibliographic Control for the New Millennium*.

http://www.loc.gov/catdir/bibcontrol/mccallum_paper.html

Nelson, T. H. (1965). A file structure for the complex, the changing, and the intermediate. In *ACM 20th National Conference Proceedings*, 84-100

Nelson, T.H. (1987). *Literary Machines, edition 87.1*. Project Xanadu.

OCLC CORC [online]

Available: <http://www.oclc.org/corc/>

O'Neill, E.T. and Visine-Goetz, D. (1989). Bibliographic relationships: implication for the function of the catalog. In Svenonius, Eleaine, ed. *The conceptual foundations of descriptive cataloging*. San Diego : Academic Press, 167-179

Ranganathan, S.R. (1964). *Classified catalogue code with additional rules for dictionary catalogue code*. 5th ed. Bombay: Asia

Remde, J.R., Gomez, L.M., and Landauer, T.K. (1987). SuperBook: an automatic tool for information exploration – hypertext? In *Proceedings of Hypertext '87, San Antonio, TX*. ACM Press., 175-188.

Schneiderman, Ben (1989). *Hypertext hand-on*. New York : Addison-Wesley.

Smiraglia, R.P. (1987). Bibliographic control theory and nonbook materials. In S. Intner and R.P. Smiraglia (Eds.), *Policy and practice in bibliographic control of non-book media*. Chicago: American Library Association

Smiraglia, R. P. (1994). Derivative bibliographic relationships : linkages in the bibliographic universe. In D. L. Anderson & T. J. Galvin (Eds.), *Navigating the networks : proceedings of the ASIS mid-year meeting, Portland, Oregon, May 21-25, 1994* (pp. 167-83). Medford, NJ: Learned Information, Inc. for the American Society for Information Science.

Smiraglia, R. P. and Leazer, G. H. (1999). Derivative bibliographic relationships : the work relationship in a global bibliographic database. *Journal of the American Society for Information Science*, 50(6), 493-504.

Sompel, Herbert Van de (1999). Reference linking in a hybrid library environment. *D-Lib Maganize*, 5 (4)

Taylor, A.G., David P. Miller. (2000). *Wynar's introduction to cataloging and classification*. 9th ed. Englewood, Colo.: Libraries Unlimited

Tillett, Barabra Ann Barnett (1987). *Bibliographic relationships: toward a conceptual structure of bibliographic information used in cataloging*. University of California at Los Angeles , Ann Arbor : University Microfilms International

Tillett, B. A. (1991). A summary of the treatment of bibliographic relationships in cataloging rules. *Library Resources & Technical Services*, 35, 393-405.

Tillett, Barbara Ann Barnett (1991b). A taxonomy of bibliographic relationships. *Library Resources & Technical Services*, 35(2), 150-158

Tillett, B. A. (1992). Bibliographic relationships: an empirical study of the LC machine readable records. *Library Resources & Technical Services*, 36, 162-188.

UNIMARC Format (1980). -- 2nd ed. -- London : IFLA International Office for UBC.

VRA Core [online]

Available: <http://www.oberlin.edu/~art/wc1.html>

Verona, Eva (1959). Literary unit versus bibliographic unit. *Libri*, 9(2), 79-104

Weintraub, D.K. (1979). The essentials or desiderata of the bibliographic record as discovered by research. *Library Resources & Technical Services*, 23, 391-405

Wilson, Patrick (1989). The second objective. In Svenonius, Eleaine, ed. *The conceptual foundations of descriptive cataloging*. San Diego : Academic Press, 5-16