

Syracuse University

SURFACE at Syracuse University

School of Information Studies - Faculty
Scholarship

School of Information Studies (iSchool)

2014

Breaking Records: The History of Bibliographic Records and Their Influence in Conceptualizing Bibliographic Data

Rachel I. Clarke
Syracuse University

Follow this and additional works at: <https://surface.syr.edu/istpub>



Part of the [Cataloging and Metadata Commons](#)

Recommended Citation

Clarke, Rachel I., "Breaking Records: The History of Bibliographic Records and Their Influence in Conceptualizing Bibliographic Data" (2014). *School of Information Studies - Faculty Scholarship*. 170. <https://surface.syr.edu/istpub/170>

This Article is brought to you for free and open access by the School of Information Studies (iSchool) at SURFACE at Syracuse University. It has been accepted for inclusion in School of Information Studies - Faculty Scholarship by an authorized administrator of SURFACE at Syracuse University. For more information, please contact surface@syr.edu.

Breaking Records: The History of Bibliographic Records and Their Influence in Conceptualizing Bibliographic Data

Description/Abstract

A bibliographic record is a conceptual whole that includes all bibliographic information about a resource together in one place. With the Semantic Web, individual data statements are linked across the web. This position article argues that the traditional conceptualization of bibliographic records affects the affordances and limitations of that data. A historical analysis of the development of bibliographic records contrasted with the Semantic Web model reveals how the “record” model shaped library cataloging and the implications on library catalogs today. Reification of the record model for bibliographic data hampers possibilities for innovation in cataloging, inspiring a reconceptualization of bibliographic description.

Keywords

bibliographic records, Semantic Web, linked data, library cataloging

Disciplines

Cataloging and Metadata | Library and Information Science

Additional Information

This is an Accepted Manuscript of an article published by Taylor & Francis in *Cataloging & Classification Quarterly* on December 22, 2014, available online: <http://www.tandfonline.com/doi/full/10.1080/01639374.2014.960988>.

Creative Commons License



This work is licensed under a [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/).

Title: Breaking Records: The History of Bibliographic Records and Their Influence in Conceptualizing Bibliographic Data

Author: Rachel Ivy Clarke

Running Head: History and Conceptualization of Bibliographic Records

Keywords: bibliographic records, semantic web, linked data, library cataloging

Abstract: A bibliographic record is a conceptual whole that includes all bibliographic information about a resource together in one place. With the Semantic Web, individual data statements are linked across the Web. This position paper argues that the traditional conceptualization of bibliographic records affects the affordances and limitations of that data. A historical analysis of the development of bibliographic records contrasted with the Semantic Web model reveals how the “record” model shaped library cataloging and the implications on library catalogs today. Reification of the record model for bibliographic data hampers possibilities for innovation in cataloging, inspiring a reconceptualization of bibliographic description.

Introduction

Cataloging and bibliographic description are arguably fundamental pillars of librarianship, supporting the selection, management, and preservation of information. When collections of information become too large to organize and access directly, librarians create surrogate “bibliographic records” to represent specific items in a library’s collection, physical or digital.¹ The purpose of these bibliographic records is ostensibly to describe a document in sufficient detail to identify it uniquely among other documents and specify where the record can be located within in a file of other bibliographic records.² Originally, mere title and author information was enough to distinguish a work among others and order its record within a collection of records. But over time, the amounts and types of data that constitute sufficient detail to uniquely describe a resource and offer access to its description have grown. As more resources arose, more data was considered necessary to distinguish one resource from another. The conceptual model of a bibliographic record, then, is one of a *collection* of data elements.

This collective model served American librarianship well in its first 100 years. However, changes in the form and distribution methods of resources—especially in digital environments—raise questions about whether librarians’ traditional conceptual model is still relevant and applicable. Contemporary data models like the Semantic Web’s model of linked Data directly conflict with this traditional conceptualization of bibliographic records. A bibliographic record is a conceptual whole that includes all the bibliographic information about a resource together in one place, like a catalog card or a MARC record. With Linked Data, bibliographic data need not be a conceptual whole. Bibliographic data about a resource is no longer collected in a single location, but rather linked from many multiple locations across the web. What then, are the implications of the traditional conceptualization of the bibliographic record for librarianship?

This position paper argues that the way in which librarians conceptualize bibliographic data—as a “record”—affects the affordances and limitations of that data, especially in digital

environments. By tracing the development of the concept of the bibliographic record from early American librarianship through the present, and contrasting that model with the current Semantic Web model, this paper will reveal how the “record” model shaped library cataloging from early physical catalogs through contemporary digital software and interfaces.

Early American bibliographic records

Early American library catalogs in the colonial period were typically made available in book form, either manuscript or printed. American library catalogs had no systematically articulated or standardized rules for recording bibliographic information.³ Rather, they looked to 17th century English library catalogs, such as the ones from the Bodleian Library, for guidance. Bibliographic description in Thomas Hyde’s 1674 Bodleian Library catalog mandated the inclusion of the following elements⁴:

- author surname
- brief title
- number and size of volumes

On occasion, when warranted, entries might also include the date and place of publication and/or a note about the identity of the donor. What was the reasoning for including these data elements and not others? During the colonial period, little attention had been given to articulating the purposes of the library catalog. Many printed book catalogs were not designed to be finding aids, but rather served as marketing tools to acknowledge donations, publicize the collection, and solicit new patrons and further donations.⁵ It wasn’t until 1826, when the American Philosophical Society suggested that purpose of library catalogs was to show library holdings by subject and editions, that catalogs started to function as user tools.⁶ Once this designation was in place, standardized sets of rules became necessary.

The most famous early cataloging rules are Panizzi's 91 rules. Many cataloging scholars believe that our current cataloging structure can be traced back to these very rules.⁷ Panizzi stipulates that entries include the following data elements:⁸

- title
- author
- “any striking imperfection in a book...and any remarkable peculiarity”
- the number of parts, volumes, etc.
- place of publication, printer's name, date of publication
- printing notes (such as whether the work was printed on fine paper, vellum, etc.)

With the advent of the card catalog in American libraries in the 1850s, the shift to catalogs as aids for library users continued. By the late 1800s, Cutter convinced librarians that the purpose of the library catalog is to assist users in finding books by title, author and subject.⁹ This requires that subject data be added to the cadre of bibliographic data recorded about a work. While book catalogs grouped descriptive bibliographic entries by subject, catalog cards recorded subject heading information alongside the descriptive cataloging information, thus adding subject data elements to the burgeoning bibliographic record. The individual nature of cards—each card representing or acting as a surrogate for one specific resource—began to solidify the conceptualization of the bibliographic record as its own entity, one that carries a collection of data elements. While each line item or multi-line description in a book catalog was a record in that it recorded data about a resource, it was the full catalog that was the conceptual whole. Book catalog entries were not separate entities that could be managed or manipulated individually. The transition to individual slips or cards affords the treatment of bibliographic records as items in and of themselves.

Bibliographic records in the 20th century

In the early 1900s, the Library of Congress demonstrated support for card catalogs when they began their Card Distribution Service. In this era, what become known as “added entries” begin to be added to bibliographic records in addition to the previously stipulated elements. This data was intended to represent conceptual relationships, such as serials that were preceded or followed by other serials.¹⁰ Rules for cataloging continued to evolve, and many rules documents emerged, from *Catalog Rules, Title and Author Entries*¹¹ in 1908 and its subsequent revisions, to *Rules for Descriptive Cataloging in the Library of Congress*¹² in 1949. These sets of rules struggled to address issues regarding selection, form, order, and arrangement of values for the data elements. However, none of these rules questioned the fundamental principle that bibliographic records were the appropriate tool for recording these values. Even the shift in the 1950s and 1960s toward cataloging based on underlying principles and objectives^{13,14} assumed the concept of the bibliographic record.

As new technological possibilities arose in the 1960s, the Library of Congress became interested in the possibility of “automating” library work, such as the printing of catalog cards. Studies demonstrated how bibliographic data from catalog cards could be input “by punching on a tape punching typewriter,” making the data “machine-interpretable”—meaning that each field of data (title, author, etc.) would be parsed separately by the machine.¹⁵ To assist with data entry, a “workform” was used to collect bibliographic data about an item.¹⁶

L.C. Class No.	L111.052A1963
Main Entry	U.S. President, 1961-1963 (Kennedy)
Title Title Page Transcription Edition Statement	Program for education. Message relative to a proposed program for education, and a draft bill to strengthen and improve educational quality and educational opportunities in the Nation.
Imprint	Washington, U.S. Govt. Print. Off., 1963
Collation	68 p. 24 cm.
Series Statement	(68th Cong. 1st sess. House of Representatives. Document no. 54)
Notes	Caption title.
Tracings	1. Education-U.S.-1945 1. Title. (Series: U.S. 68th Cong., 1st sess., 1963. House. Document no. 54)
L.C. Class No.	370.973
Dewey No.	63-60561
L.C. Card No.	

Figure 1: Sample of one of the earliest workforms used to record bibliographic data. From Lawrence F. Buckland, *The Recording of Library of Congress Bibliographical Data in Machine Form*. Washington D.C.: Council on Library Resources, Inc., 1965, figure 3.

The workform, like the catalog card, was an entity of its own, with boundaries, space limitations, and prescribed data elements, thus reinforcing the conception of the bibliographic record as a *thing*. Additionally, the computerized collection of data elements created for Machine Readable Cataloging (MARC) became known as a “MARC record,” a term still in use today. Physically, a

MARC record is formless—as a compendium of electronic bytes, it has no tangible physical form. A raw MARC record has no shape at all, it simply consists of a string of unformatted characters (see Figure 2). Structure is imposed through standardized character positions (such as in the leader), delimiters and punctuation. Character placement is also used in the directory, which defines the structure of the variable fields.¹⁷

```
02617cam 22004931a
4500001000800000005001700008008004100025035002300066035001400089040
0017001030430012001200900021001321000042001532450336001952460021005
3124600630055226000350061530000440065059001610069459001990085550000
1901054500014201073500005101215500003901266500013501305500014601440
6510043015866510036016296550047016656550063017126920028017756920025
0180369200290182870000500185770000320190770000300193970000340196970
0003402003700001902037700002202056927000902078928001302087948002302
10011974352003022713 0037.0940923s1840 enkabcf 00 0 eng u
a(UPRA)CTYXRL7078-B 9CAF1680YL cUPRAdCtY-BR an-us--- a\Za W679\
b+840s1 aWillis, Nathaniel Parker,d1806-1867.10aAmerican scenery,
or, Land, lake, and river illustrations of transatlantic nature :-
buniform with Dr. Beattie's Switzerland, Scotland, & Waldenses /-
cfrom drawings by W.H. Bartlett, engraved in the first style of the
art, by R. Wallis, J. Cousen, Willmore, Brandard, Adlard,
Richardson, & c ; the literary department by N.P. Willis.30aAmerican
scenery30aLand, lake and river illustrations of transatlantic
nature aLondon :bGeorge Virtue,c1840 a30 parts :bills., map, port.
;c29 cm. aBEIN Za W679 +840S: Original wrappers. On parts 1-15,
price ([50] cents) has been stamped on. Without subscriber list in
part 1. Bookplate of Owen F. Aldis. aBEIN 2003 +56: Original
wrappers. Variant wrapper for part 1: advertising matter on p. [2-
3]. Plates are hand-colored, probably in the 20th century. Stamp:
Carl Gustaf. Bookplate of Paul Mellon. aWrapper title. a"The first
edition of this work is printed expressly for Martin and Co., 42,
Whitehall Street, New York, and at Philadelphia and Boston." -
a"Subscriber's copy. Sold to subscribers only." aPrinted subscriber
list in part 1. aAdded engraved title pages in parts 1 and 30:
American scenery by N.P. Willis ; illustrated in a series of views
by W.H. Bartlett. aFor possible binding into two volumes, title
pages (both dated 1840) and lists of engravings for vol. 1 and vol.
2 included at end of pt. 30. 0aUnited StatesxDescription and
travel. 0aUnited StatesvPictorial works. 7aHand-colored
illustrations.2local5CtY-BR 7aPublishers' advertisementszEngland-
zLondony1840.2rbgenr14aAldis, OwenxBookplate.14aGustaf, CarlxStamp.
14aMellon, PaulxBookplate.1 aBartlett, W. H.q(William Henry),d1809-
1854.1 aWallis, Robert,d1794-1878.1 aCousen, John,d1804-1880.1 -
aWillmore, Arthur,d1814-1888.1 aBrandard, Robert,d1805-1862.1 -
aAdlard, Henry.1 aRichardson, G. K. arlin aAC031397 aRLA Beinecke
Recon□
```

Figure 2. Raw MARC data for the book *American Scenery, or, Land, Lake, and River Illustrations of Transatlantic Nature*. Data from <http://www.hray.com/5264/marc.htm>

While this raw data is intended for machine processing, the human interface continued to resemble the workform. As technology evolved, data entry shifted from physical paper workforms to direct electronic input. Yet even today these input interfaces still retain the shape and structure of the physical workform, reifying the idea that the bibliographic record is a collection of specific data elements within a bounded space.

The screenshot shows the OCLC Connexion interface for a record with OCLC number 1822785. The record is for 'Margaret Mitchell and her novel, Gone with the wind.' The interface includes a menu bar (File, Cataloging, Authorities, Edit, Action, View, Tools, Window, Help), a toolbar with various icons, and a main data entry area. The data entry area is divided into several sections: 'Books' (with a dropdown menu), 'Rec stat' (with a dropdown menu), 'Entered' (19860304), and 'Replaced' (20020731). Below these are fields for 'Type', 'BLvl', 'Desc', 'Form', 'Cont', 'GPub', 'Fast', 'Snce', 'Conf', 'Biog', 'LitF', 'Dist', 'Ctrl', 'MRec', 'Lang', 'Ctry', 'Index', and 'Dates'. A table of MARC fields is displayed below, with the first field (010) highlighted. The MARC fields include: 040 (DLC #c DLC #d CNF #d ZOY), 043 (n-us-ga), 050 (0 0 P83525 .I972 #b G686 1936), 082 (0 0 813/.52 #2 19), 049 (OCLC), 245 (0 0 Margaret Mitchell and her novel, Gone with the wind.), 260 (New York : #b Macmillan, #c 1936.), 300 (22 p. : #b ill., 2 ports. ; #c 19 cm.), 600 (1 0 Mitchell, Margaret, #d 1900-1949. #t Gone with the wind.), 651 (0 Georgia #x History #y Civil War, 1861-1865 #x Literature and the war.), and 650 (0 Women and literature #z Georgia #x History #y 20th century.). At the bottom, there are buttons for 'Delete Holdings-', 'Export-', 'Label-', 'Produce-', 'Replace-', 'Update Holdings-', 'Validate-', and 'My Status-'. The 'INS' button is also visible.

Figure 3: A contemporary digital workform in OCLC's Connexion. From <http://oclc.org/connexion/interface/client.en.html>

As technology continued to evolve, so too did the bibliographic record. The term “bibliographic record” first appears in textbooks in Wyner’s *Introduction to Cataloging and Classification*, 6th ed. (1980), where it is defined as “a transcription of the complete cataloging information for any item.”¹⁸ According to Wynar, the purpose of bibliographic records is to

1. provide all information necessary to describe an item accurately both physically and intellectually in order to distinguish it from every other item
2. to provide its location in the collection

Subsequent editions of the same textbook define bibliographic records as “catalog entr[ies] in a card, microtext, machine-readable or other form carrying full cataloging information for a given item in the library.”^{19, 20} The 2000 edition amends the definition to change “item in the library” to “information package [an instance of recorded information]”.²¹ (491, 499). Other well-known texts such as Chan’s *Cataloging and Classification: An Introduction*^{22, 23} and Downing’s *Introduction to Cataloging and Classification*²⁴ that have multiple editions in this time frame offer similar definitions. These definitions reveal the reinforcement of bibliographic record as “thing” by stipulating a form of existence, be it physical or digital. This notion of bibliographic record as “thing” is only reinforced by giving it physical shape through the use of catalog cards or conceptual shape through tools like the workforms used in the creation of machine-readable records. Just as these conceptual boundaries shape the input of bibliographic data, so too do they shape the output in the way that records are displayed [cite SHOT talk here?]. Even modern texts and instructional manuals define records as things with space and place, such as “the place where information about each item is written down, whether on a physical 3 x 5 card or in the online database.”²⁵ Definitions for all these textbook examples describe descriptive information about a resource in plural terms; that is, the multiple data elements collectively necessary to describe a resource.

Emerging definitions and statements of purpose for bibliographic records also often refer to “full,” “enough” and/or “accurate” bibliographic data as a requirement for bibliographic records. “retrieval of any item in a collection depends on a full and accurate description of the work”²⁶ But what is meant by “full” and “accurate”?

By the 1980s, “essential parts of a full catalogue entry for a modern book” (per *Anglo-American Cataloguing Rules*, 2nd edition 1.0B and *International Standard Bibliographic Description for Monographic Publications* included the following elements:²⁷

- title
- statement of responsibility
- edition
- imprint [i.e. place, publisher, year of publication]
- collation [physical description]
- series statement
- any necessary annotation [e.g., contents list, summary]
- ISBN

In addition to the data elements prescribed by accepted cataloging standards, interest arose in adding additional useful data to bibliographic records, especially MARC records. The inclusion of additional bibliographic data in records was actually discussed during the development of MARC: most library representatives favored the inclusion of as much data as possible to address unknown future uses. However, the decision of which data elements to include was ultimately left to the Library of Congress, since they were the major institution supporting the project in terms of both staff and funding.²⁸ Early MARC records included only the bibliographic data that was of use to the Library of Congress. As other libraries encountered the need for additional data elements, “enhanced” bibliographic records were developed. These enhancements added data elements based on the most popular suggestions from libraries, the most tested elements, and the most feasible additions, which mainly consisted of the following:²⁹

- increased number of subject headings
- inclusion of tables of contents

- inclusion of summaries and abstracts
- improved use of classification
- inclusion of back-of-book indexes
- including subject descriptors from other vocabularies in addition to Library of Congress Subject Headings (LCSH), including both specialty controlled vocabularies and natural language terms

At the same time, cooperative cataloging among institutions was also on the rise as the ease of sharing MARC records increased, likely due to new media formats for record distribution as well as network technologies. Enhanced records and other deviations meant that not all the shared records contained the same bibliographic data elements. In the 1990s, the cooperative cataloging community developed a set of standards for “core” records, “an agreed minimum set of bibliographic data for a catalogue record which every library participating in a cooperative programme commits to provide.”³⁰ IFLA stipulated that these minimal records included the following data elements:³¹

- title
- statement of responsibility
- edition
- place of publication
- name of publisher
- date of publication
- specific materials designation and extent
- title proper of series
- numbering within series
- standard number

Additional data elements were considered optional. AACR2r also stipulated varying levels of data elements needed for bibliographic records, from Level 1 (which omits information like subtitles, series titles, illustrations, size, and additional authors); Level 2 (which requires the inclusion of subtitles, series titles, statements of responsibility and additional physical details); to Level 3, which are “by definition, the most thorough work possible,” requiring the inclusion of every applicable detail mandated by the rules.³² Yet “every applicable detail” is still limited by the extent of the AACR2r rules, which address only the eight traditional areas of bibliographic description. A traditional bibliographic record, then, is a unit of a limited set of multiple data elements embedded in a physical or digital space meant to stand in for a physical or digital resource. These characteristics shape the current mental model of how librarians conceptualize bibliographic data.

The turn of the 21st century

Despite the long service of the bibliographic record as conceptualized in the 20th century, as user demands and digital technologies continued to advance, librarians attempted to keep pace. Years of analysis of relationships between resources³³ led to the abstract model *Functional Requirements for Bibliographic Records* (FRBR)³⁴. While certainly an advance over traditional cataloging, even the name itself reveals the insidious unconscious attachment to the bibliographic record as the unit of description for a resource. Therefore, any new cataloging codes and standards based on FRBR will continue to be based on the idea of records. *Resource Description and Access* (RDA), for example, relies on linking of records for works, manifestations, expressions, and items.³⁵ This means that catalogers must create separate records for each work, each manifestation of that work, each expression of each manifestation, and possibly each item.³⁶ These records are then related to one another using FRBR relationships. Relationships therefore exist at the record level, relating a collection of data

elements to another collection of data elements, rather than at the level of the data elements themselves.

New conceptual models

The traditional library approach to describing bibliographic data in the format of a bibliographic record—a formalized unit containing multiple data elements—is clearly embedded in the psyche of most librarians as well as in their practices and technologies. While such conceptualization proved a valuable tool for American libraries in the 20th century, new conceptual models of data at large from outside the library world may also prove useful, especially in the context of data sharing via the World Wide Web.

The idea of the Semantic Web is one such model that many agree can revolutionize librarianship. The purpose of the Semantic Web is to “bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users,” such as more precise and automated searching.³⁷ In addition to the benefits of more precise search and retrieval, Semantic Web technologies can also benefit libraries by making library data more openly accessible and interoperable with other library and non-library systems, and library resources more easily findable on the Web at large.³⁸ The Semantic Web is made possible through “Linked Data”—a method of publishing data—rather than documents—in a structured, machine-readable way, thus enabling connection (“linking”) of data among various sources for improved information retrieval and experiences.³⁹ Unlike the traditional model of bibliographic records, which associates multiple descriptive properties with one resource, the semantic web model uses multiple statements, each one individually associating a descriptive property with a resource.⁴⁰ These statements follow standards for publishing data on the web that hinge on two major conceptual foundations: 1)

Uniform Resource Identifiers (URIs) and 2) semantic web standards, such as the Resource Description Framework (RDF).⁴¹ These statements, known as “triples,” are formulated as assertions that include a subject, a predicate, and an object.⁴² For example, the author of a work may be conceptualized as follows:

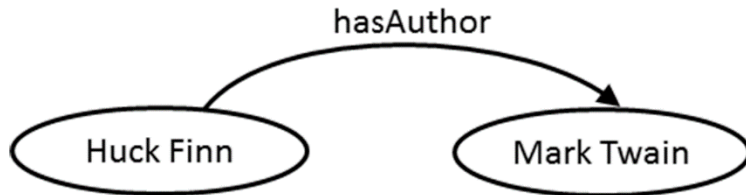


Figure 4: Visualization of an RDF triple statement declaring Mark Twain to be the author of *The Adventures of Huckleberry Finn*. Illustration by the author.

For the model to work successfully, however, subjects and predicates must be represented by Uniform Resource Identifiers (URIs)—unique character strings used to “distinguish one resource from all other resources.”⁴³ Every resource, physical or digital, concrete or abstract, is assigned a standardized character string that uniquely identifies that resource among other resources. A URI may also specify location of a resource, especially in the case of digital resources.⁴⁴ A URI is a *single* data element that uniquely describes (and more often than not, locates) a resource.

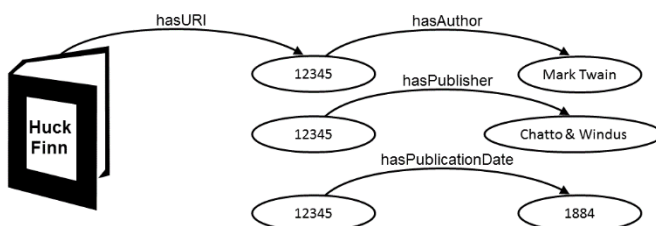


Figure 5. Visualization of RDF triple statements resolving to a single URI. Illustration by the author.

Each triple describes one, and only one, data element. Any amount of triples may be used to describe a resource. These triple statements may be asserted by the same source (such as a single library or institution), or they may be asserted by other institutions and organizations across the Web (such as VIAF or Wikipedia, for example). The major point is that individual triple statements about a single resource need not be collocated in the same physical or intellectual space.

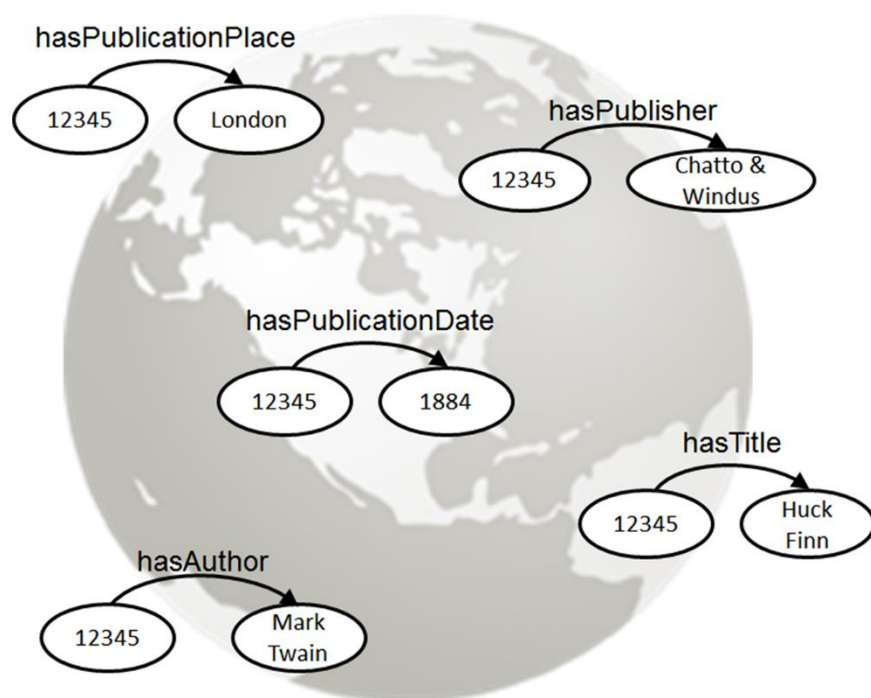


Figure 6: Visualization illustrating diverse and disparate sources of triplets. Illustration by the author.

Additionally, with the conceptual model of Linked Data, it is no longer necessary to record a complete “set” of data elements to sufficiently describe a work. Unlike a bibliographic record,

which as we have seen requires a collection of multiple data elements to describe a specific resource, the model of the semantic web relies on the idea that a single data element—the URI—uniquely identifies a resource. In the Linked Data world, bibliographic data need not be a conceptual whole. There is no need for a minimum (or maximum) number of data elements to describe a resource. In fact, the collective unit model espoused by the traditional bibliographic record directly contradicts the underlying concept of Linked Data and the Semantic Web. The purpose of the RDF model is to allow data assertions to be parsed independently of one another. While a traditional bibliographic record requires all mandatory elements to be present for that record to exist, a resource description in the Semantic Web model comes into existence with the very first URI triple. Additional descriptive triple statements may be created at the same time, or they may be created and added over time. And unlike cards or workforms that pose limitations on physical space and therefore on the amount of descriptive data, triple statements are unlimited. Libraries and other institutions may choose to only display a selected set of triple data, and different institutions may choose to display the data elements that are most useful and desirable to their unique users. It is the underlying concept of separate, individually stated data elements that allows for the accessibility and interoperability so desirable to libraries. Yet based on the review of bibliographic records presented earlier in this paper, libraries and librarians still seem entrenched in conceptualizing resource descriptions as cohesive collections of data and educating future librarians to do the same.

Implications

Throughout the 20th century, catalogers and cataloger scholars supported the notion that library cataloging is independent of technology, and that the principles of cataloging exist independently of technological medium.⁴⁵ The preceding analysis of the history and evolution of bibliographic records and the ways in which they are conceptualized demonstrates that

cataloging rules and principles were, and continue to be, closely tied to the technologies of the eras of their development. The catalog card in particular had a long-standing effect on which and how much bibliographic data could be recorded about a resource, concepts which continued to be enforced in cataloging even as new technologies (such as computing) arose that did not afford the same limitations.

This historical analysis of the discourse around bibliographic records reveals that the need for its existence in its current form has been taken for granted and gone unquestioned for more than 100 years. Early analysis of cataloging rules in the 1950s found them to be based on “a long accumulation of essentially ad hoc decisions” rather than any stipulated purposes or objectives.⁴⁶ Later rules, such as ISBD, AACR, and even RDA purport to support cataloging objectives: finding works by title, author, or subject, identifying them, selecting them from among other works, and obtaining them.⁴⁷ Yet the development of machine-readable cataloging data—which had a significant impact on the conceptualization of the bibliographic record—was originally developed not with cataloging objectives in mind but rather goals of improved internal efficiency.⁴⁸ Additionally, despite the claim of user service, it seems clear from years of user studies that the traditional cataloging objectives are not as well aligned with user needs as perhaps was once thought. Users do not use library catalogs in the same way librarians assume or expect they should. Users seem to have a different mental model of how bibliographic data and links between these data are structured. Why shouldn't the underlying structure of library catalogs evolve to reflect how users model and conceptualize information? Such a new model would be based on 1) how users conceptualize information relationships (rather than how librarians or existing bibliographic records conceptualize them) and 2) the influences of current technology, both its effect on users' conceptualizations as well as its new affordances.

Traditional bibliographic records were also conceptualized as surrogates to physical resources. Yet few contemporary libraries restrict their collections to solely physical holdings. Even as

librarians' resource collections have evolved, the definition of the bibliographic record remained the same, save for small wording changes like "book" to "information package." While this reflects the shift to describing the intellectual content of a resource separate from the physical carrier that contains it, the assumption remains that a bibliographic record comprised of collective data elements is the appropriate tool for description. Despite the acknowledgement of significant resource evolution, little to no consideration was even considered for underlying structural methods of bibliographic description. Which data elements were necessary for inclusion and how to record their values underwent constant renegotiation, yet the fundamental carrier for this data remained unquestionably accepted.

With increasingly self-describing digital resources, there is less need for surrogate bibliographic records to provide description and location information. Digital documents can represent themselves with URIs and linked triple statements that specify locations. Even physical, non-dereferencable resources that need pointers to physical locations (like physical books, archival materials, or even people) can be located so long as a triple statement asserts the location information (such as a call number, box number, or address, respectively). In a traditional bibliographic record manifested via a catalog card or MARC record, this type of location data could not be conceptually divested and used separately from the rest of the included bibliographic data. But triple statements do not need to be carried alongside other data elements. So long as a URI is assigned to the resource, any bibliographic data or other descriptive information that references that URI is available as an assertion. It need not be included in the same document, the same database, even the same library or the same country.

The conceptualization of bibliographic data under a Linked Data model would exist as individual statements rather than the traditional mentality of collections of descriptive data elements. As sets of data, traditional bibliographic records required a minimum number of elements or mandatory amount of data to be considered full and accurate. Bibliographic records offered

stable boundaries indicating which elements were included and which were excluded. With individual triple-structure assertions, this notion becomes moot. While actual implementation may vary, with a new conceptual model of cataloging based on Linked Data, catalogers would no longer think in terms of records at all—rather, they would conceive of data as individual triple statement assertions. The conceptual work of catalogers then begins to shift. Instead of seeing a bibliographic record as some sort of “checklist” as to which data can be included and which data have no fit, catalogers could be free to include any and all triple statements deemed useful to describe a resource. Catalogers may continue to be guided by traditional cataloging objectives: i.e., if one objective is to find works by author, then author is clearly a necessary triple statement (or multiple triples in the case of multiple authors). Or catalogers may choose to reexamine traditional cataloging objectives in light of the contemporary information environment and the needs of their local users. Catalogers might make much more localized and specific triple statement assertions based on local needs. Because the Linked Data model supports tracing data via URIs, patrons can follow triple statements to find additional useful data.

However, catalogers can only begin to work in this way if they are not held bound by the traditional definitions and conceptualizations of bibliographic records. Many catalogers are certainly interested in the Semantic Web, evidenced by professional interest groups⁴⁹ and even national cataloging committees.⁵⁰ However, it is almost always more difficult to break out of an existing conceptualization than to start fresh. For this reason, we need to strongly examine cataloging education and how we teach future catalogers about bibliographic description and data. The use of traditional definitions and descriptions of bibliographic records—especially continuing to use the catalog card as an illustrative analogy—is problematic in shaping the perspective of future catalogers. If we continue to educate catalogers using the traditional model of the bibliographic record, we risk remaining entrenched in this model ten, twenty, fifty years down the line, keeping libraries and library catalogs on the trailing edge of technology, perhaps

even contributing to libraries' obsolescence. Instead, let us educate the future members of our profession to be open-minded to new possibilities, both technological and conceptual, and by doing so, perhaps it will be libraries once again developing the next generation of technological innovation, whatever is to come next after the Semantic Web.

Conclusion

The idea of the bibliographic record—a collection of descriptive information that stood as a surrogate for a physical resource—enabled significant advances in American library cataloging. Yet as other components of libraries changed, such as resource types and technological tools, the model of the bibliographic record remained static. Types of data elements and the source and format of their values were often highly disputed, yet the fundamental underlying structure of the bibliographic record remained unquestioned for nearly 100 years.

However, the evolution of digital resources and technologies has spurred new ways of conceptualizing data. The Semantic Web's Linked Data model, with its individual descriptive statements, triggers an opportunity for the re-conceptualization of bibliographic records away from conceptual collective units that include all the bibliographic information about a resource together in one place, like a catalog card or a MARC record. Linked Data allows for individual data elements to be collected in and linked from many multiple locations across the web, opening up new opportunities for the ways library cataloging is conceptualized and practiced.

Libraries have already shown interest in Linked Data as a means of improving catalogs for both patrons and librarians. But if libraries and librarians continue to assume the bibliographic record as the basic unit of the library catalog, and continue to teach that model to future librarians, they risk perpetuating and reifying it. Linked Data clearly holds much more potential when used to link data instead of records. However, so long as the concept of the record as a

collective unit holds sway, library bibliographic data remains conceptually incompatible with these new models, and so long as cataloging continues to be based on records rather than data, library catalogs will not change.

References

- ¹ Joan M. Reitz, "Bibliographic record." In *Online Dictionary of Library and Information Science* (Libraries Unlimited, 2013). Retrieved November 12, 2013 from http://www.abc-clio.com/ODLIS/odlis_b.aspx#bibrecord
- ² Ronald Hagler, *The Bibliographic Record and Information Technology*, 3rd ed. (Chicago: American Library Association, 1997), 15.
- ³ Jim Ranz, *The Printed Book Catalogue in American Libraries, 1723-1900*. (Chicago, American Library Assn, 1964), 33.
- ⁴ Ranz, *Printed Book Catalogue*, 11-12.
- ⁵ Ranz, *Printed Book Catalogue*, 7-8.
- ⁶ Ranz, *Printed Book Catalogue*, 25-26.
- ⁷ Barbara B. Tillett, "Bibliographic Structures: The Evolution of Catalog Entries, References, and Tracings." In *The Conceptual Foundations of Descriptive Cataloging*, 149–166. (San Diego, California: Academic Press, Inc., 1989), 150.
- ⁸ British Museum, "Rules for the Compilation of the Catalogue." In *Foundations of Cataloging: A Sourcebook*, 1-14. (Littleton, Colorado: Libraries Unlimited, 1985).
- ⁹ Charles A. Cutter, *Rules for a Dictionary Catalog*. (Washington, D.C.: Government Printing Office, 1891).
- ¹⁰ Tillett, "Bibliographic Structures," 156.
- ¹¹ American Library Association, *Catalog Rules, Author and Title Entries*. (Chicago: American Library Association, 1908).
- ¹² Library of Congress Descriptive Cataloging Division, *Rules for Descriptive Cataloging in the Library of Congress*. (Washington, D.C. : Library of Congress, 1949).
- ¹³ Seymour Lubetzky, *Cataloging Rules and Principles; a Critique of the A.L.A. Rules for Entry and a Proposed Design for their Revision*. (Washington D.C.: Library of Congress, 1953).

- ¹⁴ International Conference on Cataloguing Principles, *Statement of Principles Adopted at the International Conference on Cataloguing Principles, Paris, October 1961*. (London : IFLA Committee on Cataloguing, 1971).
- ¹⁵ Lawrence F. Buckland, *The Recording of Library of Congress Bibliographical Data in Machine Form*. (Washington D.C.: Council on Library Resources, Inc., 1965), 2.
- ¹⁶ Buckland. *Bibliographical Data in Machine Form*. v.
- ¹⁷ Library of Congress Network Development and MARC Standards Office. *MARC 21 Specifications for Record Structure, Character Sets, and Exchange Media*. (Washington, D.C., Library of Congress, 2000). Retrieved on August 15, 2014, from <http://www.loc.gov/marc/specifications/specrecstruc.html#direct>
- ¹⁸ Bohdan S. Wynar, *Introduction to Cataloging and Classification*. 6th ed. (Littleton, Colo: Libraries Unlimited, 1980), 2.
- ¹⁹ Arlene G. Taylor, *Introduction to Cataloging and Classification*. 7th ed. (Littleton, Colo: Libraries Unlimited, 1985), 602.
- ²⁰ Arlene G. Taylor, *Introduction to Cataloging and Classification*. 8th ed. (Englewood, Colo: Libraries Unlimited, 1992).
- ²¹ Arlene G. Taylor, *Wynar's Introduction to Cataloging and Classification*. (Englewood, Colo: Libraries Unlimited, 2000), 491, 499.
- ²² Lois Mai Chan, *Cataloging and Classification : An Introduction*. (New York: McGraw-Hill, 1981).
- ²³ Lois Mai Chan, *Cataloging and Classification : An Introduction*. 2nd ed. (New York: McGraw-Hill, 1994).
- ²⁴ Mildred Harlow Downing, *Introduction to Cataloging and Classification*. 6th ed. (Jefferson, NC: McFarland, 1992).
- ²⁵ Allison G. Kaplan, *Crash Course in Cataloging for Non-Catalogers : A Casual Conversation on Organizing Information*. (Westport, Conn: Libraries Unlimited, 2009), 113.
- ²⁶ Downing, *Introduction*, 8.
- ²⁷ Mary Piggott, *A Topography of Cataloguing : Showing the Most Important Landmarks, Communications and Perilous Places*. (London: Library Association, 1988), 73-74.

- ²⁸ Buckland, *Bibliographical Data in Machine Form*, vii-ix.
- ²⁹ Stefanie A. Wittenback, "Building a Better Mousetrap: Enhanced Cataloging and Access for the Online Catalog." In *Advances in Online Public Access Catalogs*, 1: 74–92. (Westport, Conn: Meckler Publishing, 1992), 74-81.
- ³⁰ Pat Oddy, *Future Libraries, Future Catalogues*. (London: Library Association Pub, 1997), 92.
- ³¹ Jay H. Lambrecht. *Minimal Level Cataloging by National Bibliographic Agencies*. (München ; New York: Saur, 1992).
- ³² Sheila S. Intner, *Standard Cataloging for School and Public Libraries*. 4th ed. (Westport, Conn: Libraries Unlimited, 2007), 9.
- ³³ Barbara A. Tillett, "Bibliographic Relationships: Toward a Conceptual Structure of Bibliographic Information Used in Cataloging." (PhD. diss., University of California, Los Angeles, 1987).
- ³⁴ IFLA Study Group on the Functional Requirements for Bibliographic Records, *Functional Requirements for Bibliographic Records : Final Report*. (München: KGSaur, 1998).
- ³⁵ Robert L. Bothmann, "Evolving Standards: Making the Jump to RDA in Historical Context." In *Practical Strategies for Cataloging Departments*, 1–18. (Santa Barbara, Calif: Libraries Unlimited, 2011), 10.
- ³⁶ Bothmann, "Evolving Standards," 7.
- ³⁷ Tim Berners-Lee, James Hendler, and Ora Lassila, "The Semantic Web." *Scientific American*, May 2001.
- ³⁸ Emmanuelle Bermes, "Enabling Your Catalogue for the Semantic Web." In *Catalogue 2.0: The Future of the Library Catalogue*, 117–142. (Chicago: Neal-Schulman, 2013), 118.
- ³⁹ Christian Bizer, Tom Health, and Tim Berners-Lee, "Linked Data—The Story So Far." *International Journal on Semantic Web and Information Systems* 5 (3): 1-22.
- ⁴⁰ Eric J. Miller, "An Introduction to the Resource Description Framework." *Journal of Library Administration* 34 (3-4) (December 2001): 245–255, 248.
- ⁴¹ Tim Berners-Lee, "Linked Data – Design Issues." Retrieved August 15, 2014 from <http://www.w3.org/DesignIssues/LinkedData.html>

⁴² Bermes, "Enabling," 123.

⁴³ Tim Berners-Lee, Roy Fielding and Larry Masinter, "Uniform Resource Identifier (URI): Generic Syntax. Request for Comments 3986. Retrieved on November 20, 2013 from <http://tools.ietf.org/html/rfc3986>. Section 1.1

⁴⁴ Ibid, Section 1.1.3.

⁴⁵ S. Malinconico. 1979. *The Future of the Catalog : The Library's Choices*. White Plains, NY: Knowledge Industry Publications.

⁴⁶ Wyllis Wright, "General Philosophy and structure of the code.: In *Institute on Cataloging Code Revision Working Papers*. (Stanford, Calif., 1958), 1.

⁴⁷ International Federation of Library Associations and Institutions, Statement of International Cataloging Principles. (The Hague, Netherlands, IFLA, 2009). Retrieved on January 20, 2014 from http://www.ifla.org/files/assets/cataloguing/icp/icp_2009-en.pdf

⁴⁸ Christine L. Borgman, "From Acting Locally to Thinking Globally: A Brief History of Library Automation." *The Library Quarterly*: 215–49 (1997).

⁴⁹ Such as <http://lodlam.net/>

⁵⁰ Library of Congress, Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services. (Washington D.C.: Library of Congress, 2012). Retrieved on January 20, 2014, from <http://www.loc.gov/bibframe/pdf/marclld-report-11-21-2012.pdf>