User Conceptualizations of Derivative Relationships in the Bibliographic Universe

Kim Tallerås, Jørn Helge B. Dahl, Nils Pharo

OsloMet – Oslo Metropolitan University

Tallerås, K., Dahl, J. H. B., & Pharo, N. (2018). User conceptualizations of derivative relationships in the bibliographic universe. *Journal of Documentation*, 74(4), 894–916. https://doi.org/10.1108/JD-10-2017-0139

ABSTRACT

Purpose

Considerable effort is devoted to developing new models for organizing bibliographic metadata. However, such models have been repeatedly criticized for their lack of proper user testing. This paper presents a study on how non-experts in bibliographic systems map the bibliographic universe and, in particular, how they conceptualize relationships between independent but strongly related entities.

Methodology

The study is based on an open concept-mapping task performed to externalize the conceptualizations of 98 novice students. The conceptualizations of the resulting concept maps are identified and analyzed statistically.

Findings

The study shows that the participants' conceptualizations have great variety, differing in detail and granularity. These conceptualizations can be categorized into two main groups according to derivative relationships: those that apply a single-entity model directly relating document entities and those (the majority) that apply a multi-entity model relating documents through a high-level collocating node. These high-level nodes seem to be most adequately interpreted either as superwork devices collocating documents belonging to the same bibliographic family or as devices collocating documents belonging to a shared fictional world.

Value

The findings can guide the work to develop bibliographic standards. Based on the diversity of the conceptualizations, the findings also emphasize the need for more user testing of both conceptual models and the bibliographic end-user systems implementing those models.

Keywords

Bibliographic systems, User studies, Cataloguing, Linked data, Information modelling, Conceptualizations, Ontologies, FRBR, Mental models, Metadata

1. INTRODUCTION

At the time of writing, science fiction fans battle in heated debates over whether the new *Alien* and *Blade Runner* movies are part of the same fictional universe. The final season of the television series *Game of Thrones* has been launched well ahead of the remaining volumes in the book series that initially inspired it. Another anticipated television series is an adaptation of the book *Pride and Prejudice and Zombies*, which itself is a mash-up of Jane Austen's classic novel with the zombie craze in contemporary pop culture. Such entities seem to orbit each other in a *bibliographic universe*, "just as the physical universe reels with gravity and physical forces that propel, impel, and propel planets, stars, asteroids, and other bodies to exist in relation to each other" (the ideas of Wilson, 1968, as interpreted by Smiraglia, 2014, p. 10). In practice, these entities can cover the same topics or even transmedial storylines, share author and fictional characters, and belong to families of works related through various types of derivations.

When organizing bibliographic data in information systems, it is crucial to control the forces of the bibliographic universe in a way that increases the fitness for use. One particular challenge to controlling such a universe is the application of complex derivative relationships. This paper presents a study on how non-experts in bibliographic systems map the bibliographic universe and, in particular, how they *conceptualize* relationships between independent but strongly related works. The study is based on an open concept mapping task performed to externalize the bibliographic conceptualizations of novice students in library and information science. The resulting conceptualizations are analyzed statistically to reveal typical structures.

The paper has the following organization. Section 2 provides background information on bibliographic modelling and the research question, while section 3 describes the theoretical framework. Section 4 provides an overview of previous research, and section 5 presents the research methodology. Sections 6, 7 and 8 convey the results, discussion, and concluding remarks.

2. BACKGROUND

In bibliographic systems, relationships are indirectly applied based on descriptive metadata expressing *shared characteristics* (Tillett, 2001) about responsibility, topicality, and publishing events. Other bibliographic relationships, such as adaptations and non-trivial derivations, cannot be applied in as a straightforward way but are included as elements in existing bibliographic models. These include the *Library Reference Model* (LRM) (Riva *et al.,* 2017), the latest formalization of models belonging to the so-called FRBR-family[1]. The LRM includes the original FRBR entities for *works*, *expression* (of works), and *manifestations* (of expressions). Together, these *W/E/M* [2] entities enable representing both successive derivative relationships, such as new marginally changed editions (enforcing a new manifestation entity), and more significant modifications, such as a translation (enforcing a new expression entity). In addition, the LRM provides

[1] https://www.ifla.org/node/2016.

[2] The entity Item is also one of the so-called Group 1 entities in the original FRBR model and represents a concrete exemplar of a manifestation. However, in this study this physical level (item) is not considered.

derivative relationships directly between works, for example, in cases when one work has served as inspiration for another.

The LRM specifications state that the model was developed based on what are believed to be important entities and relationships for users of bibliographic systems. The users are represented through a set of specific user tasks (e.g., *to explore*), which should be facilitated by "the support of discovery by making relationships explicit, by providing contextual information and navigation functionality" (Riva *et al.,* 2017, p. 10). In models like the LRM, the included elements and, not least, their structural organization represent a *conceptualization* of the bibliographic universe, a simplified, abstract model of what exists in that particular universe of discourse. According to theories on mental models (Norman, 2013), users interacting with information systems depend heavily on their own conceptualizations when solving tasks. Thus, to facilitate, for example, the exploration task as defined by the LRM, the conceptualizations facilitating "contextual information and navigation functionality" (Riva *et al.,* 2017, p. 10) should reflect the conceptualizations of the users.

A repeated claim is that bibliographic models lack proper user testing (see, e.g., Coyle, 2016; Pisanski and Žumer, 2010a; Zhang and Salaba, 2009). The models typically reflect experts' accumulated ideas about important user tasks (for instance, the LRM builds on tasks that can be traced back to the bibliographic pioneer Charles Cutter (1904)). Pisanski and Zumer (2010a, 2010b, 2012) examined users' mental models of W/E/M entities but mostly evaluated the resemblance between mental models and W/E/M structures as they are mandated by the FRBR model. Although this approach has provided valuable insights into users' verification of that particular model, it could be beneficial to complement this research by testing users independent of an already-given structure. Another motivation for the present study is found in the bibliographic universe characterized by intertextuality and transmedia franchises generating immense numbers of complex derivative relationships, as exemplified in the introduction. Studies focusing on user conceptualizations in that particular context have not been found.

Thus, this paper is motivated by both the dearth of user testing in the domain of bibliographic modelling in general and the lack of knowledge on how users conceptualize derivative relationships in particular. These gaps lead to the following research question: how do users conceptualize derivative relationships between entities in the bibliographic universe?

3. THEORETICAL FRAMEWORK

The vision of the semantic web and the Linked data principles introduced to guide its realization have become the driving theoretical framework of recent developments in bibliographic metadata (van Hooland and Verborgh, 2014; Willer and Dunsire, 2013). This framework promotes interoperability through the establishment of a global network of metadata, facilitated by the use of standards (Berners-Lee, 2006; Hyland *et al.,* 2014). Such standards may be more or less technical and tailored to support the network structure, such as the Resource Description Framework (RDF), or they may be ontologies that reflect the conceptual structures of the entities and relationships constituting a particular domain. Although an RDF-like network is part of the present research design,

as described in section 5, the main concerns of this study are ontologies and their conceptual building blocks[3].

3.1 Conceptualizations

From an information science perspective, Gruber (1993, p. 199) defined an ontology as an "explicit specification of a conceptualization"[4]. Smith (2004, p. 161) elaborated on the implications of such conceptualizations: "As we engage with the world from day to day, we participate in rituals, and we tell stories. We use information systems, databases, specialized languages, and scientific instruments. […] Each of these ways of behaving involves, we can say, a certain conceptualization. What this means is that it involves a system of concepts in terms of which the corresponding universe of discourse is divided up into objects, processes, and relations in different sorts of ways. [...] Tools can be developed to specify and to clarify the concepts involved and to establish their logical structure".

This notion of conceptualizations corresponds to what is often referred to as *mental models* in human–computer interaction, usability, and other related research fields. Norman (2013, p. 25), a leading proponent of this tradition, defined mental models as "the conceptual models in people's minds that represent their understanding of how things work. […] People create mental models of themselves, others, the environment, and the things with which they interact". Theories on mental models derive from psychology, but since the 1940s, they have gradually been subjected to different interpretations in a number of fields (Westbrook, 2006). Like previous studies on users' internal models of bibliographic structures (e.g. Pisanski and Žumer, 2010a), this present study is based on Norman's (2013) perspective on mental models. This perspective is related to ontologies and principles underlying the development of modern bibliographic standards.

In this perspective, to improve usability for, say, data consumers who need to understand and use a particular Linked data set in a local system, it is imperative to model the Linked data set in a way that reflects the common conceptualizations shared among the potential data consumers. Ontologies can limit or enable the information architecture of end-user interfaces by providing rich, granular, simple, or shallow data structures. In such cases, ontologies should be based on an idea of how a generic user of those interfaces conceptualizes the entities in the given universe of discourse.

In bibliographic model development, designers often model users as stereotypes by defining user tasks or use cases[5]. These show the commitment of the development process to facilitating the user experience but often assume specific structures. Take, for example, a use case relevant to the research question of this study from the development

[3] Regarding the nuances between a conceptual model and a data model, see, for example, Coyle (2017).

[4] Studer *et al.* (1998) developed Gruber's (1993) ontology definition by stating that the conceptualizations should be *shared*.

[5] An example of the bibliographic extension for the schema.org vocabulary: https://www.w3.org/community/schemabibex/wiki/Use_Cases; an example of the Linked Data for Libraries model: https://wiki.duraspace.org/display/ld4l/LD4L+Use+Cases.

of the BIBFRAME model, defined as "Broadening Search (Discover Adaptations of Work)" (Library of Congress, 2013):

"Sue's research paper for Classics 201 must identify the themes in Homer's *Odyssey* as they relate to present day. She has a paperback copy of the book, but thinks that various adaptations of the Work (Movies, Plays, translations, etc.) would help with her research. In order to do this, she first finds the Work associated with the specific Instance she has in hand. From there, she explores the various relationships to other related Works."

The entity types *instance* and *work*, which represent a certain conceptualization of the bibliographic universe, are considered part of Sue's mental model. From the use case, it is clear that the implementation of these entities are preconditions to solve the task. The literature on bibliographic organization (see next section) discuss vast numbers of specific entities. However, this study examines mental models and users' conceptualizations as independent as possible from such constructs.

3.2 Derivative Constellations in the Bibliographic Universe

The universe of discourse examined in this paper is the bibliographic universe, defined as a concept space containing all recorded knowledge (Smiraglia, 2014, p. 10). Bibliographic entities refer to instances of that recorded knowledge (e.g., a novel, movie, or piece of music). Such instances can be grouped in multi-entity models reflecting their (dis)similarities (Baker *et al.,* 2014). This implies that in addition to the single entities representing a novel or a movie, there are more complex entities bringing their variations together. The mentioned W/E/M entities, for example, bring together different expressions or manifestations of a particular work. Svenonius (2000, p. 35) provided a similar grouping based on sets, including

- "The set of all documents sharing essentially the same information (work),
- The set of all documents sharing the same information (edition),
- The set of all documents descended from a common origin (superwork)" (Svenonius, 2000, p. 35)

Elaborating on superworks, Svenonius (2000, p. 38) explained that they *collocate* (a term adopted in the following analysis) works that are similar "by virtue of emanating from the same ur-work". As a concrete example of a *Hamlet* superwork, she listed collocated works such as the "original text, motion pictures, sound recordings of readings, analyses of the play, commentaries, playbills, derivative works like *Rosencrantz and Guildenstern Are Dead*" (Svenonius, 2000, p. 38). Svenonius (2000) also commented that a superwork can serve as an interesting tool for effective navigation. Smiraglia (2007) discussed *the bibliographic family*, a similar concept introduced by Wilson (1968). A bibliographic family collocates kindred works. The family structures are all "unique in the relationship the members bear to the originating work [...] yet distinct patterns occur among the members" (Smiraglia, 2007, p. 74). Smiraglia (2007) described such patterns as different types of derivative relationships that create a network of *instantiations*. An instantiation is "a concrete exemplar of a work as it has appeared at a specific point in the lifetime of the work" (Smiraglia, 2007, p. 83). Others, including Carlyle (1999) and Yee (1994), have also touched on the idea of a high-level collocating device. What all these approaches have in common is the shared premise of a specific starting point: the existence of a first

instantiation of a work that serves as the prime mover or the common ancestor of all the other works that form a family.

An instantiation is a generic term for different types of derivative relationships (Smiraglia, 2007), which were investigated extensively by both Smiraglia and Tillett in the 1990s. Tillett (1991) studied bibliographic relationships in general, including derivative relationships. Smiraglia and Leazer (1999) elaborated on Tillett's examples and definitions and listed seven types of common derivations:

- Simultaneous derivations
- Successive derivations
- Translations
- Amplifications
- Extractions
- Adaptations
- Performances

As mentioned in the background section, the W/E/M structure of the FRBR model encompasses some of these relationships. A successive derivation (e.g., a revised "second edition") can, if the intellectual or artistic content is unaffected, be represented by a new manifestation entity. Changes to the content result in a new expression or even a new work if "a significant degree of independent intellectual or artistic effort is involved" (Riva *et al.,* 2017, p. 20). A translation is widely understood as a new expression entity, whereas an adaptation is considered a new work. Other derivative relationships between works are defined with varying levels of granularity in FRBR-based models. For example, the RDA vocabulary [6] contains 14 specified sub-attributes representing various forms of adaptations, such as "is adapted as a motion picture" (P10085) and "is adapted as a television program" (P10085). The FRBRoo ontology, which harmonizes the original FRBR model with the museum-oriented CIDOC CRM model (LeBoeuf, 2012), includes a complex work concept (F15) that is quite similar to the notion of a superwork or bibliographic family. According to the FRBRoospecification it covers the notion that "The conceptual unity observed across a number of complete sets of signs, which makes it possible to organise publications into `bibliographic families.´ This is modelled as: F15 Complex Work is a F1 Work, and F15 Complex Work R10 has member (is member of) F1 Work" (Working Group on FRBR/CRM Dialogue, 2016, p. 26).

Vukadin (2014) points out that in addition to providing a practical means for collocating bibliographic entities in a superwork set, the FRBRoo F15 complex work concept can be used in cases when it is difficult to identify a common ancestor of the entities. This is common in so-called transmedia works that contain stories taking place in a shared fictional world but are often instantiated simultaneously across multiple media platforms. Such fictional worlds typically are developed through stories referencing the same characters, places, or events within or across authorships. In particular, they are studied in literary and media science as intertexts (in the tradition of Genette, 1997) or as transmedia storytelling (Jenkins, 2006).

[6] Resource Description Access (RDA) is the cataloging code developed to replace the existing AACR2 code. RDA has been described as a Linked data vocabulary (http://www.rdaregistry.info/).

4. PREVIOUS RESEARCH

As pointed out, the existing bibliographic models are based on theoretical constructs, not empirical data reflecting end users' understanding of bibliographic entities. Some researchers, though, have matched users' preferences with the FRBR model or tested interfaces for systems built upon the model. Yee (2005) evaluated the search facilities of four FRBRized catalogs and found that they were designed neither to take into account how existing bibliographic records can exploit the FRBR model nor to understand the original purpose of these records.

Carlyle and Becker (2008) conducted a survey asking if users would accept substitutes of FRBR manifestations, expressions, and related works when searching for known items. Their results showed that differences in manifestation types (e.g., a website or a printed copy) were as important as differences in expressions (e.g., different languages) when evaluating substitutability. Most surprisingly, Carlyle and Becker (2008) found that the participants accepted 30% of the related works suggested as substitutes.

Pisanski and Žumer (2010a, 2010b, 2012) compared how users' mental models of the bibliographic universes matched the W/E/M/I entities of the FRBR model. In their first study (Pisanski and Žumer, 2010a), 30 participants were given cards representing W/E/M/I entities of two books. In the first task, the participants were asked to sort the cards "into at least three groups based on the criterion of concrete/abstract (physical/non-physical) nature" and to name the groups (Pisanski and Žumer, 2010a, p. 649). Card co-occurrence was used to perform cluster analysis, which showed that no "constantly similar mental models" could be found (Pisanski and Žumer, 2010a, p. 655).

In the second task, the participants were asked to create a concept map describing the inter-relation of the cards and, specifically, "what comes out of what" (Pisanski and Žumer, 2010a, p. 655). The authors found that 14 of the 30 participants formed at least one work–expression–manifestation–item (four lengths) chain, and another 10 participants formed at least one chain of three lengths. Only two maps, however, corresponded exactly to the FRBR model.

In the third task (Pisanski and Žumer, 2010b), the participant were shown 11 pairs of items whose members differed in one W/E/M/I entity. The participants ranked the pairs according to their substitutability, and the analysis showed that the rankings matched the pairs' FRBR level. In other words, *items* were considered to be easily substitutable, whereas pairs that differed on the *work* level could not be substituted for one another (Pisanski and Žumer, 2010b).

Pisanski and Žumer (2012) followed up with a study in which the participants (120 students) were asked to select among six graphs representing potential relationships between W/E/M/I entities. The majority of the participants chose the graph representing the FRBR view, which indicates that it was the preferred way of coupling W/E/M/I entities.

A few user studies of library systems with FRBR-inspired interfaces have been conducted. Zhang and Salaba (2009) examined how users succeeded in performing different tasks in three FRBR-inspired catalogs. The users most successfully accomplished tasks that had the target of finding a work. The participants had problems with (in order of increasing difficulty) finding manifestations, identifying manifestations, and obtaining items (Zhang

and Salaba, 2009). Based on these findings, the authors developed a new prototype catalog, which they evaluated against a non-FRBRized catalog. Zhang and Salaba (2009) reported that 85% of the users preferred the FRBR prototype. Users, not surprisingly, performed tasks tailored toward works, expressions, and manifestations better in the FRBR system than the regular catalog.

Merčun and colleagues developed the FrbrVis prototype system (with FRBRized records) and assessed it against a traditional system (without FRBRized records) in two usability studies. In the first study (Merčun et al., 2016), 120 participants were asked to perform specific tasks interacting with bibliographic families representing different levels of complexity; in the second (Merčun *et al.,* 2017), they were free to explore the system. The controlled study found that the FrbrVis prototype performed better than the traditional system, both in general and when taking into account the complexity level of tasks.

To summarize, research investigating how users understand bibliographic universes have mostly used the FRBR model as their point of departure. Conceptually, users generally find different items, manifestations, and expressions of the same work to be substitutable and, to a certain degree, allow related works to be substituted for one another. When asked to map how different FRBR entities are related, users are less consistent but tend to prefer the FRBR model from among the alternatives presented. Some attempts to FRBRize existing records have been made, but evaluations indicate that these projects have been only partially successful. Promising FRBR prototype displays have been developed, and it will be interesting to see whether these can be implemented in future catalogs.

In contrast to previous research, the users in this study are not presented with existing solutions or bibliographic records but, rather, conceptualize bibliographic families based on their own understandings of what the documents' important characteristics are.

5 METHOD

5.1 Concept Mapping

Concept mapping serves as a method to reveal the bibliographic conceptualizations held by the participants in a study. The literature describes two forms of concept maps: hierarchical concept maps and network concept maps (Ruiz-Primo and Shavelson, 1996). Novak and Cañas (2006) deemed the hierarchical model to be not flexible enough for the purposes of studies such as the present one. Chang (2007, p. 107), who studied novice students' modelling of the homeostasis of blood sugar, concluded that the network concept map "is suitable for knowledge encompassing complex processes or interrelationships". As well, networks, or *graphs*, represent both flexible and sophisticated tools for organizing entities in a cultural heritage context (Murray and Tillett, 2012). Accordingly, the participants in this study were asked to draw a network representing how they view the documents, their essential attributes, and the (derivative) relationships between them. This method does not favor any hierarchical understandings of the bibliographic universe, leaving the participants free to draw any kinds of concepts and relationships.

The instructions for handling entity identification and organization in the concept mapping process may still affect the outcome. The task model should not limit the

elicitation of the conceptualizations by either its complexity or difficulty of application, but neither should it provide a means to represent the conceptualizations too abstractly. In this study, an RDF-like network was used as a guide for the concept maps. Since RDF is based on a graph model with named nodes and directed edgesit is a relatively intuitive and expressive guide for a concept mapping task.To provide the participants with a concrete guideline before they drew their concept maps, they were shown an example network conceptualizing an alternative universe of discourse (ships and persons related to the ships). They also received a short introduction explaining the task in detail. Finally, the participants were presented with a contextual purpose: their finalized conceptualizations should mirror what they believe would be a reasonable organization and selection of information for a general multiuser information system "like the ones used in libraries".

5.2 Participants

The study participants were all first-year bachelor students in library and information science. The task was given in the students' first lecture on bibliographic metadata. The participants' competency in this field of research was expected to be low and comparable to that of ordinary users of information systems. The participants also completed a post-task questionnaire, with questions on their gender, age, and previous experience with metadata, cataloguing, and programming related to their education, work, and hobbies. In addition, the participants could comment on the task in a text box.

5.3 Documents

Each participant was given three pieces of paper depicting three documents representing a book, a movie, and a music record. Before the main experiment, a pilot study was performed with five participants. The pilot testers performed the same tasks that we planned to use. Based on the pilot study some adjustments to the introduction were made. Apart from that, the study design remained unchanged. Documents from two different bibliographic families were used. Family *PG* contained:

1) The title page (recto and verso) of *Peer Gynt* by Henrik Ibsen, a Norwegian edition from 1962, published by Gyldendal
2) The DVD cover (front and back) of the 2006 television adaption of *Peer Gynt* directed by Bentein Baardson and produced by the Norwegian Broadcasting Corporation
3) The CD and the liner notes of *Music from the Mountains,* a collection of Peer Gynt suites composed by Edvard Grieg and Harald Sæverud, conducted by Ari Rasilainen, performed by the Norwegian Radio Orchestra and published by Finlandia Records in 1997

Family *RJ* contained:

1) The title page (recto and verso) of *Romeo and Juliet* by William Shakespeare translated to Norwegian by André Bjerke and published by Aschehoug in 2000
2) The DVD cover (front and back) of the 1996 movie *Romeo + Juliet* directed by Baz Luhrmann and published by Twentieth Century Fox

3) The CD and the CD cover (backside) of the *Romeo + Juliet* soundtrack published by Capitol Records in 1996

Documents that represent works belonging to large bibliographic families were selected, increasing the probability that the participants were familiar with them. The documents also represent typical entities that can be found in a library. In order to avoid constraining the tasks for the informants, no representations of the original manuscripts of the plays were included. Smiraglia and Leazer (1999) found that the size of a bibliographic family grows with the age of the progenitor work; therefore, universes that contain relatively old items were selected. *Peer Gynt* was written in 1867, while *Romeo and Juliet* was first printed in 1597. These two works have given rise to a great variety of creative inspirations and interpretations, so their accumulated bibliographic entities realized in a variety of media platforms have contributed to shared fictional worlds where stories unfold based on (or at least referencing) a set of given characters, places, and events.

The two families contained similar but not identical relationships between the documents. They both contained a play presenting a version of the original work. They also contained a movie and a musical record. In the PG family, the movie and the music represented independent adaptations. In the RJ family, the movie was an adaptation of the play, but the music contained already-published songs by different artists collected as a soundtrack for the movie. It was therefore less connected to the original play. The differences in the document families and relationships were incorporated into the research design to control for these variables in the experiment. In the following, *movie* represents the DVDs containing the movies, *book* represents the books containing the plays, and *music* represents the CDs containing the musical recordings.

6 ANALYSIS

6.1 Cluster Analysis

A total of 107 participants was recruited for the experiments. Their concept maps were interpreted and encoded by two researchers in two iterations. The first iteration provided an overview of the maps' common characteristics, such as the main nodes and the relationships between them. Eight concept maps could not be further analyzed due to a lack of identifiable or interpretable attributes. The remaining 99 concept maps were drawn according to the task instructions. They all contained a minimum of three nodes that could be identified as representations of the three documents from the handouts and the relationships connecting them directly or indirectly. The nodes were depicted as named circles or boxes, relationships as arrows or lines. Many relationships were named. Document nodes were identified as those being related to a minimum number of attributes, such as title, publisher, publication year or carrier/expression type (see Section 6.5 and 6.6 for details). In addition, indirect relationships between such document nodeswere often formalized through a *central node*, as in the example concept map shown in Figure 1.

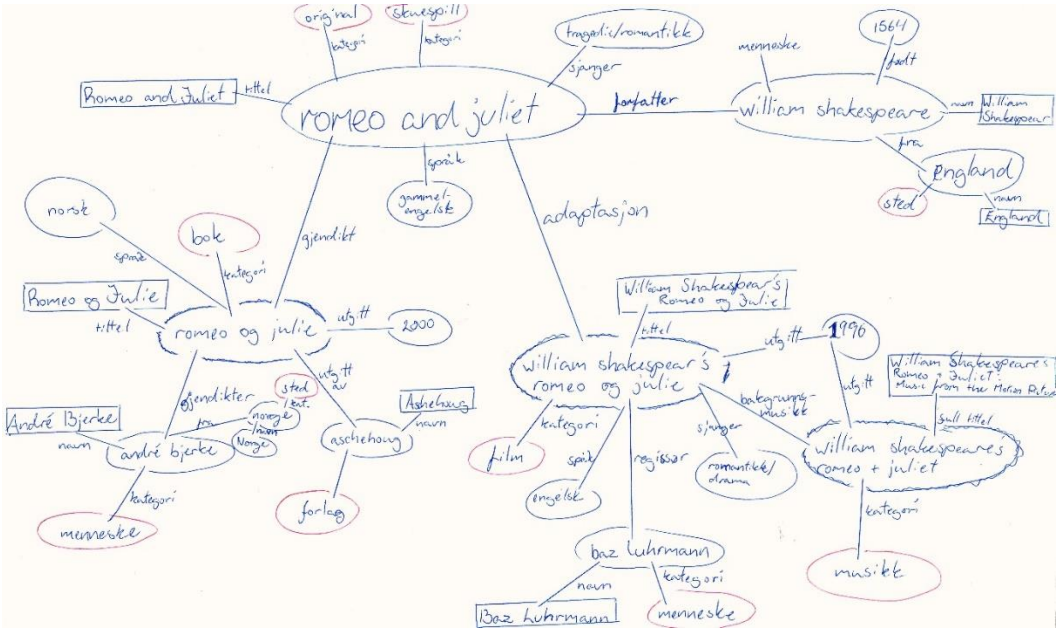
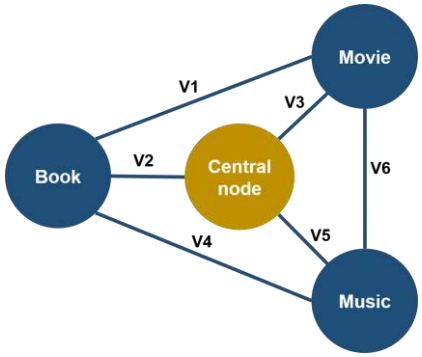


Figure 1 A representative concept map from the RJ family.

In the second iteration, the relationships between the document nodes were encoded in a spreadsheet as present or absent based on the criteria described above. To identify and group common relationship models, a hierarchical cluster analysis of six binary variables representing the identified relationships between the main nodes (Figure 2) was performed. In addition to the document nodes, the cluster analysis included the central node among the main nodes.

Cluster analysis offers a set of methods for grouping objects based on their characteristics and structures already present in data (Kaufman and Rousseeuw, 2009). Specific methods are chosen based on the types of variables (e.g., interval scaled, nominal, or binary). In order to perform the cluster analysis one need an operation to calculate the dissimilarities between objects and one to cluster the results. The well-known, simple matching coefficient (Sokal and Michener, 1958) was used for the (symmetric) binary data to develop a distance matrix, and the average linkage method (from the hclust package in R[7]) was utilized to build hierarchies.



[7] https://stat.ethz.ch/R-manual/R-devel/library/stats/html/hclust.html.

Figure 2 The main nodes identified, the concept maps, and the relationships between them treated as binary variables (present/not present) in the cluster analysis.

The results of the cluster analysis are visualized as a dendrogram, shown in Figure 3.

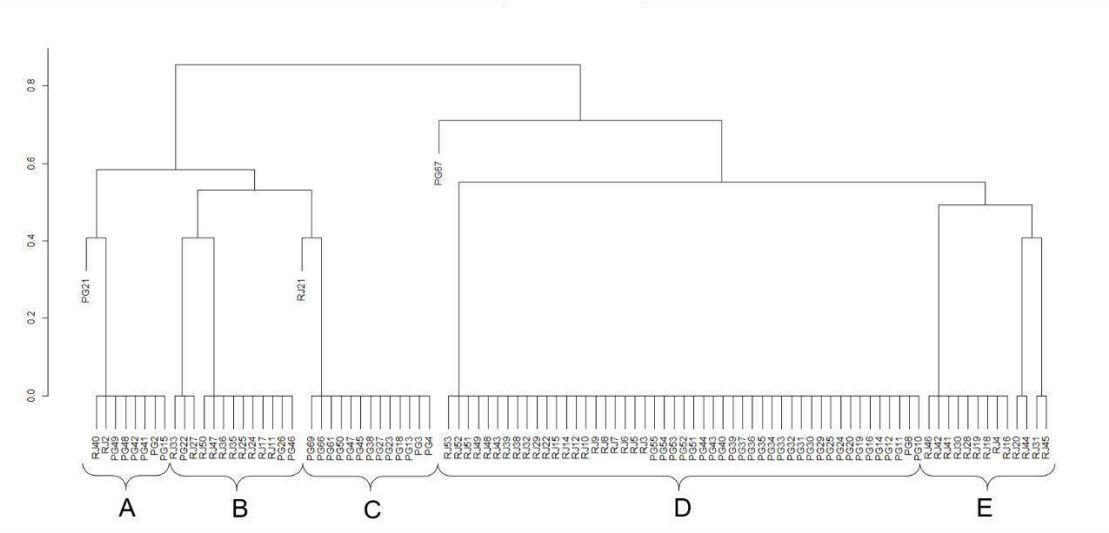


Figure 3 Dendrogram showing the results of the cluster analysis, with two main clusters and five sub-clusters identified.

The results of the cluster analysis reveal that the concept maps mostly belong to two clusters with five sub-clusters (A-E). Table 1 shows the distribution of the concept maps in each sub-cluster, the two main clusters, and each document universe. The most notable difference between the two main clusters is whether they include a central node for handling relationships. The 62 concept maps clustered in clusters D and E all include such a node; the universes belonging to the three other clusters (A, B, and C) do not. One concept map, placed between the C and D clusters in the dendrogram, is an outlier with a unique combination of relationships. In the following analysis, the attributes characterizing the five sub-clusters are examined. The outlier conceptualization is considered so atypical that it is removed from the statistics. Thus, 98 concept maps are included in the examinations. The analysis of the common properties in the various sub-clusters examines the directions of the relationships, primarily based on explicit naming (e.g., “adaptation of”, “version”, and “belongs to”) but also other expressed features indicating direction (e.g., arrows).

Cluster	PG family	RJ family	Total	Main clusters in %
A	7	2	9	37%
B	3	10	13	
C	13	1	14	
D	28	21	49	63%
E	0	13	13	
Total	51	47	98	100%

Table 1 Distribution of concept maps by cluster. Clusters A, B, and C contain central nodes; clusters D and E do not.

6.2 Clusters B and C: Document Nodes with Directed Relationships

Directly applying derivative relationships between document nodes, the concept maps in clusters B and C share an evident document orientation; the document nodes represent the essential entities in the bibliographic universe. Furthermore, the document nodes are related in sequences in which the book is interpreted as the originator and starting point. In cluster B (Figure 4), a chain of document nodes starts with a node representing the book and continues to the movie node and then to the music node. This cluster mostly contains concept maps of RJ documents (10 of 13). In cluster C (Figure 5), the book is also interpreted as an originator but is related to the other document nodes in two directions: from the book to the music and from the book to the movie. The movie and music nodes thus are not related in cluster C. Cluster C contains, with one exception, concept maps of PG documents.

These two clusters contain concept maps with essentially the same structure based on their shared document orientation. As well, their sequential aspects reflect an understanding of the book as the prime mover in the given bibliographic family. They differ in the direction and order of relationships, most likely due to actual differences between the two families. The music of the PG family was written drawing inspiration from the book, whereas the music of the RJ family is a collection of music created independently of the book. The participants who chose the "wrong" structure here (the four RJ participants with concept maps in cluster B) likely perceived the PG music as a soundtrack to the PG performance.

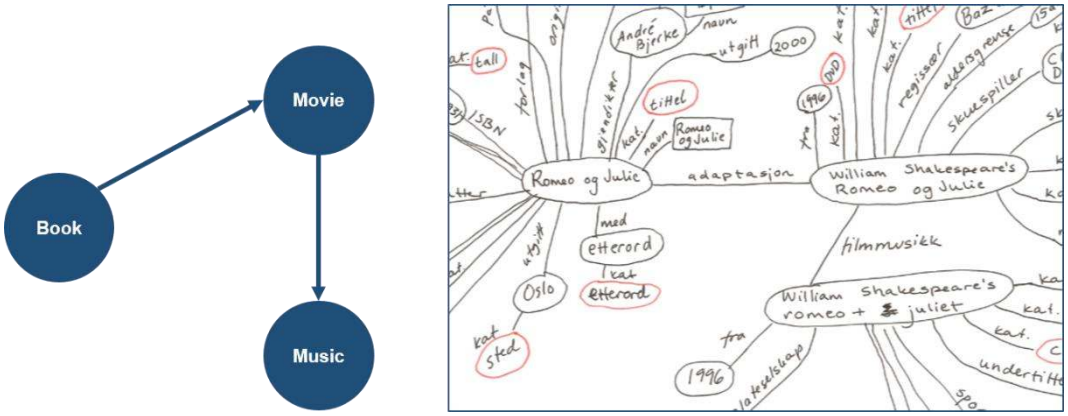


Figure 4 Cluster B with an example of a concept map. The document nodes are directly related from the book to the music via the movie.

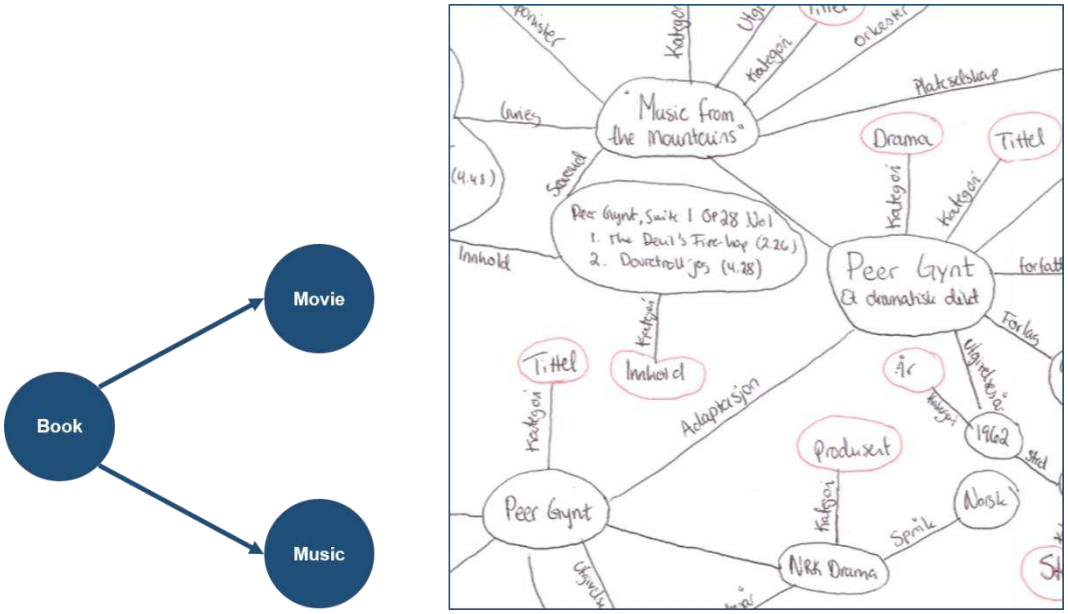


Figure 5 Cluster C with an example of a concept map. The document nodes are directly related from the book to the movie and the music.

6.3 Cluster A: Document Nodes with Relationships Based on Shared Characteristics

The document orientation of the concept maps in cluster A are similar to those in clusters B and C. What distinguishes the maps in cluster A is the lack of derivative relationships between the documents. The documents are instead linked indirectly via shared characteristics, such as authors, dates, genres, or topics. The relationships between the documents seem more arbitrary, as illustrated with dotted lines in Figure 6.

Seven concept maps in this cluster describe the PG family, while three describe the RJ family. No particular characteristics that can explain the skewed distribution are identified.

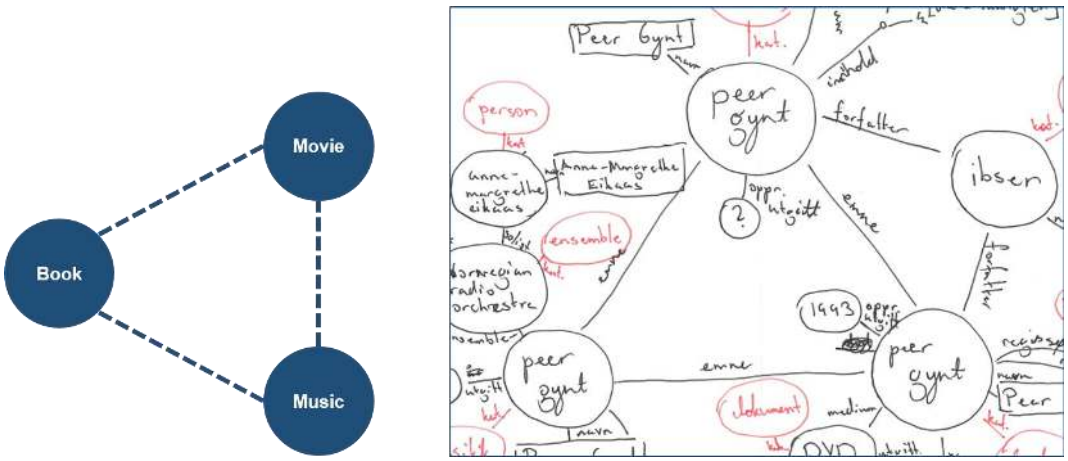


Figure 6 Cluster A with an example of a concept map. The document nodes are related via shared characteristics.

6.4 Clusters D and E: Document Nodes with Relationships Based on a Central Node

Clusters D and E comprise more than 63% of all concept maps. Of the 98 concept maps, 49 belong in cluster D alone. The concept maps in these clusters differ from those in the others due to their central nodes handling the relationships between the document nodes. As shown in Figure 7, the central node in cluster D handles all the relationships between the documents.

In the analysis, a central node was defined as constituting a separate semantic entity based on a single criterion: as a minimum, one attribute should be related to it. This criterion was set to differentiate between central nodes that actually represent the intended autonomous entities from the seemingly more arbitrary relationships applied in cluster A. The semantics of a central node, however, can be interpreted in different ways, as discussed in section 7. In cluster D, 28 concept maps of documents are from the PG family, and 21 from the RJ family.

As seen in Figure 8, the concept maps in cluster E, similar to the concept maps in cluster D, include a central node to handle relationships and have a direct relationship between the music node and the movie node. Cluster E comprises 13% of the concept maps, exclusively representing RJ documents. The relationship from the music to the movie in cluster E probably results from an interpretation of the RJ family documents similar to cluster B.

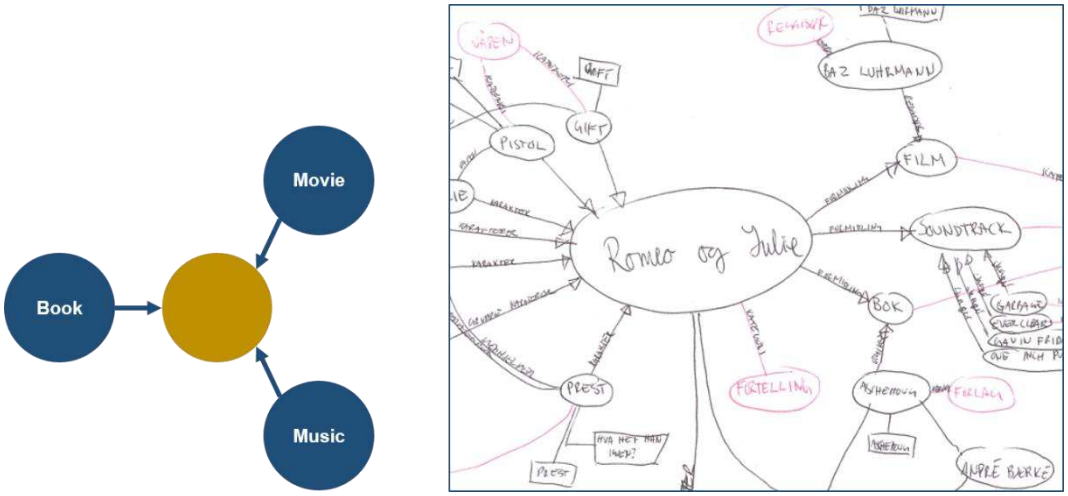


Figure 7 Cluster D with an example of a concept map. The document nodes are indirectly related via the central node.

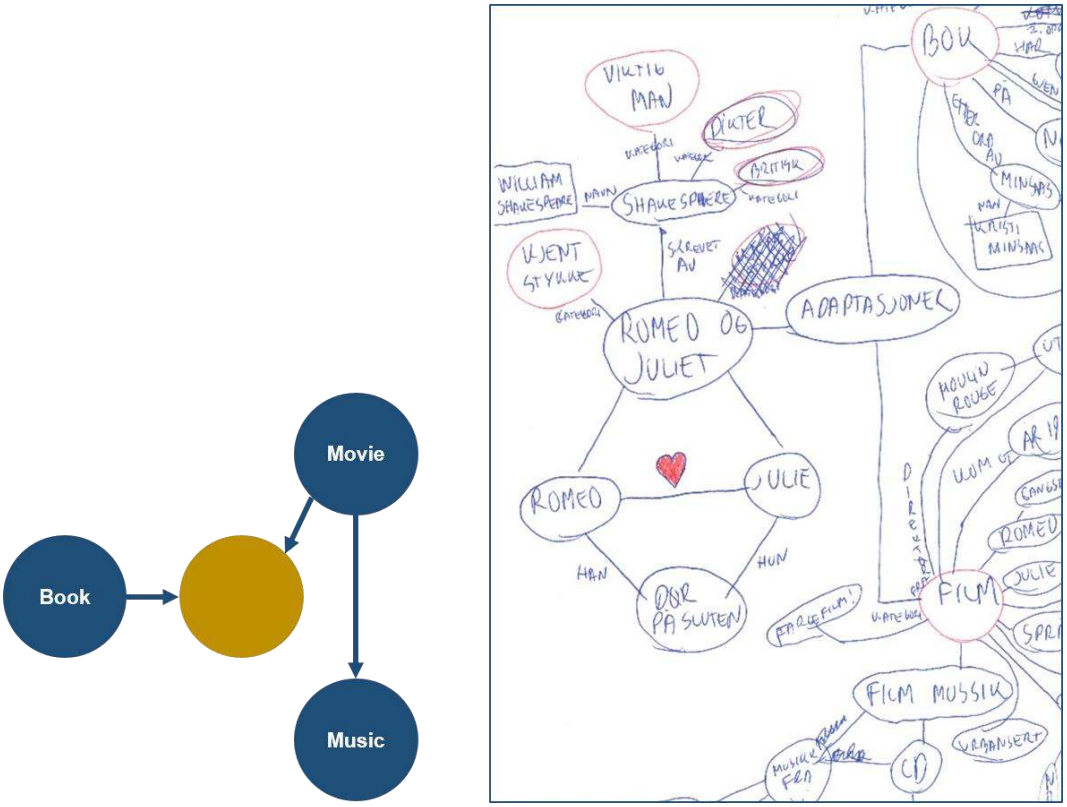


Figure 8 Cluster B with an example of a concept map. The document nodes are partly related indirectly via the central node (the book and the movie) and partly related directly (the movie and the music).

6.5 Attributes

Altogether, 72 different attributes, or descriptive characteristics of the documents, were identified. The concept maps each contained 18 attributes on average. The nature of the supplied material likely was a contributing factor to which attributes the informants included. For example, visually clear attributes (e.g., a publisher presented in a large font) were included in the concept maps more frequently than visually weaker ones. Although this study was more concerned with the overarching structures than the details of the attributes, the representations of the three most common attribute types were examined. Table 2 shows the distribution of the attributes across the clusters. Due to the different genres of the two music documents, the responsible composer for the music document in the PG family and the artists in the music document in the RJ family were included.

		A (n=9)	**B** (n=13)	**C** (n=14)	**D** (n=49)	**E** (n=13)	**All models**
Title	*Book*	89%	85%	79%	37%	46%	55%
	Movie	100%	92%	57%	41%	54%	57%
	Music	100%	77%	64%	59%	54%	66%
	Central node				100%	100%	63%
Responsibility	*Book (author)*	100%	100%	100%	41%	46%	64%
	Movie (director)	89%	77%	50%	55%	77%	63%
	Music (componist/artist)	78%	69%	79%	67%	38%	67%
	Central node (author)				69%	85%	46%
Date of publication	*Book*	78%	77%	71%	73%	85%	77%
	Movie	56%	92%	64%	69%	85%	73%
	Music	78%	54%	64%	76%	69%	71%
	Central node				4%	0%	2%

Table 2 Distribution of key attributes across the clusters.

Table 2 shows that many concept maps in clusters D and E have title and responsibility attributes directly linked to the central node; this is not the case for the date of publication. The date of publication is a typical manifestation attribute (in FRBR terminology), and the analysis reveals that the concept maps in the central-node clusters mostly attach these attributes to the document node.

Table 3 presents the distribution of different attribute types for the central node. Of the concept maps in clusters D and E, 73% have an author related to the central node, whereas 31% have a genre related to it. Only a few concept maps have a date of origin or an original language (the latter applies solely to concept maps in the RJ family) related to the central node. Of the concept maps, 40% have central nodes related to various fictional characters (e.g., "Mor Åse" and "Juliet"), while 15% have other attributes from the fictional world of the relevant documents, such as places ("Verona") and events ("The death of Romeo by poison").

	% of central node conceptualizations
Author	73%
Date of origin	7%
Original language	7%
Genre	31%
Related fictional characters	40%
Related fictional places or events	15%

Table 3 Attributes related to the central node.

6.6 Naming

Beyond a general request to make the nodes interpretable, the task instructions gave the participants no specific guidance on how to name the nodes. Examining the concept maps found that this creative freedom yielded additional insights into the conceptualizations.

The central nodes are exclusively named “Peer Gynt” or “Romeo and Juliet”. Such a naming practice was interpreted to indicate, or at least to originate from, a title. The naming of the document nodes is somewhat more complex. In addition to the use of document titles, two other sources of names are identified: carrier and expression types. The carrier category includes names that specify a carrier device, such as a CD or DVD. The expression category contains content or media types including names such as “text”, “music” and “video”. Table 4 shows the distribution of the naming categories across models and document types. For all document types and models, on average, 58% of the participants name their document nodes with a title, 34% a expression type, and 6% a carrier type. A closer look at the distribution across the different models reveals a dominant tendency: concept maps belonging to the central-node-only cluster (cluster D) include fewer titles and more carrier and expression types than the other clusters of concept maps. The concept maps in cluster A have the highest frequency of titles, while the other non-central-node-clusters (B and C) also include more titles than the central-node clusters.

Models	Title	Carrier	Expression
A (n=9)	82%	0%	19%
B (n=13)	77%	8%	10%
C (n=14)	64%	2%	33%
D (n=49)	20%	16%	57%
E (n=13)	46%	5%	49%
Average (n=98)	58%	6%	34%

Table 4 Types of the names of document nodes across models.

6.7 Participants

The results from the post-task questionnaire show no significant differences in the gender or average age of the participants creating the concept maps across the clusters. Overall, 25% of the participants reported that they had some prior experience with metadata, which seems to have influenced their conceptualizations. In cluster A, 60% of the participants reported that they had previous metadata experience, whereas only 17% of the participants with concept maps in cluster D did so. The other clusters had 20%–30% participants with prior experience, similar to the total average. An interesting possible explanation may be found in the cataloguing tradition of the Norwegian library sector, where the participants most likely gained their experience. In Norway, cataloguers are trained to catalogue documents according to standards (AACR2 and MARC) that mandate few relationships representing derivations between documents. This document orientation may have influenced the arbitrary relationships in the conceptualizations found in cluster A. Moreover, the central nodes found in cluster D concept maps created by participants with at least some experience are very different from the conceptualizations mandated by the current standards.

6.8 Main Findings

This study was intended to examine conceptualizations of derivative relationships. Cluster analysis of the relationships between the main nodes in the concept maps resulted in five clusters. Two clusters (D and E, representing 63% of the concept maps) include a central node used to relate all or some of the document nodes. In the other clusters, the

document nodes are related directly (clusters B and C) or indirectly via shared characteristics (cluster A). Cluster A thus represents a significant document orientation which does not include the derivative relationships between the documents.

Statistical analysis of the attributes and naming of the nodes confirmed the identified clusters. The concept maps without a central node tend to have titles as the names of document nodes, whereas clusters with a central node tend to use the names of document nodes to explicitly identify the type of expression or carrier the documents represent (e.g., “video” or “music”). Clusters with a central node tend to have persons of responsibility related to the central node but provide other attributes at the document level, such as the date of publication. Many concept maps belonging to the central-node clusters also relate to the central node information from the fictional world to which the documents belong, such as related fictional characters, places, and events.

If the concept maps are considered expressions of the participants’ conceptualizations, the findings suggest that the participants hold conceptualizations that:

- relate documents solely via shared characteristics (cluster A)
- relate documents directly (clusters B and C)
- relate documents through a central node (cluster D)
- combine a central node with direct relationships between the documents (cluster E)

Regarding bibliographic modelling, two different approaches to conceptualizing the entities and relationships of the bibliographic universe are identified. The document-oriented nodes and relationships in clusters A, B, and C can be generalized into a single-entity model, with the documents themselves at the center; the book is “a book”. Clusters D and E, in contrast, introduce a level of abstraction with their central nodes and indicate a multi-entity model; the book can be differentiated into several entities reflecting its meaning, expression, and physicality (Baker *et al.,* 2014).

Based on these groups of concept maps, a spectrum can be established (Figure 9), ranging from document-oriented conceptualizations constituting a single-entity model to conceptualizations with relationships handled by an entity representing an abstraction of the documents, constituting a multi-entity model.

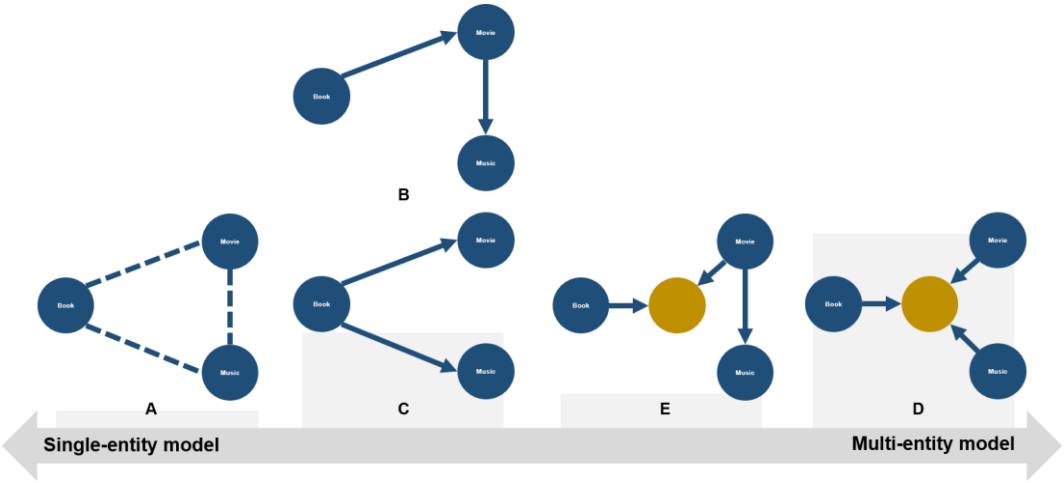


Figure 9 The five clusters along a spectrum from single-entity to multi-entity conceptualizations. The bars in the background indicate the amount of conceptualizations distributed across the spectrum

Conceptualizations with (single) entities that contain all the attributes of the documents describing both the formal characteristics and the aspects of their content and functionality can be found on the left side of the spectrum. “Nothing” appears to exist outside those entities. On the right side of the spectrum, conceptualizations with multiple entities differentiated by their varying views of the document attributes can be found. In these conceptualizations, the attributes appear to float more freely in a concept space that can be organized according to different views.

As explained in Section 5.3, slightly different document families were used to detect any influences on the conceptualizations. Some of the resulting clusters clearly relate to a particular document family and illustrate that the families, to some extent, do influence characteristics, such as the direction of relationships (clusters B and C) and the semantics of the central node (cluster E). The analysis, however, reveals that the variations are equally distributed among the clusters. Clusters B and C are nearly the same size and are considered to represent the same conceptualization. Cluster E, which is interpreted to represent an independent conceptualization, contains only RJ concept maps. The concept maps in cluster E handle the music node (the soundtrack of the movie adaption) in a way (excluding it from the central node entity) that would have been unlikely in the context of the PG family documents, where the music node represents an independent adaptation of the Ibsen play. It, therefore, is assumed that some participants in the PG group with concept maps in the D cluster would have made concept maps belonging to the E cluster if they were in the RJ group.

Considering the number of attributes included, relatively large differences exist between the families. Whereas, for instance, 87% of the participants include a genre in the PG group, only 32% include a genre in the RJ group. In the case of actors, the distribution is 48% and 72%, respectively. The differences are most likely due to the graphical presentations in the supplied document representations (e.g., font size and color) and the participants’ greater familiarity with some actors than others.

These figures show quite clearly that the document universes are interpreted differently when it comes to specific information but similarly in terms of the expressed high-level relationships between the document nodes.

7 DISCUSSION

As described in section 4, previous research has been concerned with verifying FRBR structures. Although the present study was designed to avoid an initial influence from particular bibliographic models, it is interesting to reflect on the findings in light of bibliographic entities, as they are outlined in section 3.2. In particular, the significant presence of a central node that signals a form of abstraction leads toward multi-entity structures like the W/E/M entities. In cluster E, two documents (the movie and the book on which the movie is based) are related through a collocating central node. The third document node (the music document not directly based on the book but strongly related to the movie) is related not to the central node but directly to the movie node. Hence, in the conceptualizations of cluster E, the central node organizes two cultural artifacts with strongly related content but excludes a third that is obviously related but also has quite different content. This recalls the logic behind W/E/M entities that conceptually collocates and separates varieties of cultural products based on similarities in content. Although cluster D does not directly contain ordinary W/E/M-entities, several of its factors are also reminiscent of FRBR logic. According to FRBR (e.g., in the form of LRM relationships between an agent and its W/E/M entities; Riva *et al.,* 2017, pp. 66–67), information about primary responsibility "for the creation of the intellectual or artistic content" is linked to the work entity, while the translators or the record company is linked to the expression or the manifestation. Information about responsibility for elements other than the content mostly is linked directly to the document nodes of cluster D. Thus, in the conceptualizations of cluster D, an FRBR-like chain of entities which includes almost half of the concept maps can be sensed.

Of the concept maps in clusters D and E, 73% have Shakespeare or Ibsen (the main persons of responsibility in the bibliographic families) related to the central node. However, it is quite apparent from the document representations that neither Ibsen nor Shakespeare has any responsibility for the content of the music node, which is part of the entity represented by a central node in cluster D. Thus, the central node cannot entirely be interpreted as an FRBR work. Some participants may have used the central node as a representation of the original work, although only a few participants included attributes like the original language and the original year of publication (see Section 6.5).

Perhaps it is also plausible to interpret the central node as an even more abstract collocating device. Instead of being responsible for all the documents attached to the central node, Ibsen and Shakespeare are linked to the central node because they are responsible for the originator works of the bibliographic families. From this perspective, the central node more resembles a superwork entity. This may also be the case for cluster E conceptualizations. A movie adaptation or a dramatization of a text is usually interpreted as a new work within FRBR, so the collocation of the book and the movie node is perhaps more accurately understood as a superwork function, collocating works belonging to a larger bibliographic family.

In addition, the study demonstrates that the central nodes are attached to much information belonging to an even more abstracted level: the fictional world related to the content of the documents. Thus, at least three types of conceptual abstractions can be drawn from the analysis of the central nodes:

- FRBR work
- Superwork
- Fictional world

The FRBR work has only a weak presence, though. Nevertheless, this study shows that for the purpose of expressing derivative relationships, users seem to prefer multi-entity conceptualizations including a superwork entity or characteristics of the relevant fictional world. It is also worth mentioning that 27% of the participants include direct derivative relationships between document nodes in their concept maps.

Another interesting question for this research field is whether any existing models reflect the present findings. Clarke (2015) claims that the current framework for bibliographic development based on semantic-web and Linked-data principles is progressing towards a new way of conceptualizing data. In this perspective, metadata are not necessarily exchanged as "units that include all the bibliographic information about a resource together in one place [...] like a MARC record", as in traditional cataloging, but also as limited individual statements (RDF-triples) "from multiple locations" (Clarke, 2015, p. 300). This requires bibliographic models that contribute to and enable flexible, complex semantics on different levels of abstraction. Several prominent libraries have attempted to facilitate such exchange and interoperability by publishing their bibliographies online as Linked data based on application profiles that include FRBR entities (Tallerås, 2017). Although the present findings show that some participants hold document-oriented conceptualizations reminiscent of traditional cataloging, the majority applies models dividing the universe into different levels of abstraction.

Emerging models, such as RDA, and Linked-data-based bibliographic models, such as BIBFRAME, also enable the modelling of FRBR works. The superwork level may be inferred from the explicitly expressed derivative relationships between works, but no models dedicate a conceptual entity at the superwork level, with the prominent exception of FRBRoo. Regarding the fictional-world level of abstraction, FRBRoo certainly can be used to collocate relevant works but does not provide any sophisticated semantics for expressing such relationships.

The identified conceptualization thus, to some extent, can be realized through existing models, but it appears that few utilize the available opportunities in current systems. Previous research has also mostly been concerned with the established FRBR levels, both in terms of user verification and the information architecture in user interfaces. Based on the present findings, more attention should be paid to the more high-level abstractions of superworks and fictional worlds.

8 LIMITATIONS AND FURTHER RESEARCH

Although the distinction between the two main clusters of drawings is evident, some of the five subclusters are rather small. Further research with more participants is therefore necessary to provide deeper insight.

The participants were not observed during or interviewed after the concept mapping task. The knowledge about their motivations and strategic decisions during development of their concept maps is therefore limited. An alternative research design would have been to observe and interview a smaller group of participants.

The two bibliographic families used in the study have similar origins, in that both stem from a play. Thus, they represent a specific selection of documents. It is necessary with additional studies including more and other types of documents to improve the understanding of derivative relationships in bibliographic universes in general. Other documents, such as a translated text presented together with a representation of the text in its original language, probably would provide more insight into user conceptualizations of bibliographic structures. To compare the results with those of previous FRBR-oriented research, documents with such relationships should be prominently incorporated in the design of future research.

The informants were all library and information science students, and therefore do not represent a general population. The students, however, were in their second week of the first semester in a bachelor degree and had no previous formal training in bibliographic metadata standards. One fourth of the students, on the other hand, did report some degree of previous experience with metadata (as discussed in section 6.7).

9 CONCLUDING REMARKS

A bibliographic universe comprises “the totality of things over which bibliographical control is or might be exercised” (Wilson, 1968, p. 6). Wilson (1968) further described two distinct kinds of control. Descriptive control provides the means, traditionally by cataloging, to create (arbitrary) lists that enable retrieval of all the entities characterized by certain attributes (“All plays by Ibsen”). Exploitative control, in contrast, is the ability to procure the best entities available serving a specific purpose. The first kind of power is evaluative neutral; the second involves appraisal by the user (Wilson, 1968, p. 22). According to Wilson (1968), exploitative control is more important, but descriptive control is a precondition for achieving exploitative control; to identify the best entities, these entities must be known, and to be known, they must be described. The same is true for the gravitational forces in the bibliographic universe manifested by relationships between entities. The present study shows that users conceptualize such relationships quite differently. Some utilize attributes that describe the shared characteristics of the documents—the traditional apparatus of descriptive control. Others directly relate documents by applying accurate derivative links. The majority of participants, however, applies a multi-entity model in which document entities are related through nodes at a higher level of abstraction describing the characteristics of a bibliographic family or a shared fictional world. Such information is essential to exercise exploitative control in an ever-expanding bibliographic universe containing the storylines of transmedia franchises and the derivative accumulations of popular bibliographic families.

Today, one perhaps could argue that the problem of descriptive control has more or less been solved, especially given the provision of digitized content enabling the automated generation of descriptions. Existing bibliographic ontologies provide means to describe complex derivative relationships. However, the ability to exploit these descriptions—to exercise exploitative control—is still an open problem and a holy grail for the world's leading libraries, search engines, and recommender systems.

REFERENCES

Baker, T., Coyle, K. and Petiya, S. (2014), "Multi-entity models of resource description in the Semantic Web", *Library Hi Tech*, Vol. 32 No. 4, pp. 562–582.

Berners-Lee, T. (2006), *Linked Data: Design Issues*, W3C, available at: http://www.w3.org/DesignIssues/LinkedData.html (accessed 7 November 2012).

Carlyle, A. (1999), "User categorisation of works: toward improved organisation of online catalogue displays", *Journal of Documentation*, Vol. 55 No. 2, pp. 184–208.

Carlyle, A. and Becker, S.R. (2008), "ASIS&T 2008 annual meeting poster 'FRBR and the "known-item" search.'", *Proceeding of the American Society for Information Science and Technology*, Vol. 45 No. 1, pp. 1–9.

Chang, S.-N. (2007), "Externalising students' mental models through concept maps", *Journal of Biological Education*, Vol. 41 No. 3, pp. 107–112.

Clarke, R. (2015), "Breaking records: The history of bibliographic records and their influence in conceptualizing bibliographic data", *Cataloging & Classification Quarterly*, Vol. 53 No. 3–4, pp. 286–302.

Coyle, K. (2016), *FRBR, before and after: a look at our bibliographic models*, American Library Association, Chicago.

Coyle, K. (2017), "Two FRBRs, many relationships", *Coyle's InFormation*, available at: http://kcoyle.blogspot.no/2017/05/two-frbrs-many-relationships.html (accessed 5 October 2017).

Cutter, C. (1904), "Rules for a dictionary catalog", *The US Bureau of Education Special Report on Public Libraries-Part II*.

Genette, G. (1997), *Paratexts: Thresholds of interpretation*, Cambridge University Press, Cambridge.

Gruber, T.R. (1993), "A translation approach to portable ontology specifications", *Knowledge Acquisition*, Vol. 5 No. 2, pp. 199–220.

van Hooland, S. and Verborgh, R. (2014), *Linked Data for Libraries, Archives and Museums: How to clean, link and publish your metadata*, Facet publishing, London.

Hyland, B., Atemezing, G. and Villazón-Terrazas, B. (2014), "Best Practices for Publishing Linked Data", *W3C Working Group Note*, available at: https://www.w3.org/TR/ld-bp/ (accessed 5 October 2017).

Jenkins, H. (2006), *Convergence culture : where old and new media collide*, New York University Press, New York.

Kaufman, L. and Rousseeuw, P. (2009), *Finding groups in data: an introduction to cluster analysis*, Wiley, Hoboken, NJ.

LeBoeuf, P. (2012), "A strange model named FRBRoo", *Cataloging & Classification Quarterly*, Vol. 50 No. 5–7, pp. 422–438.

Library of Congress. (2013), "BIBFRAME Use Cases and Requirements", available at: http://bibframe.org/documentation/bibframe-usecases/ (accessed 22 September 2017).

Merčun, T., Žumer, M. and Aalberg, T. (2016), "Presenting bibliographic families", *Journal of Documentation*, Vol. 72 No. 3, pp. 490–526.

Merčun, T., Žumer, M. and Aalberg, T. (2017), "Presenting bibliographic families using information visualization: Evaluation of FRBR-based prototype and hierarchical visualizations", *Journal of the Association for Information Science and Technology*, Vol. 68 No. 2, pp. 392–411.

Murray, R.J. and Tillett, B.B. (2012), "Cataloging theory in search of graph theory and other ivory towers", *Information Technology and Libraries*, Vol. 30 No. 4, pp. 170–184.

Norman, D. (2013), *The design of everyday things: Revised and expanded edition*, Basic Books, New York, NY.

Novak, J.D. and Cañas, A.J. (2006), "The origins of the concept mapping tool and the continuing evolution of the tool", *Information Visualization*, Vol. 5 No. 3, pp. 175–184.

Pisanski, J. and Žumer, M. (2010a), "Mental models of the bibliographic universe. Part 1: mental models of descriptions", *Journal of Documentation*, Vol. 66 No. 5, pp. 643–667.

Pisanski, J. and Žumer, M. (2010b), "Mental models of the bibliographic universe. Part 2: comparison task and conclusions", *Journal of Documentation*, Vol. 66 No. 5, pp. 668–680.

Pisanski, J. and Žumer, M. (2012), "User verification of the FRBR conceptual model", *Journal of Documentation*, Vol. 68 No. 4, pp. 582–592.

Riva, P., Bœuf, P. Le and Žumer, M. (2017), "IFLA Library Reference Model", The International Federation of Library Associations and Institutions, Den Haag, available at: https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla_lrm_2017-03.pdf.

Ruiz-Primo, M. and Shavelson, R. (1996), "Problems and issues in the use of concept maps in science assessment", *Journal of Research in Science*, Vol. 33 No. 6, pp. 569–600.

Smiraglia, R.P. (2007), "Bibliographic families and super works", in Taylor, A.G. (Ed.), *Understanding FRBR: What It Is and How It Will Affect Our Retrieval Tools*, Libraries Unlimited, Westport, CT, pp. 73–86.

Smiraglia, R.P. (2014), *The elements of knowledge organization*, Springer, Cham.

Smiraglia, R.P. and Leazer, G.H. (1999), "Derivative Bibliographic Relationships: The Work Relationship in a Global Bibliographic Database.", *Journal of the American Society for Information Science*, Vol. 50 No. 6, pp. 493–504.

Smith, B. (2004), “Ontology”, in Floridi, L. (Ed.), *The Blackwell Guide to the Philosophy of Computing and Information*, Blackwell publishing, MA.

Sokal, R.R. and Michener, C.D. (1958), “A statistical method for evaluating systematic relationships”, *University of Kansas Science Bulletin*, available at:https://doi.org/citeulike-article-id:1327877.

Svenonius, E. (2000), *The Intellectual Foundation of Information Organization*, MIT Press, Cambridge, Massachusetts.

Tallerås, K. (2017), “Quality of linked bibliographic data: the models, vocabularies, and links of data sets published by four national libraries”, *Journal of Library Metadata*, Vol. 17 No. 2, pp. 126–155.

Tillett, B. (1991), “A taxonomy of bibliographic relationships”, *Library Resources and Technical Services*, Vol. 35, pp. 150–158.

Tillett, B.B. (2001), “Bibliographic relationships”, in Bean, C.A. and Green, R. (Eds.), *Relationships in the Organization of Knowledge*, Springer Netherlands, pp. 19–35.

Vukadin, A. (2014), “Bits and pieces of information: bibliographic modeling of transmedia”, *Cataloging & Classification Quarterly*, Vol. 52 No. 3, pp. 285–302.

Westbrook, L. (2006), “Mental models: a theoretical overview and preliminary study”, *Journal of Information Science*, Vol. 32 No. 6, pp. 563–579.

Willer, M. and Dunsire, G. (2013), *Bibliographic information organization in the semantic web*, Chandos Publishing, Oxford.

Wilson, P. (1968), *Two kinds of power: An essay on bibliographical control*, University of California Press, Berkeley, CA.

Working Group on FRBR/CRM Dialogue. (2016), “Definition of FRBRoo: A conceptual model for bibliographic information in object-oriented formalism”, The International Federation of Library Associations and Institutions, Den Haag, available at: https://www.ifla.org/files/assets/cataloguing/FRBRoo/frbroo_v_2.4.pdf (accessed 5 October 2017).

Yee, M.M. (1994), “The Concept of Work for Moving Image Materials”, *Cataloging & Classification Quarterly*, Vol. 18 No. 2, pp. 33–40.

Yee, M.M. (2005), “FRBRization: a method for turning online public finding lists into online public catalogs”, *Information Technology and Libraries*, Vol. 24 No. 3, pp. 77–95.

Zhang, Y. and Salaba, A. (2009), “What Is next for functional requirements for bibliographic records? A delphi study”, *The Library Quarterly*, Vol. 79 No. 2, pp. 233–255.