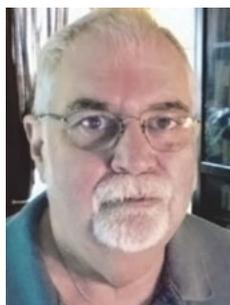


# Work†

Richard P. Smiraglia

Institute for Knowledge Organization and Structure, Inc., Shorewood WI 53211, USA,  
<xxxx@knoworg.org>



Richard P. Smiraglia holds a PhD in information from the University of Chicago. He is Senior Fellow and Executive Director of the Institute for Knowledge Organization and Structure, Inc. and is Editor-in-Chief of this journal. He also is Professor Emeritus of the iSchool at the University of Wisconsin-Milwaukee. He was 2017-2018 KNAW Visiting Professor at DANS (Data Archiving and Networked Services division of the Royal Netherlands Academy of the Arts and Sciences), The Hague, The Netherlands, where he remains visiting fellow and was the 2018 recipient of the 2018 Frederick G. Kilgour Award for Research in Library and Information Technology.

Smiraglia, Richard P. 2019. "Work." *Knowledge Organization* 46(4): 308-319. 58 references. DOI:10.5771/0943-7444-2019-4-308.

**Abstract:** A work is a deliberately created informing entity intended for communication. A work consists of abstract intellectual content that is distinct from any object that is its carrier. In library and information science, the importance of the work lies squarely with the problem of information retrieval. Works are mentefacts—intellectual (or mental) constructs that serve as artifacts of the cultures in which they arise. The meaning of a work is abstract at every level, from its creator's conception of it, to its reception and inheritance by its consumers. Works are a kind of informing object and are subject to the phenomenon of instantiation, or realization over time. Research has indicated a base typology of instantiation. The problem for information retrieval is to simultaneously collocate and disambiguate large sets of instantiations. Cataloging and bibliographic tradition stipulate an alphabetico-classed arrangement of works based on an authorship principle. *FRBR* provided an entity-relationship schema for enhanced control of works in future catalogs, which has been incorporated into RDA. *FRBRoo* provides an empirically more precise model of work entities as informing objects and a schema for their representation in knowledge organization systems.

Received: 21 September 2018; Accepted: 22 September 2018

Keywords: works, mentefacts, catalogs, disambiguation, collocation, FRBR, FRBRoo, information retrieval, semiotic, superworks, canonicity

† Derived from the article of similar title in the ISKO Encyclopedia of Knowledge Organization, Version 1.0; published 2018-09-18, last edited 2019-05-02. Article category: Core concepts in KO

## 1.0 Introduction

A work is the essence of a creation, such as a novel, a symphony, a painting, a statue, a thesis, etc., intended by its creator to be communicated with some audience. More formally, a work is a deliberately created informing entity intended for communication. Known variously in humanistic disciplines as work, oeuvre, opus, etc., a work can be a critical entity for ordering and retrieval in bibliographic information systems such as catalogs or indexes. A work consists of abstract intellectual content that is distinct from any object that is its carrier. This distinction between the work and the item that carries it is critical for information retrieval, because the attributes of works are those of their abstract intellectual content, whereas the attributes of items are those of (usually physical) informing objects.

Works need not be literary or even textual. Although most libraries contain mostly books, a wide variety of other means of expression are represented in information retrieval systems. Certain works become very well known, and it is

these works for which many iterations might come to be represented together in information retrieval systems. In culture at large, some works serve iconic roles—think of *Mona Lisa* (Da Vinci) or *Eroica Symphony* (Beethoven) or *Iliad* (Homer), or even *Gateway Arch* (Eero Saarinen) for examples associated with individual creators, or *Bible* or *Kama Sutra* for examples of works that are not associated with specific individual creators. These works, which serve somehow iconic roles, can be viewed as semiotic entities—signs, in other words—imbued with various cultural reputations that might extend beyond their intellectual content.

In this article we will consider the importance of distinguishing between works and items and we will look briefly at the history of the treatment of works in information retrieval. We will consider carefully the nature of works including their cultural meaning. We will look at the phenomenon of instantiation that underlies the evolution of sets of works derived from a common progenitor, and we will consider the major conceptual schema currently in use for representing works.

## 2.0 The importance of works for information retrieval

The concept of the work is of great interest to scholars in the humanities, most notably in literary criticism, philosophy, and musicology, as well as in the interdisciplinary exercise of textual criticism that crosses the boundaries of bibliography and literary criticism. Several authors have called for a theory of the work, notably Foucault (1984) and Tanselle (1989), and others have explored the cultural meaning of works, such as Talbot (2000) and Goehr (1992). But in library and information science the importance of the work lies squarely with the problem of information retrieval. It is widely accepted that information retrieval systems should collocate the works of a particular author, and furthermore, that within that collocated list the iterations—translations, editions, etc.—of a particular work should likewise be collocated. But the problem arises, that individual publications, from which indexed citations are harvested by transcription, do not necessarily have identical identifying marks. That is to say, under Charles Dickens, entries for editions of books with the title *Bleak House* will file together, but translated editions with titles *Hokumeikan monogatari* and *Maison d'Aprè-Vent* will not collocate. So a convention is required to cause all of the iterations of *Bleak House* to be filed together in the system, as well as to keep them distinct from the other works by Charles Dickens. The uniform title was used to good effect for this purpose in Anglo-American and other cataloging traditions through most of the twentieth century, recently renamed the preferred title.

However, a second problem arises, and that is the subsequent problem of disambiguation in a file of apparently identical entries. What do we do with a list of hundreds of English-language editions of *Bleak House*? Typical solutions are to sub-file by publisher or date, but that does not necessarily sort the intellectual content in ways that might be appropriate for retrieval. The same is true for illustrations of Saarinen's *Gateway Arch*—none of them is the arch, and all of them are representations of the arch, but all of them are also different from each other. So, disambiguation of large retrieval clusters that might otherwise appear identical requires understanding the nuances of the iterations of works, and the resulting instantiations that might all be present in such a cluster. "Instantiation," then, is the phenomenon associated with works (and also with all other informing objects) that describes the patterns of iteration over time that result in large, ambiguous clusters in retrieval systems. What has been required is a means of separating the work entity from other entities such as documents (Buckland 2018), sources or information objects in information retrieval systems (Smiraglia 2002a).

## 3.0 A brief history of the treatment of works in information retrieval

We do not have space here to rehearse the entire history either of the catalog or of information retrieval. But, it is important to understand that in the development of the library catalog the movement from "inventory of books" to "device for indexing works" has been a long haul. One must appeal, of course, to Strout's (1956) famous history of the catalog. And there one will learn that most early attempts to create catalogs look to our eyes like inventories. One will also learn that the likely crux that caused greater sophistication in the construction of catalogs (things like entry of works under author names, subject gatherings, and so forth) were introduced not by librarians but by booksellers. This should not be surprising. In a time when few were literate the librarian likely had a good grasp of the books under his charge. It was only when it became important to sell books, and various diverse iterations of books, that catalogs required ever greater sophistication.

Nevertheless, by the Enlightenment, the work had begun to receive the attention of librarians such as Thomas Hyde, and by the mid-nineteenth century, Anthony Panizzi. These famous librarians (and engineers of catalogs, and they were engineers) had to concern themselves with the issue of disambiguation. The famous hearings before Commons in which Panizzi ([1848] 1985) defended his catalog structure made that clear. The reader was not so much interested in a particular book, Panizzi asserted, as in a particular work, no matter in what book it appeared. By the twentieth century the issue had become critical for modern librarianship. Eva Verona (1985) wrote about the notion of the literary unit, which was to be a collocating device for a work. The reason—that a nascent information explosion had begun to present librarians and the reading public with many diverse editions, translations, and commentaries of the most sought works.

The catalog was originally structured as an inventory of books. Describe this "item" by transcribing its title page. Then disambiguate the description as need be to make it serve several filing masters by adding subject headings, author headings, and so forth. The uniform title was added in rare instances to collocate editions of canonical works for which a library might have dozens of iterations. Still, however, the work itself remained without operational definition. The second edition of the *Anglo-American Cataloguing Rules (AACR2 1988)* led to change by presenting a modern approach to information retrieval requiring the use of uniform titles for works that had appeared under different titles. This offered a superstructure of works—a set of alphabetico-classified solutions for organizing (and disambiguating) large files of works (see below).

In the latter decades of the twentieth century, research began to provide empirical evidence of the extent of the panoply of iterations of works that might require simultaneous collocation and disambiguation for information retrieval. Studies were conducted using random samples of works gathered from various bibliographical sources, including academic libraries, bibliographic utilities, music and motion picture libraries, and canonical lists of works (Smiraglia 1992; Yee 1993; Vellucci 1997; Smiraglia 2001, 2007b; Petek 2007). User studies suggested more natural groupings of works could be achieved in the catalog (Carlyle 1999). Major results indicated that large proportions of the works that make up library collections exist in multiple iterations, that cultural phenomena seem to play a role in determining which works will do so, and that clustering works by meaningful identifiers would be most useful.

#### 4.0 The nature of works

##### 4.1 Cultural meaning

Works are mentefacts (Gnoli 2018). That is, they are intellectual (or mental) constructs, but as such, they serve as artifacts of the cultures in which they arise. The meaning of a work is abstract at every level, from its creator's conception of it, to its reception and inherence by its consumers. Various semiotic theories have been used to describe this phenomenon. In Saussurian terms, a work at the time of its origin is fixed in its creator's intellect and is theoretically, if only for a moment, immutable. But once the work has been offered to the public it is overtaken by its reception, and it becomes infinitely mutable. Furthermore, such works then function along the lines of Peircian symbols—they have cultural meaning in a broad sense that is determined both collectively and individually. Thus, a popular work becomes the property of those for whom it is popular. What is *Mona Lisa*? A painting? Yes, but much more; it is a human mystery, a mysterious woman, a sign of the human condition—these and many other interpretations keep the painting alive in the consciousness of millions who have never seen the painting and millions more who have never even seen its likeness. Yet they all know what a “Mona Lisa” is—it has iconic semiotic status (that is, its role in culture is larger than life). This is the status works may achieve. And once they do, the job of the library-and-information science community is to curate them.

Works are said to be entities of what Patrick Wilson ([1968] 1978) called the bibliographic universe, a sort of concept-space in which all recorded knowledge exists, represented by specific texts, all in relationship with each other. This metaphor has been extended to show the role of works in knowledge organization (Smiraglia and van den Heuvel 2013, 373):

Therefore in our construct the metaphorical bibliographical universes are populated by entities—knowable elements of reality—that can be seen to exist in relationship to each other—relationships of nearness and distance, of joint motion, of evolution over time, etc. Smiraglia (1996) extended the metaphor logically by suggesting that works and their instantiations cluster in metaphorical constellations, having orbital and therefore gravitational relationship to each other, and that there are different sorts of celestial bodies in the bibliographical universe. These “constellations” are groupings of instantiations of works—not only the progenitor work itself, but also its editions, translations, abridgments, adaptations, excerpts, etc., and their instantiations as well. These have been termed variously “bibliographic families” (Wilson), “superworks” (Svenonius 2001), “textual identity networks” (Leazer and Furner 1999), and “instantiation networks” (Smiraglia 2008).

Curatorship demands understanding and elucidation of the ineluctable qualities of these mentefacts. Thus, a music librarian must know not only Beethoven or his *Eroica symphony* but also the story of the Napoleonic wars, the history of the Hapsburgs, the cultural and scientific evolution of the symphony as an icon of western musical sophistication, the history of the rise and fall of the symphony orchestra, the appreciation of Beethoven and his rise to cult status, and so on. Placing the editions, Ur-texts, scores and parts, recordings, and other iterations of this “work” requires multidisciplinary knowledge. Such is the task of the cataloger and this is but a single example.

Day (2008, 44–45) drew a distinction between literary works and works of art, suggesting that works should be viewed as “events that are constitutive of meaning by virtue of their negotiation of cultural and social horizons through material forms and techniques.” Works of art (sculpture, painting, etc.) are seen as the result of “work” (in the sense of labor expended) that results in a physical object that may be understood as site-specific and time-valued. Drawing at once on Heidegger and the Visual Resources Association *VRA Core* standards for description of visual works,<sup>1</sup> Day situates works of art and literary works in different epistemic traditions, such that literary works that might begin as mentefacts occupy an epistemic zone in which the transference of their content (i.e., instantiation) is seen as a metaphysical property of their social position, but works of art occupy an epistemic zone in which their technological realization as objects is post-metaphysical. In this scenario the re-presentation of a work of art does not belong necessarily to the same categorically-bounded set as the work itself, whereas the literary work and its re-presentations are all members of the

same set. Day reports that the VRA Core standard distinguishes works of art by attributes of entities such as time period of creation, location of discovery, and current curation. Similarly, the CIDOC Conceptual Reference Model (CRM)(<http://www.cidoc-crm.org/>), which is a multi-disciplinary international meta-level ontology for cultural heritage information sharing, approaches all works (art or otherwise) as results of the events that brought them into being as well as those associated with their persistence across time.

Works also are ontological realities, which makes them objects for knowledge organization. A clear example is the manner in which representations of instantiations of works are gathered in information retrieval systems under name-title appellations—kinds of nominal historicist epistemological anchors—but then subdivided in detailed schema based on mutation of the work’s ideation or its actual expression in text. Works are vehicles for communication (Smiraglia 2001, 57), which also means that they are social entities shaped by culture. Because works are core narratives in every part of human experience—from sacred texts to legal foundations to iconic structures to iconic novels—they have been studied as constructs in many disciplines.<sup>2</sup> Philosophy is the discipline that most closely touches knowledge organization, and in that field many voices combine to extol the meaning of a text even as the same voices seek to promote their own texts. We already have mentioned Foucault and his search for the “author”; Barthes’ metaphor of text as tissue (1975)—impermanent, non-persistent, and utterly interpreted only on its reception—makes it clear that even textual works, like statues, lie in the conscious minds of those who behold them. This aligns with Goehr’s reception theory of musical works (1992). Works are culturally critical, but they also are impermanent fixtures in the minds of people.

#### 4.2 The properties of works

Works are mentefacts, which means they are abstract and made up of ideas. Works, obviously, are not the only kinds of mentefacts that occur in information retrieval, but they are unique in the fact that, as creative expressions, they are in some sense ideationally fixed.<sup>3</sup> Works have two properties, which are referred to as ideational content—the ideas expressed—and semantic content<sup>4</sup>—the mode of expression of the ideas. Either ideational or semantic content might be changed in subsequent iterations of a work. Stories abound about authors arriving at print shops in the middle of a press run and changing one word here or there, thus resulting in very slight alterations of semantic content that (likely) do not affect ideational content. But more often, works are translated or abridged or reissued with illustrative matter. In these cases, it is possible to trace over

time the evolution of a work as its semantic content changes—such iterations have been termed “derivations” (Smiraglia 2002b). In other cases, works might be adapted for reuse in children’s versions, as screenplays, as librettos, etc. In these cases, it is possible to trace the evolution of the alteration of the ideational content—such iterations have been termed “mutations.”

Works are a kind of informing object, alongside more usually physical entities, such as documents (Buckland 2018), or archival records or naturally occurring artifacts. Like all such informing objects, then, works are subject to the phenomenon of instantiation. The term “instantiation” describes the phenomenon of realization over time (Smiraglia 2008). We have learned that the majority of works exist only in their original instantiation, but that significant numbers (likely around one-third in the bibliographic universe but one-half or more in library collections) exist in multiple instantiations. For every two or three one-off books there is a work like Steinbeck’s *Grapes of Wrath* that exists in hundreds of editions, and translations, and becomes a screenplay, a motion picture, and so on. The first-published iteration of a work in such a set has been termed the “progenitor,” and we know that older progenitors are associated with larger networks of instantiations, but more recent progenitors are associated with more complex networks of instantiations. That is, works that originated centuries ago are likely to have large networks of editions and translations, but works that originated recently, if they are associated with large instantiation networks, are more likely to have many mutations in their instantiation networks.

We have seen above the two main categories of instantiation—derivation and mutation. From research it has been demonstrated that bibliographic works have at least the following types (Smiraglia 2002, 11):

<b>Derivations</b>
simultaneous editions
successive editions
predecessors
amplifications
extractions
accompanying materials
musical presentation
notational transcription
persistent works
<b>Mutations</b>
translations
adaptations
performances

Research has shown that the majority of instantiations are simultaneous editions (works published in more than one place at one time), successive editions (second, third, etc., subsequent editions), and translations (Smiraglia 2001). The distinction between derivation and mutation is the degree of alteration in ideational content. We assume alteration of the semantic content occurs from one edition to the next. But major change in the presentation of ideas occurs as the work evolves over time in accord with cultural stimuli, which act as market forces to compel motion pictures, musical realizations, and so forth (e.g., the motion pictures *Prisoner of Zenda* 1937 and 1952, with identical screenplays and music, based on Anthony Hope's novel of 1894, or the motion pictures *The Bishop's Wife* 1947 and *The Preacher's Wife* 1996, based on Robert Nathan's 1928 novel *In the Barley Fields*, but with almost unrecognizable characters and locales). Market forces are very much present in the research on instantiation (Smiraglia 2007b). We have seen, for example, that works that are associated with very large instantiation networks are more likely to have been published simultaneously at the outset—a strategy well known in publishing.

In a somewhat different vein, Furner (2009, 10) has suggested that works can be described both as relations among things, and as identities of the properties of relations themselves:

There are a number of different kinds of entities that are capable of entering into relations with one another. We might find it convenient to distinguish in some way between worlds, works, words, persons, and so on. It doesn't really make much difference whether we decide to treat these entities as substances that somehow exist separate from their properties, or simply as bundles of properties. Whatever system of fundamental categories of entities we settle on, we can also use it as the basis of a taxonomy of relations between entities. Depending on our purposes, we might want to distinguish (a) relations between works and people from (b) relations between works and other works, for instance.

This Furner relates to the philosophical positions of nominalism (2010, 186), by which he says they exist as sets of relationships between linguistic expressions, and realism, e.g., the concreteness of a stipulated text. That is, a work is made up of ideation expressed semantically. Judgments about how two instantiations might be exemplars of the same work rely on both nominalist points of view about whether both texts bear the same sets of relationships to other texts, and to realist points of view about the exact match between the textual strings that constitute the expression of the work. He demonstrates that the relation-

ship between two documents that might be instantiations of the same work has the same identity as the relationship between two documents that might be about the same concept, because they are both members of the same class (12):

If I decide that Doc 1 instantiates Work A, what that amounts to is a judgment—an entirely subjective judgment made by me on a particular occasion—that Doc 1 has the property of being an instance of Work A, that Doc 1 is a member of the class of documents that share the property of instantiating Work A ... These are just different ways of saying the same thing, and again we can also say that Doc 1 and Doc 2 are similar in the sense that they share the same property or that they are members of the same class. And again, if it turns out that Work B has exactly the same extension as Work A does—in other words, if it turns out that all and only the documents that instantiate Work A instantiate Work B—then we can say that Work A is the same work as Work B.

A related aspect of the phenomenon of instantiation is the re-presentation of content, and it is this property that is endemic in the incorporation of information objects in retrieval systems in which large clusters of seemingly similar content must be simultaneously gathered and disambiguated. Empirical analysis of this phenomenon in museums and archives demonstrated the means by which not only visual representations of specific objects but also metadata associated with them require control—gathering and disambiguation—around a nominal anchor that usually is the identifier for the work (Smiraglia 2006, 2007d, 2008). This research aligns with similar observations from Coleman (2002) concerning scientific models,<sup>4</sup> and Greenberg (2009) with regard to life-cycle modeling of data records from evolutionary biology. The concept was recently extended (Smiraglia 2017) to the re-presentation of data in repositories. It is perhaps at this point that the epistemic distinction drawn by Day (2008)—between literary works and works of art helps distinguish between the clusters of instantiated realizations of works that reside primarily as texts and the clusters of instantiated re-presentations of metadata associated with works that reside as objects—helps both to inform the understanding that works of different kinds possess different properties in fact as well as socially and culturally, as well as the comprehension that the central problem of the work for knowledge organization is its treatment as the nominal anchor for clustering in knowledge organization systems.

Still, probably the most important empirical finding from the empirical research is the discovery that there is a cultural catalyst for the growth of a family of works all de-

rived from a common progenitor. Initially, borrowing a phrase from Wilson ([1968], 1978), these were called bibliographic families. In *Functional Requirements for Bibliographic Records* (FRBR, IFLA 1998) they are called superworks. These are works like *Gone with the Wind* that have achieved iconic status, and thus for which potentially thousands of iterations have come forth, all of which can be associated with a common progenitor through shared ideational and semantic content (Smiraglia 2007a). Nonetheless, not everything in the superwork set *Gone with the Wind* is equivalent with Margaret Mitchell's novel. Instead, ideational nodes within the set (such as a screenplay) are related works that have their own instantiation sets. The problem for information retrieval, as stated earlier, is to simultaneously collocate and disambiguate these large sets of instantiations.

## 5.0 The major conceptual schema currently in use for representing works

### 5.1 AACR2

All of this means that bibliographic works are very complex entities to handle in systems for information retrieval. The simple style of cataloging described earlier is insufficient to disambiguate the large collocated networks of instantiations associated with many bibliographic works. The *Anglo-American Cataloguing Rules, Second Edition* (AACR2 1998) contains within its complex rules for "Headings, Uniform Titles, and References" a set of requirements for attribution, denomination, collocation, and disambiguation of the instantiations of works. Initially, works are divided into those that are associated with a specific creator and those that are not, such as:

Hemingway, Ernest  
Sun Also Rises  
and  
Episcopal Church  
Book of Common Prayer  
but  
Cloud of Unknowing

where works are entered in creator-title citation form, but title alone for works not associated with a particular creator. Denomination of works is dependent on their period of origin, with works promulgated primarily after the invention of printing from movable type (actually, the year 1500 is stipulated) entered under the version of the original title by which they have become known:

Dickens, Charles  
Pickwick papers

but  
Beowulf  
or  
Artistotle  
Meteorologica

a well-established English title is used for works originating before 1501. A set of terms (e.g., Selections, or Works, or Plays) are allowed for collocating collections under a single author. Parts of a specific work published separately are entered first under the original work and then qualified with the name of the part:

Tolkien, J.R.R.  
Two towers

for part two of Tolkien's trilogy *Lord of the Rings*, and "Selections" may be used also to designate a set of extracts from a work:

Gibbon, Edward  
History of the decline and fall of the Roman Empire.  
Selections.

A translation is entered under the heading for the original work and qualified with the language of translation thus:

Caesar, Julius  
De bello Gallico. French & Latin.

The net effect is an alphabetico-classed arrangement of works under their headings. For example, we might find a set like the following in the catalog of a single library:

Dickens, Charles. Works  
Dickens, Charles. Selections  
Dickens, Charles. Bleak house  
Dickens, Charles. Bleak house. French  
Dickens, Charles. Great expectations  
Dickens, Charles. Great expectations. Selections. German  
Dickens, Charles. Pickwick papers

and so on. This effect is perhaps most pronounced under headings for legal works and the Bible:

Bible. English. New Revised Standard.  
Bible. N.T. Timothy  
Bible. O.T. Pentateuch

and so forth. The effect of this arrangement is to accomplish collocation under specific headings and sub-headings, but it leaves disambiguation to chance or to the ex-

expertise of a user. In a small library the user will likely not have difficulty making a selection from such a file, but in a bibliographic utility one can retrieve a search result of hundreds of headings for *Pickwick Papers* with no identifiable distinction among them readily apparent.

These rules sit at the apex of Anglo-American cataloging tradition stretching from the late eighteenth century to the late twentieth. This tradition relies on an authorship principle that has been shown to occasionally override cultural discourse in favor of assigning any work to a personal name, no matter how distant the named person might be from the creative task (Smiraglia and Lee 2012). It was only in the late nineteenth century that sufficient commercial interest arose in the profitable marketing of authorship, which took the form of increased production of works associated with specific names (Smiraglia, Lee, and Olson 2010). Discourse analysis of the authorship principle revealed multiple meanings for “author” (Martínez-Ávila et al. 2015, 1110):

In cataloging tradition, and to some extent in classical bibliography, an author is foremost a named entity to whom intellectual creativity is attributed. But also, and almost more importantly, in cataloging and bibliographical tradition, as the discourse has been transformed to this date, an author is the name of a class of related works that can be collocated with the iconic representation of the named entity ... The discussion leads inexorably to the conclusion that an author is not so much a person who writes, as it is the name of a class of works that can be related, either through power structures or lived experience, with a specific named entity.

The work, then, has been used as the core historical anchor for an alphabetic-classed arrangement of instantiations in the library catalog (Smiraglia 2003). This is all to change with the incorporation of the *FRBR* (IFLA 1998) conceptual model.

The following illustration (Figure 1) demonstrates the limits of librarianship’s ability to comprehend either the core ontological importance of works or the complexity revealed by empirical research. Derived from Tillett (2004, 4; 2001, 23) the figure arrays categories of work content relationships in the form of a trajectory that embraces the point at which library cataloging rules distinguished between “versions” of a work and emergent “new” works.<sup>5</sup>

Without reference to the research on mutation or that on expression and manifestation entities, this diagram shows some of the kinds of publications in library catalogs that require collocation and disambiguation of instantiations of works. The column to the left identifies copies, the central column lists kinds of derivative “mutation” in-

stantiations that we saw above in Table 1, and the column on the right identifies things like book reviews, that are not part of the instantiation network of a work, but rather are works about the work.

## 5.2 The *FRBR* conceptual model and RDA

*FRBR* (IFLA 1998) sets out an entity-relationship model for the bibliographic record that separates the inventory functions of the catalog that are item-based from the searching functions that are work- or subject-based. A simple schema represents the entities as works, expressions, manifestations, and items. In this schema, works remain abstract, and items represent physical entities. Expressions and manifestations are the entity names for all forms of instantiation, wherein expressions identify specific realizations of works, particularly with regard to semantic content, and manifestations identify physical embodiments of expressions. J.S. Bach’s *Art of the Fugue* is a work, the score of it or a performance of it are different expressions, and Breitkopf & Härtel and Schirmer editions of the score, or Deutsche Grammophon and Nonesuch recordings of a performance are manifestations. The manifestations then reside in specific, physical items. This set of distinctions allows the separate inventory of instantiations according to their intellectual attributes. Change in ideational content results in a new work, change in semantic content results in a new expression, and the role of the publisher who brings an expression to market is recognized in the production of manifestations. Because much of the complexity of current online catalogs results from the admixture of entity data in bibliographic records, *FRBR* promised a more articulate, if still complex, approach to access to works and their iterations.

Problems with the *FRBR* conceptual model inhibited its full implementation although many approaches were undertaken in the library and bibliographic utility community (see, for example, Smiraglia 2013). The most difficult problems were with the precise implementation of the expression entity and with gaps in the model, principally for aggregate works (works that include other works, such as anthologies, or journals) (Le Boeuf 2006; Smiraglia 2012). Research into the nature and treatment of aggregates fed into a 2015 report by an IFLA working group and is carefully reported in O’Neill, Žumer and Mixter (2015). Aggregates, which are arguably themselves “works,” were determined to be very common occurring more than 20% of the time in one sample (128); the majority were anthologies, conference proceedings, scholarly journals and compilations (127). A multiplicity of aggregations of expressions of works—original essays, reprinted articles, translations, etc.—commonly occur in aggregates. A conceptual model consisting of three types of aggregation—collec-

# Family of Works

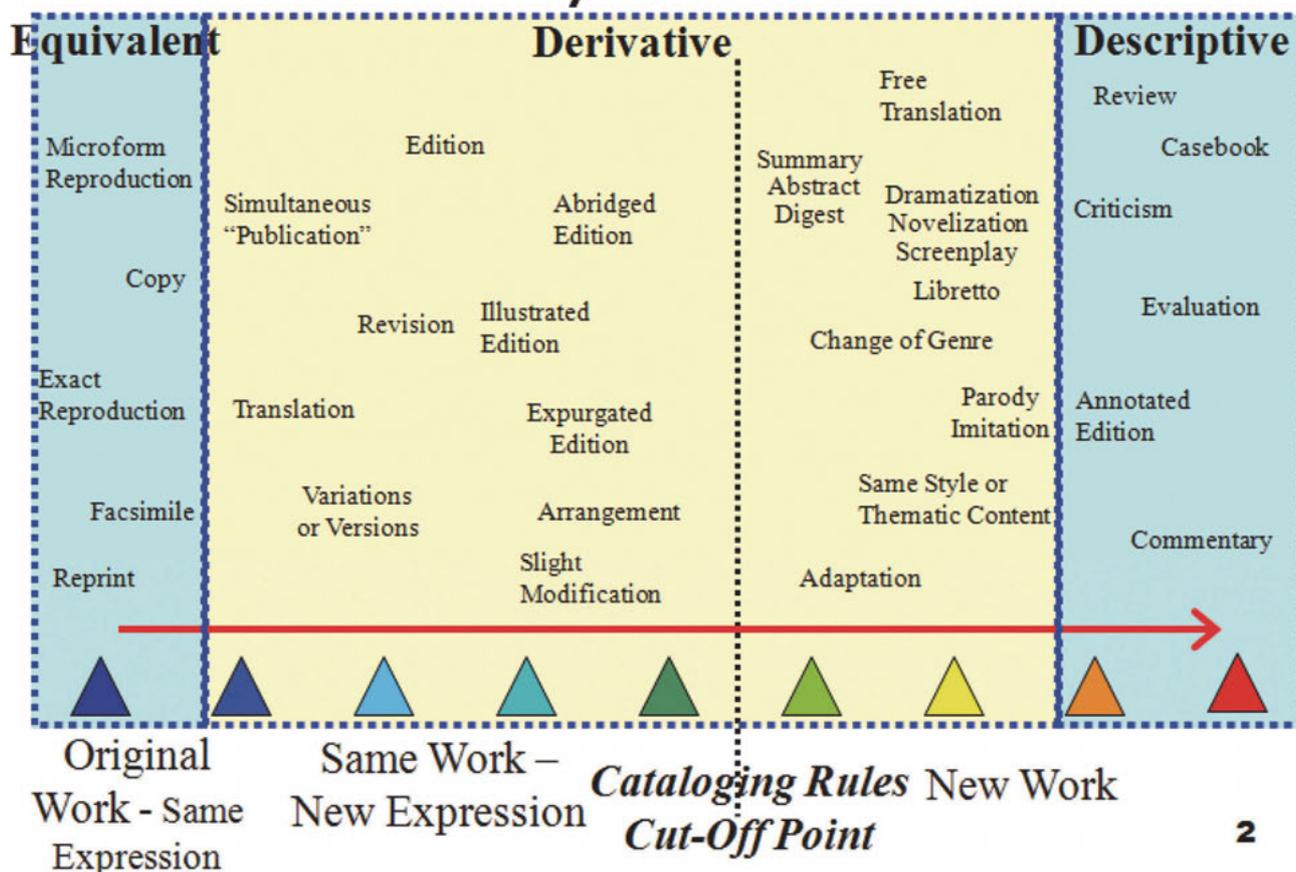


Figure 1. Tillett's (2004, 8; 2001, 23) content relationships.

tions, augmentation aggregates, and parallels—was reported in O’Neill and Žumer (2012).

Many of these problems have been tackled in a reinterpretation of the FRBR conceptual model from its original entity-relationship model into an object oriented model (Bekiari et al. 2015), which is harmonized with the cultural heritage information sharing ontology known as the CIDOC -CRM. Known as FRBRoo, the empirically-based object-oriented model overcomes the earlier difficulties by removing temporal requirements for expressions and allowing aggregation (Smiraglia 2015, 297):

Entities are broken into associated phenomena named “objects,” and are oriented to each other by their attributes. The various models, then, do not rely on temporality or mutually exclusive classes, but rather on associative principles of linked attributes. FRBR’s “W-E-M-I” (works-expressions-manifestations-items) entities in FRBRoo become objects that may be associated multiply according to their related attributes. In a given instantiation network derived from an ideational conception there might be many

works, for each of which there might be many expressions. Not all expressions spawn manifestations, and so forth. Also a distinction is made between the intellectual work, and publication events, which might spawn manifestations.

In 2010 a new set of international (but mostly Anglo-American in practice) cataloging instructions were introduced under the rubric “Resource Description and Access” (RDA). RDA embraces the FRBR entity-relationship conceptual model and divorces problems in transcription from manifestations or inventory control of items from those of representing works and their creators. There is very little difference between RDA and AACR2 in the outcome of work identifiers—an alphabetic-classed system of works, ordered by their preferred titles<sup>6</sup> under authorized access points for their creators still is used—but more flexibility is available for multiple attribution where creatorship is complex or even unattributed. Fuller implementation of the FRBR conceptual model through the use of RDA is leading to the more appropriate representation of works in information retrieval systems, with both better

clustering and better disambiguation. Problems still remain in RDA with works that are performed and recorded (Smiraglia 2007c)—so that no distinction is made between the work that is a recording of a performance of *Eroica Symphony* and the work that is its performance—although FRBRoo provides a mechanism to do so.

The most recent development incorporating works into a functional conceptual model is represented by the 2017 *Library Reference Model (LRM)*, which offers a harmonization of the entire family of FRBR conceptual models and is itself harmonized with the CIDOC CRM. The *LRM* defines the “work” entity as “the intellectual or artistic content of a distinct creation” and its expression as “a distinct combination of signs conveying intellectual or artistic content” (Žumer 2018, 312).

### 5.3 Incorporating works in classified arrays

Recognition of the classified nature of lists of works produced by cataloging rules has led to the interesting idea that FRBRoo-designated work entities might themselves be used to classify instantiations by their incorporation as auxiliaries into faceted classifications such as the Universal Decimal Classification (UDC). Such treatment relies on understanding the role of works as taxonomic elements of canons made up of “the concatenation of mutable mutating instantiations” (Smiraglia and van den Heuvel 2013, 378):

A canon is the literature accepted as foundational for a domain, and therefore, a canon can be as broad or narrow as its domain. It is canonicity, or acceptance into a canon, that has been demonstrated to be associated with a high degree of instantiation. Put more simply, a work or a set of works, once accepted into a canon, become in demand, which causes more editions, translations, adaptations, commentaries, etc. to be generated by the domain. These canons provide the warrant for most classificatory activity in KO. Instantiation has been shown to be a continuum along which ideation is combined with intellectual force into the expressions of works (Smiraglia [2008]). Motion is the pathway of ideation in the process of instantiation.

A mechanism for using the works entity to link elements of traditional conceptual classification strings either to external ordering systems (such as document retrieval systems) or to W-E-M-I-specific identifiers with as auxiliaries has been demonstrated to be consistent with concepts of multi-versal or multi-dimensional knowledge organization (Smiraglia, van den Heuvel and Dousa 2011). Two cases using FRBRoo to delineate instantiated works, one biblio-

graphic and the other encompassing a musical sound recording aggregate were demonstrated (Smiraglia 2015), and an experiment using FRBRoo to do the same with instantiated open government data also has proven fruitful (Park and Smiraglia 2017).

### 6.0 Conclusion

In this essay, a variety of points of view about the work and its nature have been surveyed, many of them stemming from diverse epistemological understandings. It is now widely accepted in librarianship that a work is a deliberately created informing entity intended for communication, and that a work consists of abstract intellectual content that is distinct from any object that might be its carrier. Works of art inhere in less abstract form in the objects that result from the activity of technologically creating them; those that persist do so in specifics of time and place. Works are mentefacts—mental constructs—but as such they also are cultural artifacts reflecting social values. Works also are ontological realities, which makes them objects for knowledge organization, with properties of ideation and communicative attributes (often referred to as semantic) that are used to positively identify and then bound them. Works and their re-presentations instantiate across time and thereby lies the unity that links all competing definitions. In knowledge organization, the importance of a work is its role as nominal anchor. It matters not whether a work has a sequence of instantiations or exists as an object with representations; what matters in knowledge organization is the identity associated with a work, which itself becomes an iconic conceptual entity in knowledge organization systems.

If bibliographic reality conforms to Patrick Wilson’s vision ([1968] 1978) of a bibliographic universe made up of a vast concept space in which related entities move variously in consort dependent on the intensity or vagueness of their inter-relationships, works constitute the celestial bodies that populate it. Works lie at the center of galaxies of instantiating points. However appealing we find such a metaphor, the reality is that works are essential entities both as cultural mentefacts and as targets for information retrieval. Although this reality has been recognized for a long time, it is only of late that we have gathered sufficient empirical evidence of the works-phenomenon to allow the more powerful relational structure that will underlie future information retrieval systems.

### Notes

1. Day (2008) references several works by Heidegger, but the most critical to his point seems to be Heidegger ([1964] 1977). VRA Core is a set of online data stand-

- ards and schema for the “description of images and works of art and culture” maintained by the Visual Resources Association and the Library of Congress (<https://www.loc.gov/standards/vracore/>).
2. Copyright legislation as a source of cultural warrant for works is discussed in Smiraglia (2001, 68-72), specifically with reference to copyright protection, which subsists in “works of authorship” that may be “literary, musical, dramatic, pantomime or choreographic, pictorial, graphic or sculptural, motion picture or audiovisual, sound recordings, and architectural” (71). An anonymous referee made reference to Warner (1993), which relies on similar material to establish a historical connection between the development of computing from pre-existing information technologies.
  3. An anonymous reviewer asks “are all kinds of mentefacts works?” The answer is no, because a work is, as defined in the first paragraph, “a deliberately created informing entity intended for communication.” A mentefact is not by nature a work, but only becomes one if it is created in a form intended for communication.
  4. An anonymous referee suggests “symbolic” is a better term than semantic, to distinguish the signified aspects of a work. The choice of terms is not so simple. The term “semantic content” as is explained in Smiraglia (2001, 31ff.) is derived from research by Carpenter (1981, 118-20) who relied on work by Wilson ([1968] 1978) and Domanovsky (1974). Thus, the term is the result of the inheritance of empirical thought into the nature of a work in information science. That musical notation, for example, is not “semantic” in the same way as verbal text is obvious but also disregards the purpose of musical notation. Musical notation might be symbolic to some, but to a musician it is entirely semantic. Thus, the term “symbolic,” which is admittedly enticing, is incorrect. Works are mentefacts, embodied by ideational content and communicated by semantic content. That syntactic content might also be useful is discussed in Smiraglia and van den Heuvel (2013).
  4. An anonymous referee asks “Is Einstein’s theory of relativity a work?” and suggests that “work seems not to be [an] important concept in (natural) scientific communities.” But, there are citations throughout this article to Greenberg’s work on life-cycle modeling from botanists, and Coleman’s groundbreaking work on scientific models as works has long needed replication. In fact, what the referee observes is that the hard sciences are less populated by instantiating monographs and more populated by un-tracked instantiating models. Is Einstein’s theory a work? No, but the text in which he introduced it is.
  5. The color illustration is from a Library of Congress pamphlet, which itself was reprinted from a Library of

Congress magazine *Technicalities* (v. 25, no. 5 2003). The illustration originated in Tillet’s 2001 chapter in Bean and Green’s *Relationships in the Order of Knowledge*. The pamphlet and chapter are cited here.

6. RDA Toolkit (<http://access.rdatoolkit.org/>) 5.5 “When constructing an authorized access point to represent a work or expression, use a preferred title for work (see 6.2.2) as the basis for the access point” and 6.2.2.4 “For works created after 1500, choose as a preferred title for work the title or form of title in the original language by which the work is commonly identified either through use in manifestations embodying the work or in reference sources.”

## References

- Anglo-American Cataloguing Rules, Second Edition. (AACR2)*, ed. Michael Gorman and Paul W. Winkler. 1988 revision. Chicago: American Library Assn.
- Barthes, Roland. 1975. *The Pleasure of the Text*, trans. Richard Miller with a note on the text by Richard Howard. New York: Noonday Press.
- Bekiari, Chryssoula, Martin Doerr, Patrick Le Boeuf and Pat Riva, eds. 2015. *FRBR Object-Oriented Definition and Mapping from FRBR<sub>ER</sub>, FRAD and FRSAD*. Version 2.1. Heraklion: International Working Group on FRBR and CIDOC CRM Harmonisation. [http://www.cidoc-crm.org/frbr\\_drafts.html](http://www.cidoc-crm.org/frbr_drafts.html)
- Buckland, Michael. 2018. “Document Theory.” *Knowledge Organization* 45: 425-36. doi:10.5771/0943-7444-2018-5-425
- Carlyle, Allyson. 1999. “User Categorisation of Works: Toward Improved Organisation of Online Catalogue Displays.” *Journal of Documentation* 55:184-208.
- Carpenter, Michael. 1981. *Corporate Authorship: Its Role in Library Cataloging*. Westport, Conn.: Greenwood Press.
- Coleman, Anita S. 2002. “Scientific Models as Works.” *Cataloging & Classification Quarterly* 33, nos. 3-4: 129-59.
- Day, Ronald E. 2008. “Works and Representation.” *Journal of the American Society for Information Science and Technology* 59: 1644-52.
- Domanovsky, Ákos. 1974. *Functions and Objects of Author and Title Cataloguing: A Contribution to Cataloguing Theory*, English text ed. by Anthony Thompson. Budapest: Akademiai Kiado.
- Foucault, Michel. 1984. “What is an Author?” In *Foucault Reader*, ed. Paul Rabinow. Harmondsworth: Penguin, 101-20.
- Furner, Jonathan. 2009. “Interrogating ‘Identity’: A Philosophical Approach to an Enduring Issue in Knowledge Organization.” *Knowledge Organization* 36:3-16.

- Furner, Jonathan. 2010. "Philosophy and Information Studies." *Annual Review of Information Science and Technology* 44. Medford, NJ: Wiley, 159-200.
- Gnoli, Claudio. 2018. "Mentefacts as a Missing Level in Theory of Information Science." *Journal of Documentation* 74: 1226-42. doi:10.1108/JD-04-2018-0054
- Goehr, Lydia. 1992. *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music*. Oxford: Clarendon.
- Greenberg, Jane. 2009. "Theoretical Considerations of Lifecycle Modeling: An Analysis of the Dryad Repository Demonstrating Automatic Metadata Propagation, Inheritance, and Value System Adoption." *Cataloging & Classification Quarterly* 47, no. 4: 380-402.
- Heidegger, Martin. (1964) 1977. "The End of Philosophy and the Task of Thinking." In *Basic Writings: From Being and Time (1927) to The Task of Thinking (1964)*, ed. David Farrell Krell. New York: Harper & Row, 370-92.
- IFLA (International Federation of Library Associations and Institutions, Study Group on the Functional Requirements for Bibliographic Records). 1998. *Functional Requirements for Bibliographic Records*. München: K.G. Saur.
- Leazer, Gregory H. and Furner, Jonathan. 1999. "Topological Indices of Textual Identity Networks." In *Proceedings of the 62nd Annual Meeting of the American Society for Information Science, October 31, 1999*, ed. L. Woods. Medford, NJ: Information Today, 345-58.
- LeBoeuf, Patrick, ed. 2006. *FRBR: Hype or Cure-all?* Binghamton, NY: Haworth Information Press.
- Martinez-Ávila, Daniel, Richard P. Smiraglia, Hur-Li Lee and Melodie Fox. 2015. "What Is an Author Now? Discourse Analysis Applied to the Idea of an Author." *Journal of Documentation* 71: 1094-114.
- O'Neill, Edward and Maja Žumer. 2012. "Modeling Aggregates in FRBR." *Cataloging & Classification Quarterly* 50, nos. 5-7: 456-72. doi:10.1080/01639374.2012.679547
- O'Neill, Edward, Maja Žumer and Jeffrey Mixter. 2015. "FRBR Aggregates: Their Types and Frequency in Library Collections." *Library Resources & Technical Services* 59: 120-9.
- Panizzi, Antonio. (1848) 1985. "Mr. Panizzi to the Right Hon. the Earl of Ellesmere." In *Foundations of Cataloging*, ed. Michael Carpenter and Elaine Svenonius. Littleton, Colo.: Libraries Unlimited, 15-47.
- Park, Hyoungjoo and Richard P. Smiraglia. 2017. "Ontological Data-Sharing of Open Government Data for Data Curation." Hyoungjoo Park and Richard P. Smiraglia. *Canadian Journal for Information and Library Science* 41: 285-307.
- Petek, Marija. 2007. "Derivative Bibliographic Relationships in the Slovenian Online Catalogue COBIB." *Journal of Documentation* 63: 398-423.
- RDA (Resource Description and Access). 2010- Chicago: American Library Association; Ottawa: Canadian Library Association. In RDA Toolkit: <http://www.rda-toolkit.org>
- Smiraglia, Richard P. 1992. "Authority Control and the Extent of Derivative Bibliographic Relationships." Ph.D. diss., University of Chicago.
- Smiraglia, Richard P. 1996. "The Light in the Piazza: Glimpses of the Bibliographic Universe." In *Ventures in Research Series 17*, ed. Sheila McKenna. Brookville, NY: Faculty of the C.W. Post Center, Long Island University, 99-120.
- Smiraglia, Richard P. 2001. *The Nature of "A Work": Implications for the Organization of Knowledge*. Lanham, MD: Scarecrow Press.
- Smiraglia, Richard P., ed. 2002a. *Works as Entities for Information Retrieval*. New York: Haworth Press.
- Smiraglia, Richard P. 2002b. "Works as Signs, Symbols, and Canons: The Epistemology of the Work." *Knowledge Organization* 28: 192-202.
- Smiraglia, Richard P. 2003. "The History of 'The Work' in the Modern Catalog." *Cataloging & Classification Quarterly* 35 no. 2/4: 553-67.
- Smiraglia, Richard P. 2006. "Empiricism as the Basis for Metadata Categorization: Expanding the Case for Instantiation with Archival Documents." In *Knowledge Organization and the Global Learning Society, Proceedings of the 9th International ISKO Conference, Vienna, July 4-7, 2006*, ed. Gerhard Budin, Christian Swertz and Konstantin Mitgutsch. Advances in Knowledge Organization 10. Würzburg: Ergon Verlag, 383-88.
- Smiraglia, Richard P. 2007a. "Bibliographic Families and Superworks." In *Understanding FRBR: What it is and How it Will Affect our Retrieval Tools*, ed. Arlene G. Taylor. Libraries Unlimited, 73-86.
- Smiraglia, Richard P. 2007b. "The 'Works' Phenomenon and Best Selling Books." *Cataloging & Classification Quarterly* 44, no. 3/4: 179-95.
- Smiraglia, Richard P. 2007c. "Performance Works: Continuing to Comprehend Instantiation." In *Proceedings of the North American Symposium on Knowledge Organization, June 14-15 Toronto, Canada*, ed. Joseph T. Tennis. [http://dlist.sir.arizona.edu/view/conference/North\\_American\\_Symposium\\_on\\_Knowledge\\_Organization\\_2007.html](http://dlist.sir.arizona.edu/view/conference/North_American_Symposium_on_Knowledge_Organization_2007.html)
- Smiraglia, Richard P. 2007d. "When is a Terracotta Hut Urn like a Sailor's Deck-log? Meaning Instantiated Across Virtual Boundaries." Paper delivered at Museums & the Web 2007, April 11-14, San Francisco, California. <https://www.museumsandtheweb.com/mw2007/papers/smiraglia/smiraglia.html>
- Smiraglia, Richard P. 2008. "A Meta-analysis of Instantiation as a Phenomenon of Information Objects." *Culture del testo e del documento* 9 no. 25: 5-25.

- Smiraglia, Richard P. 2012. "Be Careful What You Wish for: FRBR, Some Lacunae, a Review." *Cataloging & Classification Quarterly* 50 no. 5/7: 360-68.
- Smiraglia, Richard P., ed. 2013. *The FRBR Family of Conceptual Models*, ed. Richard P. Smiraglia; co-ed. Pat Riva and Maja Zumer. London: Taylor and Francis.
- Smiraglia, Richard P. 2015. "Bibliocentrism Revisited: RDA and FRBRoo." *Knowledge Organization* 42: 296-301.
- Smiraglia, Richard P. 2017. "Data and Metadata Instantiation: Use Cases and a Conceptual Model." Paper read at DC-2017, DCMI International Conference on Dublin Core and Metadata Applications, October 26-29, Washington, D.C. <http://dcevents.dublincore.org/IntConf/dc-2017/paper/view/516>
- Smiraglia, Richard P. and Charles van den Heuvel. 2013. "Classifications and Concepts: Elementary Theory of Knowledge Organization." *Journal of Documentation* 69: 360-83.
- Smiraglia, Richard P. and Hur-li Lee. 2012. "Rethinking the Authorship Principle." *Library Trends* 61, no. 1: 35-48.
- Smiraglia, Richard P., Charles van den Heuvel and Thomas Dousa. 2011. "Interactions Between Elementary Structures in Universes of Knowledge." In *Classification & Ontology: Formal Approaches and Access to Knowledge: Proceedings of the International UDC Seminar 19-20 September 2011, The Hague, Netherlands*, ed. Aida Slavic and Edgardo Civallero. Würzburg: Ergon Verlag, 25-40.
- Smiraglia, Richard P., Hur-li Lee and Hope A. Olson. 2010. "The Flimsy Fabric of Authorship." In *Information Science: Synergy through Diversity, Proceedings of the 38th Annual CAIS/ACSI Conference, Concordia University, Montreal, Quebec, June 2 - 4 2010*, ed. Elaine Ménard and Valerie Nessel.
- Strout, Ruth French. 1956. "The Development of the Catalog and Cataloging Codes." *Library Quarterly* 26: 254-75.
- Svenonius, Elaine. 2001. *The Intellectual Foundation of Information Organization*. Cambridge: MIT Press.
- Talbot, Michael. 2000. *The Musical Work: Reality or Invention?* Liverpool Music Symposium 1. Liverpool: Liverpool University Press.
- Tanselle, G. Thomas. 1989. *A Rationale of Textual Criticism*. Philadelphia: University of Pennsylvania Press.
- Tillett, Barbara B. 2001. "Bibliographic Relationships." In *Relationships in the Organization of Knowledge*, ed. Carol A. Bean and Rebecca Green. Dordrecht: Kluwer, 19-36.
- Tillett, Barbara B. 2004. "What is FRBR?: A Conceptual Model for the Bibliographic Universe." Washington, DC: Library of Congress, Cataloging Distribution Service. <https://www.loc.gov/cds/downloads/FRBR.PDF>
- Vellucci, Sherry L. 1997. *Bibliographic Relationships in Music Catalogs*. Lanham, MD: Scarecrow Press.
- Verona, Eva. 1985. "Literary Unit Versus Bibliographical Unit." In *Foundations of Cataloging*, ed. Michael Carpenter and Elaine Svenonius. Littleton, CO: Libraries Unlimited, 155-75.
- Warner, Julian. 1993. "Writing and Literary Work in Copyright: A Binational and Historical Analysis." *Journal of the American Society for Information Science* 44: 307-21.
- Wilson, Patrick. (1968) 1978. *Two Kinds of Power: An Essay in Bibliographical Control*. California Library Reprint Series. Berkeley: University of California Press.
- Yee, Martha M. 1993. "Moving Image Works and Manifestations." PhD diss. University of California, Los Angeles.
- Žumer, Maja. 2018. "IFLA Library Reference Model IIFLA LRM)-Harmonisation of the FRBR Family." *Knowledge Organization* 45: 310-18. doi:10.5771/0943-7444-2018-4-310