

# ENTITIES, ATTRIBUTES, AND BIBLIOGRAPHIC RELATIONSHIPS

RE-READING BARBARA B. TILLET'S  
PHD DISSERTATION THIRTY YEARS AFTER

**Lapo Ghiringhelli and Mauro Guerrini**

*Dipartimento di Storia, Archeologia, Geografia,  
Arte e Spettacolo (SAGAS), Università degli  
Studi Firenze, Firenze, Italia*

## **KEYWORDS:**

*bibliographic relationships,  
cataloguing, Barbara B. Tillett*

## **ABSTRACT**

*A bibliographic relationship is the logical connection that is created between two or more bibliographic entities. These connections are an instrument of primary importance to assist the user in the context of the catalogue. This essay intends to illustrate the PhD research study carried out by Barbara B. Tillett on bibliographic relationships from 1981-1987, the first exhaustive theoretical analysis on the bibliographic universe whose consequences were at the basis of those changes that have made a great impact on cataloguing and bibliographic knowledge organization over the last thirty years. Entities, attributes, and bibliographic relationships have played a central role in the creation of the conceptual model known as **Functional Requirements for Bibliographic Records (FRBR)** and continue to be essential in the design of standards, such as **RDA, Resource Description and Access**, and **IFLA LRM, IFLA Library Reference Model**. The study of entities, attributes, and bibliographic relationships is an important and active field that is greatly indebted to Tillett's work.*

## The PhD Thesis, *Bibliographic relationships: toward a conceptual structure of bibliographic information used in cataloguing*

In 1987, upon publication of Barbara B. Tillett's PhD dissertation *Bibliographic relationships: toward a conceptual structure of bibliographic information used in cataloguing*, the study of bibliographic relationships had not been fully explored by researchers of Library and Information Science. At that time, the most significant theoretical results were both the definitions of relationships in the UNIMARC: Universal MARC Format (1980) and the 1981 analysis of hierarchical relationships proposed by Goossens and Mazur-Rzesos in the article *Hierarchical Relationships in Bibliographic Descriptions: Problem Analysis*.

On the one hand, Goossens and Mazur-Rzesos introduced a schematic representation for hierarchical relationships to express complex tree structure in order to provide the theoretical basis able to express such relationships for manual and computerized solutions. This structure, in examples, was limited to three levels called *set level* for the highest level, *subset level* for the intermediate, and *piece level* for the lowest. Each level used special notation. This study represented, for Tillett, the only in-depth analysis of bibliographic relationships. While recognizing that the schema seemed to work for true hierarchies, such as series and subseries, Tillett found the schema less effective when expressing other types of material.

On the other hand, the definitions of bibliographic relationships found in the UNIMARC Format suggested a categorization where three types of relationship were recognized: *vertical*, or the hierarchical relationship of the whole and its parts, and the parts to a whole; *horizontal*, or the relationship between versions of a work; *chronological*, or the relationship in time between issues. Tillett observed that the three UNIMARC categories did not reflect the taxonomic principles for which each category must be mutually exclusive and totally exhaustive. They were not exhaustive because they omitted some relationships, such as those that exist between copies, between a supplement and the work it accompanies, or between a book review and the book under review. In the same way, they did not respect the requirements of exclusivity, because the distinction between *horizontal* and *chronological* was vague. This is the case with handbooks issued annually in updated versions with slight title changes. Here the relationship between ensuing editions is both horizontal, as it appears in various versions, and chronological, as it is issued annually with successively different titles. The definitions developed in UNIMARC represented an important step forward in the process of describing bibliographic relationships, but were incomplete. Therefore, Tillett

considered a new categorization necessary.

From these theoretical assumptions, Barbara Tillett's pioneering work took off. Her work is divided in two parts: an analytical study and an empirical study.

## The analytical study

In the first part, the analytical study, Tillett, through an examination of the most relevant Anglo-American cataloguing codes, created a taxonomy of bibliographic relationships and outlined a history of linking devices used in catalogues. The empirical study was conducted in order to evaluate the extent of bibliographic relationships as reflected through their frequency of occurrence in the machine-readable database (MARC) of the Library of Congress during the period between 1968 and July 1986. Tillett's intent, as stated in the introductory pages of her Ph.D. thesis, was to provide the groundwork for understanding the conceptual structure of the library catalogue in terms of bibliographic relationships.

Tillett chose twenty-four codes, selected on the basis of their well-recognized influence on cataloguing at both the Library of Congress and major libraries in the United States: from the Antonio Panizzi's 91 rules of 1841 to various editions of *Rules for a Dictionary Catalogue* by Charles A. Cutter, to American Library Association's codes, to the *Paris Principles* of 1961, up to the Anglo-American Cataloguing Rules, 2nd edition of 1978. From the review of the twenty-four cataloguing codes consulted, Tillett developed a taxonomy in which seven categories of bibliographic relationships were identified, whereby the requirements of total exhaustivity and mutual exclusivity are satisfied; at the same time, the essence of the three UNIMARC categories is kept. The seven categories identified are: a) equivalence relationships, b) derivative relationships, c) descriptive relationships, d) whole-part (or part-whole) relationships, e) accompanying relationships, f) sequential relationships, and g) shared characteristic relationships.

### *a) Equivalence relationships*

Equivalence relationships are those that hold between exact copies of the same manifestation of a work or between an original work and reproductions of it. The types of resources participating in this relationship are copies, issues and reissues, facsimiles and reprints, photocopies, microforms and other similar reproductions. The cataloguing rules suggested different methods to indicate

equivalent works in bibliographic records using linking devices as dash entries<sup>1</sup> for the equivalent work on the record for the original work; notes to link separate records for the original and related works; notes on the record for the original work acknowledging the equivalent work; as well as notes on the record for the equivalent work, acknowledging the original with the same uniform title heading used in the records for both the equivalent work and the original. An equivalence relationship presupposes that the content and intellectual or artistic authorship are preserved. Should an alteration occur, the relationship is no longer equivalent, but derivative.

In an updated article about bibliographic relationships, Tillett recognizes the trickiness of the idea of equivalence. This fact is due to the subjectivity which leads every user to consider each distinction that can show up between two entities more or less important. Thus, equivalence relationships, Tillett affirms, are “subjective and can be viewed at various levels of the bibliographic entity hierarchy, but its determination is typically up to the user” (Tillett, 2001, p. 24).

### ***b) Derivative relationships***

Derivative relationships, called horizontal relationships in the UNIMARC definition, are those between a work and a modification based on that work. One work is derived from another when it enlarges, abridges, or modifies the entire work or portions of it. Derivative relationships cover a broad range of variations from an original work. They can occur between an original work and a variation of that work, such as editions, revisions, translations, summaries, abstracts, and digests; adaptations and modifications that result as new works; changes of genre that result as new works and new works based on the style or thematic content of the original work, such as free translations, paraphrases, imitations, and parodies. The connection among such diverse relationships, from those nearly identical to those distinctly separate, is due to the act that there is some original work from which another work is derived, whether in intellectual content, style, or thematic content. As a consequence, an equal diversity of linking devices is used to express these relationships in a catalogue, such as references, dash entries for added editions, editions statements, notes, uniform titles, cross references, subject headings, common main entries, and added entries. There is no rationale

1 dashed-on entry: A pre-AACR2 convention of indicating accompanying material and additional versions on the catalogue entry for the main item, a practice that economized on the number of catalogue cards needed for items sharing basic bibliographic description. This type of entry was eliminated in AACR2, reflecting a shift from card catalogues to MARC-based electronic catalogues. Synonymous with dash entry and dash on entry. (Reitz, 2004).

in the rules for preferring one particular device over another for linking a work and its derivations, and indeed, the rules are often inconsistent.

### **c) Descriptive relationships**

Descriptive relationships are between a work and a description, criticism, evaluation, or review of that work. Criticisms, casebooks, annotated editions, and commentaries are also included. When the description is predominant or issued separately, cataloguing rules use notes and added entries or subject entries to link the description with the work being described. On specific situations, particularly for manuscripts, incunabula, and maps, the descriptive work, such as a reference source, is noted in the bibliographic record for the work being described.

### **d) Whole-part relationships**

Whole-part (or, vice versa, part-whole) relationships correspond to vertical relationships in UNIMARC and hierarchical relationships by Goossens and Mazur-Rzesos. These relationships are between a component part of a work and its whole, for example, an individual selection taken from an anthology and the entire anthology, or between an individual chapter of a monograph and the monograph itself. The components may be part of some particular physical manifestation of a work. Tillett recognizes, in the taxonomy, three sub-relationships of the whole-part relationship: those referring to a physical unit are defined either as *containing relationships*, when involving the component parts of a physical unit, or *extractive relationships*, when involving the parts of a work that have been physically extracted and issued separately as individual selections. Finally, *abstract relationships* refer to the relationship between an abstract, intellectual whole and its intellectual components, that is, works that are not just contained in a physical unit but are parts of an intellectual work.

Cataloguing codes identify a wide variety of linking devices to represent whole-part relationships within bibliographic records, according to the type of sub-relationship that needs to be expressed. The devices mainly used are analytical entries, content notes listing specific parts, added entries for either the encompassing work or the part, multilevel descriptions, dash entries to incorporate all components within one bibliographic description, uniform title headings for a larger work as the main entry headings for the description of the part, and explanatory references identifying the parts of a work. Tillett acknowledges that

“these types of relationships are particularly interesting with electronic resources where images and text and/or sound become components of the whole and need to be addressed and brought together for displays, yet often are stored as separate components” (Tillett, 2001, p. 25).

### ***e) Accompanying relationships***

Accompanying relationships are between a work and the work it accompanies. This can occur between two works augmenting each other equally, and hence intended to be used as a unit, or between one work augmenting another principal or predominant work. In the most common situation, where there is a subordinate and a predominant work, the subordinate work may continue, supplement, illustrate, or, more generally, add in some other way to the usefulness of the principal work. In AACR2, rules for accompanying and supplementary materials are indicated in two separate sections, but the treatment for both is the same. In the taxonomy, however, Tillett includes the two types of materials in the same accompanying relationship as long as the supplement adds to the main work, but cannot continue it. Thus, once an element of continuation is introduced, the relationship that exists is a sequential relationship described below. The methods for indicating accompanying materials vary according to the type of resource. In case of material dependent on the principal work, in the bibliographic record for the main work, linking devices can be used as follows: a description of supplementary material, or a note, when the material is a minor work, a multilevel description, and a dash entry on the record for the main work, as in the case of indexes. In case of material independent of the principal work it augments, cataloguing rules suggest a separate record for the supplement with a note about the principal work and, in some cases, an added entry.

### ***f) Sequential relationships***

Sequential relationships, called chronological relationships in UNIMARC, hold between works that continue or precede one another. Tillett claims that the term *sequential* seems preferable to *chronological* in order to emphasize the fact that a set of works is sequential in nature, not that they can be arranged in chronological order. The most common works exhibiting sequential relationships are series, serials, and sequels. As for serials (certainly the predominant works exhibiting this relationship), Tillett points out how cataloguing codes have

had to deal with the matter, not unusual and definitely not marginal, of how best to record title changes. To represent sequential relationships between works with title changes, rules identify different devices, such as notes of all earlier titles or all later titles, and especially notes of immediately preceding or succeeding titles, uniform titles, and assembling added entries.

### ***g) Shared characteristic relationships***

The shared characteristic relationship is between works that are not related by any of the elements described in the other types of relationship but coincidentally share a common element. Such relationship can be established for whatever element is considered useful such as, the same language, publication date, country of publication, subject, format, or other characteristic used as an access point in a catalogue. As it is a more general relationship, it was not included in the empirical study carried out by Tillett. However, because it is the most pervasive of all relationships, Tillett recognizes its usefulness in the context of online catalogues, for the help it can provide to users in retrieving information, limiting, or filtering the search results.

## **The empirical study**

The empirical study was conducted as a supplement to the analytical study. The study was an attempt to examine the extent of bibliographic relationships as reflected in their frequency of occurrence in the machine-readable database of the Library of Congress during the period from 1968, when the Library's machine-readable records were first created, to July 1986, when Tillett conducted her research. The number of records in the Library's database at that time was 2,854,252. The percentage of records for each bibliographic format was diverse: ranging from 81.6% of records for books to 0.8% of records for music. The examined records included nearly all languages even though only English language monographs were included in the database until 1976-1977. Similarly, nearly all types of works were included, such as books, serials, audio-visual materials, music, and maps. Only manuscripts and machine-readable data files records were excluded, accounting for less than 0.4% and 0.08%, respectively.

Frequency of occurrence was calculated by counting the incidences of specific codes in the MARC format representing each type of relationship. Some

relationships were explicitly identified by a unique MARC code, with a one-to-one correspondence, making the counting straightforward. On the other hand, some relationships were expressed by general codes with a one-to-many correspondence, encompassing more than one type of relationship. The only general code selected for the empirical study was the '500' which tags a general note. About half of the Library's database records had '500' general notes including both relationship and non-relationship information; determining the presence of information regarding any relationship would have involved scanning each record in order to find words or phrases associated with each bibliographic relationship. Given the prohibitive cost of such an operation, the study was conducted on a sample of records with the '500' general note. Hence, the author's choice was to separate the empirical study into two parts: a study of those fields in which relationships are explicitly coded and a study of those relationships embedded in the '500' tag of general notes.

Regarding the study of explicitly coded relationships, the analysis of tags, indicators, and subfield codes of MARC format resulted in a list of 134 specifically coded bibliographic relationships, each of which was assigned to an appropriate taxonomic category from the analytical study: equivalence, derivative, descriptive, whole-part, accompanying, sequential, and shared characteristic. For the category of descriptive relationships there were no specific codes other than the '500' tag of general note. Furthermore, the category of shared characteristic relationships was dropped out from the study. Its general and pervasive nature would have made its calculation excessively complex, due to the fact that nearly every MARC code could have indicated such relationships. 14 of the 134 coded relationships were not presented in any bibliographic records. The remaining 120 relationships were translated into computer queries at the Library of Congress and analysed in terms of five factors:

- *bibliographic format*, or each MARC format for books, serials, maps, music, and audio-visual material;
- *subject*, determined from the call number field, tag '050', and grouped into four broad areas based on the Library of Congress classification: *Sciences*, *Social sciences*, *Humanities*, and *General*, to which was added a fifth group called *Unknown*;
- *publication date*, grouped by decade for the 20th century, by century for the 19th and 18th centuries, and another group for all pre-1700 dates;
- *language*, where English, French, German, Italian, Portuguese, Russian,

and Spanish records were separately analyzed, with all records in other languages as a single group, and

- *country of publication*, grouping each country in 20 categories.

The five factors were studied hypothesizing that they could reveal the presence of a specific bibliographic relationship, reflecting distinctive patterns for each type of relationship. The existence of specific patterns referring to each bibliographic relationship would have been useful in the context of cataloguing rules and catalogue system's design.

The second part of the study was conducted on a statistically significant random sample of 1841 records of the Library of Congress database, in which 1141 had the '500' general note. Since general notes contain both relationship and non-relationship information, it was necessary to scan each note with the help of a list of keywords in order to segregate those notes expressing relationships. The result was a list of 798 records containing relationship information in the '500' note. The study indicated that every category of bibliographic relationship, except the shared characteristic relationship, was represented in the general notes.<sup>2</sup> In spite of all expectations, no new relationship categories were discovered in the study, confirming the exhaustivity of the taxonomic categories derived from the analytical study. Comparing the distribution of types of relationships in the two parts of the empirical study, there were some differences: sequential and whole-part relationships were more likely shown by explicitly coded fields, while equivalence, derivative, descriptive, and accompanying relationships were more likely to be embedded in a general note. The results of findings from the empirical study revealed that most records containing bibliographic relationships were records for English language works, published in the United States, between 1970 and 1979. An exception was resources involved in equivalence relationship, predominantly published in the 1980s.

From the data collected for distribution by subject, it was found that Science was mainly associated with accompanying relationships, Social science with whole-part relationships, Humanities with derivative relationships, and the Unknown category with equivalence relationships. Books was the primary format associated with equivalence, derivative and whole-part relationships. Audio-visual materials were involved in mostly accompanying relationships. Not

2 The predominant relationship type was the accompanying relationship with 326 occurrences (representing 17.7% of the total records of the sample), derivative relationships followed with 302 occurrences, equivalence with 123, whole-part with 40, sequential with 5, and descriptive relationships with 2 occurrences.

surprisingly, serials were involved almost totally with sequential relationships.

As for the results of the first part of the empirical study, Tillett highlights three conditions that complicated the analysis: the counting method, the limitations of the Library of Congress's computers, and the nature of bibliographic records. Regarding the counting method, data were not collected in a manner that allowed the author to correlate factors referring to a particular bibliographic record, but rather were collected for each factor separately for each group of bibliographic records with a specific code. Therefore, it was not possible to statistically correlate the data collected for each factor. The second condition was a limitation on the use of the Library of Congress's computer, which could not be used exclusively for the purposes of the study because, while running the queries, it was conducting its routine operations. Finally, the analysis of results was complicated by the nature of bibliographic relationships; one bibliographic record may present many bibliographic relationships, as in the case of a work that has a copy, a revised edition, a supplement and is at the same time an anthology of short stories. Additionally, a bibliographic record may include more than one instance of a given type of relationship, such as a record for bibliographic work with an earlier edition and a later edition, which thereby has two derivative relationships to different works. It follows that in the count of records associated with any relationship type, the same record may be counted more than once.

## Conclusion

The concepts discussed in the doctoral dissertation of Barbara B. Tillett of 1987 are the basis for any scientific study on entities, attributes, and relationships of bibliographic resources. The themes were authoritatively taken up in 1989 by Elaine Svenonius (Tillett's PhD advisor) and Tom Delsey (also a student of Svenonius) in the Cataloguing Section of IFLA. Svenonius presented her ideas in *Bibliographic entities and their uses*, the invited paper for the *Seminar on Bibliographic Records*, the Pre-IFLA seminar held in Stockholm in 1990, and organized by IFLA UBCIM Programme and IFLA Division of Bibliographic Control (Svenonius, 1992). The Section approved the effort to develop what in 1998 was called FRBR, *Functional Requirements for Bibliographic Records*, the most important theoretical document issued by IFLA in the last decades. FRBR is at the base of the ICP, *International Cataloguing Principles* of 2009 (with updates and later versions), the new IFLA consolidated model LRM: *IFLA Library Reference Model* of 2017 (Riva et al.,

2017), and of the RDA: *Resource Description and Access* (Delsey, 2016), the international metadata standard, intended for the digital age and structured according to entities, attributes and especially bibliographic relationships.

The fact that Svenonius's essay was republished in 2018 by the *Cataloging & Classification Quarterly* (Svenonius, 2018), twenty-eight years after it was presented, is highly significant because it confirms the fundamental role of the research carried out by Elaine Svenonius and her students Barbara Tillett and Tom Delsey, the research that is being considered the essential reference point for contemporary cataloguing literature.

## REFERENCES

- DELSEY, T. (2016). *The Making of RDA*. *JLIS.it*, 7 (2), 25-47, DOI:10.4403/jlis.it-11706.
- GOOSSENS, P. and MAZUR-RZESOS, E. (1982). Hierarchical Relationships. In *Bibliographic Descriptions: Problem Analysis, Library Systems Seminar*, Essen, 1981. Essen: Gesamthochschulbibliothek Essen.
- REITZ, J.M. (2004). *Dictionary for Library and Information Science*. Westport, Conn.: Libraries Unlimited, [https://www.abc-clio.com/ODLIS/odlis\\_d](https://www.abc-clio.com/ODLIS/odlis_d) (20-01-2019).
- RIVA, P., LEBŒUF, P. and ŽUMER, M. (2017). *IFLA Library Reference Model*, [https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017\\_rev201712.pdf](https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017_rev201712.pdf) (20-01-2019).
- SVENONIUS, E. (1992). Bibliographic entities and their uses. In R. Bourne (ed.) *Seminar on Bibliographic Records: Proceedings of the Seminar held in Stockholm, 15-16 August 1990*(pp.3-18). München; New York: Saur.
- SVENONIUS, E. (2018). *Bibliographic entities and their uses*. *Cataloging & Classification Quarterly*, 56 (8), 711-724, <https://doi.org/10.1080/01639374.2018.1524284>
- TILLET, B.B. (1987). *Bibliographic Relationships: Toward a Conceptual Structure of Bibliographic Information Used in Cataloging*. Thesis (Ph. D.). University of California, Los Angeles.
- TILLET, B.B. (2001). Bibliographic Relationships. In C. Bean and R. Green (eds) *Relationships in the Organization of Knowledge* (pp. 19-35). Dordrecht: Kluwer Academic Publishers.
- UNIMARC: Universal MARC Format. (1980). (2<sup>nd</sup> revised ed.). London: IFLA International Office for UBC.

# ENTITETI, ATRIBUTI I BIBLIOGRAFSKI ODNOSI PONOVO ČITANJE DOKTORSKOG RADA BARBARE B. TILLET NAKON TRIDESET GODINA

## KLJUČNE RIJEČI:

*bibliografski odnosi, katalogizacija,  
Barbara B. Tillett*

## SAŽETAK

*Bibliografski odnos logična je veza koja se stvara između dviju ili više bibliografskih cjelina. Te su veze najvažniji instrument za pomoć korisniku u kontekstu kataloga. U ovom se članku želi prikazati istraživanje koje je od 1981. do 1987., u sklopu svog doktorskog rada, provela Barbara B. Tillett o bibliografskim odnosima. To je prva iscrpna teorijska analiza bibliografskog svijeta čiji se rezultati nalaze u temeljima promjena koje su imale veliki utjecaj na katalogizaciju i bibliografsku organizaciju znanja u posljednjih trideset godina. Entiteti, atributi i bibliografski odnosi odigrali su središnju ulogu u stvaranju konceptualnog modela Uvjeti za funkcionalnost bibliografskih zapisa (FRBR – Functional Requirements for Bibliographic Records), a i dalje su od ključne važnosti u oblikovanju standarda, kao što su RDA: Opis i pristup građi (RDA – Resource description and access) te IFLA-in knjižnični referentni model (IFLA LRM – IFLA Library Reference Model). Proučavanje entiteta, atributa i bibliografskih odnosa važno je i aktivno polje koje je svojim radom zadužila Barbara B. Tillett.*