

Postprint version. Please refer to the publisher's version as follows:

Zapounidou, S., Sfakakis, M. and Papatheodorou, C. (2017). Representing and integrating bibliographic information into the Semantic Web: A comparison of four conceptual models. *Journal of Information Science*, 43(4), pp.525-553.

DOI: [10.1177/0165551516650410](https://doi.org/10.1177/0165551516650410)

Representing and integrating bibliographic information into the Semantic Web: A comparison of four conceptual models

Sofia Zapounidou, Michalis Sfakakis, Christos Papatheodorou
Department of Archives, Library Science and Museology, Ionian University, Corfu, Greece

Abstract

Integration of library data into the Semantic Web environment is a key issue for libraries and is approached on the basis of interoperability between conceptual models. Several data models exist for the representation and publication of library data in the Semantic Web and therefore inter-domain and intra-domain interoperability issues emerge as a growing number of web data are generated. Achieving interoperability for different representations of the same or related entities between the library and other cultural heritage institutions shall enhance rich bibliographic data reusability and support the development of new data-driven information services. This paper aims to investigate common ground and convergences between four conceptual models, namely Functional Requirements for Bibliographic Records (FRBR), FRBR Object-Oriented (FRBRoo), Bibliographic Framework (BIBFRAME) and Europeana Data Model (EDM), enabling semantically-richer interoperability by studying the representation of monographs, as well as of content relationships (derivative and equivalent bibliographic relationships) and of whole-part relationships between them.

Keywords

Bibliographic Framework (BIBFRAME); conceptual models; Europeana Data Model (EDM); Functional Requirements for Bibliographic Records (FRBR); FRBR Object-Oriented (FRBRoo); interoperability; Linked Data; Semantic Web

1. Introduction

Libraries and other memory institutions, such as museums and archives, anticipate to enhance the visibility of their collections and to augment their impact as major contributors in research, teaching and learning, by actively participating in the explosive Semantic Web universe. In this universe, all supplied information and resources have to follow well-defined structures, widely accepted publication principles and be deeply interpretable and accurate in semantic level, especially during these times where large aggregation portals are under expansion (Europeana, Digital Public Library of America [DPLA]). An increasing number of initiatives are considering the exposure of library and other cultural heritage resources into Semantic Web. Depending on the initiative's objectives, the scope and the intended use of the resources, each initiative developed its own interpretation of how its resources may be integrated into the Semantic Web, providing its own conceptual model. In the library domain, the most known of them are Functional Requirements for Bibliographic

Records (FRBR) [1], FRBR Object-Oriented (FRBRoo) [2] and Bibliographic Framework (BIBFRAME) [3]. However, all these different views may cause interoperability problems and prevent data integration.

Furthermore, domain-specific as well as national and international aggregation services have been developed that collect rich descriptions about cultural heritage objects with the aim to provide advanced research support services. The most well-known are the European aggregation portal of Europeana (<http://www.europeana.eu/>) and DPLA (<http://dp.la/>). Both Europeana and DPLA have developed data models, namely the Europeana Data model (EDM) [4] and the DPLA Metadata Application Profile (DPLA MAP) [5] – to enable proper harvesting and integration of metadata from a variety of data providers. Europeana is a very active aggregator and has developed recommendations for aligning library data to EDM. While EDM is not built on any particular community standard, it adopts Semantic Web representation principles and reuses other existing vocabularies, such as OAI-ORE [6], Dublin Core [7] and SKOS [8], and is also inspired from the CIDOC–CRM [9], a well-known and established model in the cultural heritage domain. CIDOC–CRM is an event-centric ontology expressed as an object-oriented semantic model that describes ‘concepts and relationships used in cultural heritage documentation’ [9].

Among the library models, FRBR has been extensively studied due to its early development in the library domain. Many reliable tools have been developed for extracting FRBR entities from library catalogue records and used in various contexts [10–14]. Moreover, some comparative studies exist that try to investigate the expressiveness of the mentioned models as an effort to provide mappings between them and to tackle interoperability problems. Chen and Ke [15] test FRBRoo as a shared ontology for the integration of heterogeneous metadata used in digital libraries and museums settings. They provide a mapping between Dublin Core elements from the Taiwan E-Learning and Digital Archives Program Union Catalog to FRBRoo using a path-oriented approach. Chen and Ke’s study [15] focuses on museum artifacts and literary works. Doerr et al. [16] study the expression of FRBR semantics through EDM using FRBRoo in the framework of EDM–FRBRoo Application Profile Task Force. This task force attempted to create a library application profile for Europeana using as a test case three indicative data types: monographs, plays and musical works. They also investigate how different modelling patterns could affect the mappings between these models. Zapounidou et al. [17] compare four models testing their expressiveness regarding monographs. The models studied were FRBR, FRBRoo, BIBFRAME and EDM, while the test case was an English translation of the two parts of Cervantes’ *Don Quixote* bound together in a single volume.

It is worth mentioning that, in the bibliographic universe, the same intellectual content could be commonly realised in different expressions, which may be embodied with different formats into different media and may also be published from various publication procedures. Moreover, new intellectual content could be derived from existing content, as well as different intellectual contents could be compiled or aggregated in order to produce new content. Hence the ability of the models to express this complexity in accordance to their community’s intended functional requirements is the matter of question. While conceptual models are tools used in a specific context, at the same time they should be interoperable to enable data-exchange and reuse [18].

This study, inspired from our previous work [17], covers a wider range of monograph types in different formats and materialisations, as well as alternative modelling patterns offered by the models. More specifically, the expressiveness, the common ground and the divergences of the four mentioned models (FRBR, FRBRoo, BIBFRAME and EDM) are explored and compared by investigating resources covering a wide range of representation categories. Multipart monographs, in single parts or aggregated publications, their translations, adaptations and other derivations in various formats and media were the categories considered by this study. As it is presented in the next section, three studies [19–21] estimated that the class of these monograph types is represented in great numbers in WorldCat. Regarding the alternative modelling patterns, their existence is investigated, as well as their influence in the expressiveness and interoperability of the models. In the case of multiple representations, it is also investigated whether using specific patterns in each model proves to be more expressive or interoperable between the models.

The well-known literary work *Don Quixote* was selected to serve as a representative case for different categories of monographs and to enable identification and study of monograph representation issues.

In the next sections of the paper, the approach and short descriptions of the studied models will be presented. Representations of single-volume monographs, as well as of content relationships (derivative and equivalent bibliographic relationships) and of whole-part relationships between monographs using each model’s semantics and modelling patterns will follow. The paper concludes with a discussion regarding similarities and differences between the studied models based on the representations findings.

2. Approach

Three library data models, namely FRBR, FRBRoo and BIBFRAME, as well as the cultural heritage data model of Europeana (EDM), are compared. The case of EDM was selected in order to scrutinise common ground between library

data models and that of Europeana in the framework of sharing library data with third-party services and systems. More information about each model is given in the next section.

Our study focuses on monographs since the majority of library holdings described through individual records are monographs. Many published works are, of course, individual papers, e.g. as parts of conference proceedings; in this case, these collections of conference papers are considered as component parts of monographs. Two studies using WorldCat data performed in 2002 [19] and in 2011 [20] provide valuable information about categories of monographs in terms of percentages and absolute numbers. According to Bennett et al. [19], 78% of all works in WorldCat have only one manifestation, 16% only a single expression but multiple manifestations and a 6% are works that have multiple expressions (revisions, translations, aggregates, augmentations). Žumer and O’Neill [20] performed a more recent estimate regarding aggregates, according to which there are at least 30 million aggregates in WorldCat.

For the purpose of investigating how specific categories of monographs, such as single-volume monographs, multipart monographs, collections, integrating resources and other aggregates, as well as translations and other variations, may be represented in specific conceptual models, *Don Quixote* was selected. *Don Quixote* consists of two separate works: the first one was published in 1605 and the second one was published in 1615. These two parts have been both published and translated afterwards as independent volumes, as well as in a single volume. Moreover, there are many reproductions of materials other than the original publications, as well as other works based on many variations of the original work. Therefore, the case of *Don Quixote* may serve as a representative case for the millions of WorldCat records providing the ability to build more complex representations and to assess each model’s expressiveness. *Don Quixote* was also used by the EDM–FRBRoo Application profile task force [16] in the context of creating an application profile. In our study we use five MARC21 records from the National Library of Spain – three of them (records 1, 2 and 5) are the same with the ones used by the EDM–FRBRoo Application profile task force [16]. It is worth mentioning that in the following text, *Don Quixote* refers to the whole set of concepts appeared during the coherent evolution of the initial ideas expressed in the 1605’s first publication of the *El ingenioso hidalgo don Quixote de la Mancha*. Our set of bibliographic records is used to represent various cases of monograph types in four different models; alternative representations enabled by each model’s semantics are also studied. Each selected record may be considered as a representative case for one or more monograph categories. The selected records describe:

- (1) The first part of *Don Quixote* published in 1605 with the title *El ingenioso hidalgo don Quixote de la Mancha*. This title has been later used in editions publishing both parts of *Don Quixote*. In this study we use the original titles that Cervantes used. This part is denoted as ‘First Part’ in the fourth section’s representations. This First Part is a representative case of single-volume monographs (hereafter referred as single monographs), parts of multipart monographs and digitised monographs.
- (2) The second part of *Don Quixote* published in 1615 with the title *Segunda parte del ingenioso cauallero don Quixote de la Mancha*. This part is denoted as ‘Second Part’ in the fourth section’s representations. This Second Part is a representative case of single monographs, parts of multipart monographs and digitised monographs.
- (3) A single volume annotated edition containing both parts published by the Real Academia Española and the Asociación de Academias de la Lengua Española. This edition, entitled *Don Quijote de la Mancha* (denoted as ‘Annotated’ in the fourth section’s representations), was published in 2004 on the occasion of the 400-year anniversary of the first edition of the First Part in 1605. This Annotated edition, whose editor was Francisco Rico, was augmented with other significant works written by notable authors, such as Mario Vargas Llosa, Francisco Ayala, Martín de Riquer and José Manuel Blecuá. These works are mostly studies on Cervantes and his work, as well as on the Quixote character. To determine the type of aggregate, as well as the sufficiently significant distinct content to form works on their own, we consulted other records of the same edition from WorldCat, Library of Congress and University of Southern California Libraries. This Annotated edition is a representative case of aggregates. It must be noted that for readability reasons, not all works incorporated in this Annotated edition are represented in the fourth section’s representations. The works considered as the most significant for representation are *Una novela para el siglo XXI* (denoted as ‘Novela’) by Mario Vargas Llosa, *La invención del Quijote* (denoted as ‘Invención’) by Francisco Ayala and *La lengua de Cervantes y el ‘Quijote’* (denoted as ‘Lengua’) by José Manuel Blecuá. The works by Llosa and Ayala review the story and the character of Don Quixote, while Blecuá’s work focuses on the language Cervantes used in his works and in *Don Quixote* in particular.
- (4) A French translation of both parts (denoted as ‘French translation’ in the fourth section’s representations). The French title of both parts of *Don Quixote* was *Histoire de l’admirable don Quichotte de la Manche*, the translator was Filleau de Saint-Martin. The French translation was published in 1677. Filleau de Saint-Martin changed the end of Cervantes’ work where Don Quixote dies. Then he wrote a whole new work with the continuing story

of Don Quixote. This new version was a very popular one and was translated into many languages [22]. Even though much information can be found regarding how the French translator altered the end of the original *Don Quixote*, no information was found regarding the exact edition/variation of *Don Quixote* he used as a basis for his free translation. Thus this free translation is considered as a new Work that derives from the original *Don Quixote*. This French translation is a representative case of derivations (adaptation) of works. This work has also been digitised and is available through the Biblioteca Digital Hispánica. Therefore, it is also considered as a representative case of digitised monographs.

- (5) An English translation of both parts by John Phillips (denoted as ‘English translation’ in the fourth section’s representations). John Phillips based his translation *The history of the most renowned Don Quixote of Mancha* on Filleau de Saint-Martin’s French text. His translation is considered by experts as a bad one and even Putnam, another translator of *Don Quixote* in English, characterised John Phillips’ translation as the worst ever made [23]. This English translation is closer to the French text and is a representative case of translations. This work has also been digitised and is available through the Biblioteca Digital Hispánica. Therefore, it is also considered as a representative case of digitised monographs.

For each case, graph representations are generated for all models using each model’s semantics, classes and properties. The nodes of a graph denote the corresponding classes, while the edges denote the relationships between the classes of each model. In particular, the upper box of a node denotes the class of the respective model, while the lower box provides its instance. For clarity reasons a small description is used instead of each object’s full title and/or level’s related details. This small description is followed by a letter signifying the class, e.g. First Part_W (First Part as FRBR *Work* class instance) and First Part_M (First part as FRBR *Manifestation* class instance). The names of models’ classes/entities and properties are written in the text in italics, e.g. *Work* represents the FRBR *Work* entity, while the concept of work is written as normal text.

3. Conceptual models

3.1. Functional Requirements for Bibliographic Records (FRBR)

The FRBR conceptual model has been developed by the International Federation of Library Associations and Institutions (IFLA) [1] and it has been considered as a significant contribution in the library domain, often used as a point of reference by other library data modelling efforts. It is an entity-relationship model that defines key bibliographic entities, their attributes and the relationships between them. Key bibliographic entities are grouped into three groups: Group 1 entities refer to intellectual or artistic products (*Work, Expression, Manifestation, Item*); Group 2 entities refer to people (*Person*) and/or corporate bodies (*Corporate Body*) that create, publish or preserve the Group 1 entities; and Group 3 entities are used as topics in the Group 1 entities (*Concept, Object, Event and Place*). Group 1 entities are linked to each other, as well as the other two groups’ entities by a set of relationships.

Group 2 and Group 3 entities are analysed in subsequent reports. Namely, Group 2 entities are defined in detail in the Functional Requirements for Authority Data (FRAD) report [24], Group 3 entities are defined in the Functional Requirements for Subject Authority Data (FRSAD) report [25]. Recent efforts are considering the consolidation of the family of the three models into one [26]. The consolidated FRBR model, named as FRBR Library Reference Model (FRBR-LRM) [26], is expected to be announced in 2016. Yet, as presented in [26], major changes have been made with regard to Group 2 and Group 3 entities, while Group 1 entities, on which this study focuses, have been retained from the original FRBR model [1].

Group 1 entities (*Work, Expression, Manifestation, Item*, also known as WEMI) are defined from the abstract to the concrete level of a bibliographic record. *Work* and *Expression* reflect the content, while *Manifestation* and *Item* reflect the physical form of the content. *Work* refers to the intellectual or artistic creation, *Expression* refers to the realisation of the intellectual or artistic creation, *Manifestation* refers to the physical embodiment of an *Expression* of a *Work*, and *Item* refers to a single copy of a given *Manifestation* usually produced from a publication process. As far as Group 1 relationships are concerned, Tillett [27] identifies inherent relationships, e.g. the relationships between the Group 1 entities, as well as content relationships viewed ‘as a continuum from works/expressions/manifestations/items’ enabling ‘the collocation of related items and navigation through the sometimes complex network of the bibliographic universe’. A debated construct with regard to the FRBR Group 1 entities is aggregates. The Final Report of the Working Group on Aggregates [28], after a long period of investigations and deliberations, proposed a definition for aggregates and recommended modelling practices for them. Aggregates are a significant category of monographs; in 2011, 30 million aggregates were estimated in WorldCat [20]. There are many kinds of aggregates, such as anthologies, individual monographs

in a monographic series, collection of private papers, a website consisting of text and images, two items bound in one volume. In the original FRBR report, aggregates ‘operate in the same way as entities at the integral unit level’ and may be ‘defined – and related to one another in the same way as entities at the integral unit level’ [1]. Therefore, *has Part* relationships may exist at all WEMI levels, depending on the available information regarding when and how the aggregation took place. The FRBR report allows recursive relationships. As an example, a *Work* may be part of other aggregate *Works*, and an aggregate *Work* may have its own multiple *Expressions* and *Manifestations*. Representation of aggregates according to the FRBR model is complex; ‘requires creation of work-of-works and their expressions’, as well as ‘whole/part relationships to link individual works to the work-of-works’ [29].

In 2005, the IFLA FRBR Working Group (WG) on Aggregates was established ‘to investigate practical solutions to the specific problems encountered in modelling (a) collections, selections, anthologies – , (b) augmentations, (c) series, (d) journals, (e) integrating resources, (f) multipart monographs, all of which are gathered under the generic term “aggregates”’ [30]. The WG’s final report was published six years later [28] because there were serious differences between the working group members in regards to the approach they would adopt for representing aggregates in FRBR. Tillett proposed the aggregates as works approach [31], while Žumer and O’Neill proposed the aggregates as manifestations approach [32]. From the deferent definitions and approaches investigated by the WG on Aggregates, finally, the WG adopted the approach proposed by Žumer and O’Neill, while Appendix B of the report describes another approach; the report considers an aggregate as a *Manifestation* embodying multiple *Expressions*. The WG’s final report was approved on 12 September 2011 and is expected to be incorporated in the revised FRBR Library Reference Model.

Our study focuses on representation of monographs – Group 1 entities in FRBR terms – and therefore it should be noted that the next section’s graphs have been created according to the original FRBR report published in 2009 [1]. The cases of aggregates from our *Don Quixote* test case are represented according to the FRBR WG on Aggregates report published in 2011 [28] in the ‘Whole-part relationships’ section.

3.2. FRBRoo, object-oriented FRBR

FRBRoo [2] is another conceptual model, initially developed for the harmonisation of the CIDOC–CRM [9] and the FRBR models [1]. Yet it has developed as a separate model, while its by-products added a bibliographic flavour to the CIDOC–CRM, as well as re-contextualised bibliographic entities. In FRBRoo, FRBR Group 1 entities are further analysed, while more specialised classes regarding intellectual or artistic creations and their physical embodiments are identified. The event-centric orientation of the CIDOC–CRM provides FRBRoo the ability to represent not only the instances from the bibliographic entities, but also the ability to express the processes for their existence during the time. Therefore, FRBRoo could represent the history of the bibliographic entities by expressing all creation and publishing processes ‘starting from initial idea to the item a user holds’ [2]. FRBRoo is regularly updated during FRBR-CIDOC–CRM harmonisation meetings.

The FRBRoo model enables different representations of bibliographic objects according to the view that a cataloguing agency or a library wishes to implement. The official FRBRoo text [2] provides two views of expressing the FRBR *Work* and *Expression* classes: a static and a dynamic one. The static view ‘shows the relations that exist between works and expressions and the subclasses of both concepts, independently from any dynamic aspects involving the activities of creation and modification’, while the dynamic view ‘shows the dynamic process through which products of the mind come into being’ [2]. The static view is closer to the FRBR model, even though FRBR classes are further analysed in FRBRoo. The dynamic view enables detailed description of creation and publication processes and the time that these processes took place.

In the *Don Quixote* test case of the EDM-FRBRoo application profile report [16], another two FRBRoo modelling patterns (referred as paradigms) with regard to translations are explicitly stated: derivation and realisation based. Both modelling patterns conform to the FRBRoo model’s semantics and may be used in both static and dynamic FRBRoo views. The derivation-based pattern could be considered as a more detailed modelling approach where ‘every distinct intellectual contribution is modelled using individual work and self-contained expression with derivative relationships between works that are adaptations, translations etc’. The realisation-based pattern is a simpler one where translations are modelled as realisations of the same work and therefore no separate works for each translation are used.

FRBRoo is a granular model enabling multiple representations for bibliographic entities and for activities related to the intellectual creation or physical production of these bibliographic entities; thus giving cataloguing agencies and libraries the ability to select the modelling approach that best fits to their collection and cataloguing policy.

3.3. Bibliographic Framework data model (BIBFRAME)

BIBFRAME is a model developed by the Library of Congress within the Bibliographic Framework Initiative aiming to accommodate many bibliographic formats and data models. The only prerequisite stated by the Bibliographic Framework initiative is the use of Linked Data technologies and the Resource Description Framework [33] as implementation language. The BIBFRAME data model is going to serve as a transition model of legacy MARC records to linked data.

Its main classes are: *Creative Work*, *Instance*, *Authority* and *Annotation* [3]. The class *Creative Work* (or simply *Work*) reflects the ‘conceptual essence of the cataloguing item’ [3]. The class *Instance* reflects ‘an individual, material embodiment of the Work’. The class *Authority* is used to identify People, Places, and Organisations involved in the creation or publication of a *Work*. The class *Annotation* expresses comments made about a BIBFRAME *Work*, *Instance* or *Authority*. Examples of BIBFRAME annotations are: library holdings cover arts, sample texts, reviews, etc.

Four draft specifications regarding the BIBFRAME data model have been published: Annotations, Authorities, Relationships and Profiles. The Annotation model [34] defines better the *Annotation* class and provides five annotation examples (cover art, holding, review, description and annotation using an external namespace). The BIBFRAME Authorities specification [35] is accompanied by a discussion paper [36] where BIBFRAME authorities are described in detail and guidelines are given regarding their use. The BIBFRAME Relationships specification [37] defines better the relationships between *Works*, between *Instances* and between *Works* and *Instances*. The BIBFRAME Profiles specification [38] provides a formal syntactic constraints vocabulary that may be used by libraries in a specific context. A BIBFRAME profile may be used for expressing local cataloguing policies, for using another content description standard within BIBFRAME or for interoperability reasons.

On the BIBFRAME website, discussion papers, proposals, draft specifications and experimentation tools have also been published for exchange of ideas within the BIBFRAME community. According to the latest draft specifications, it seems that there is no tendency to modify the representation and the distinction between the conceptual essence of the cataloguing item and its material embodiments. In the published draft specifications the classes *bf:Creative Work* and *bf:Instance* remain with the same semantics. Exemplars of the *bf:Instance* class shall be represented with a new class called *bf:Item*. This new class will replace existing *bf:Annotation* subclasses used for this purpose, namely *bf:HeldItem* and *bf:HeldMaterial* classes. The new class *bf:Item* seems semantically close to the FRBR *Item* entity and will be subclass of the generic *bf:Resource* class and not of the *bf:Annotation* class.

Other parallel initiatives regarding BIBFRAME have been undertaken by ten early implementers. These initiatives test and examine aspects of BIBFRAME with regard to specific content types or content description standards, or experiment with BIBFRAME tools for editing (creation of new bibliographic data) or for converting legacy data. BIBFRAME is still under development and updates are anticipated.

3.4. Europeana Data Model (EDM)

Europeana (<http://www.europeana.eu/>) is an aggregation portal providing access to web-accessible digital representations of cultural heritage content provided by memory institutions all over Europe. Europeana initially used a Dublin Core based vocabulary called Europeana Semantic Elements [39]. The need for meaningful accommodation of the various domain specific metadata used by the cultural heritage community in the Europeana rationalised the development of the EDM [4]. Following Semantic Web representation principles, EDM reuses and extends existing vocabularies and supports interoperability while preserving original metadata descriptions. EDM is model-agnostic, able to aggregate descriptions of various Cultural Heritage Objects (CHOs) in order to enrich resource discovery for all memory institutions including libraries, archives and museums.

For each provider, EDM distinguishes between real provided cultural heritage objects (e.g. painting), their digital representations (e.g. digital image of painting) and their descriptions (e.g. metadata about painting). This key consideration in the EDM context, as depicted in Figure 1, is represented with two core classes, namely *edm:ProvidedCHO* (for provided Cultural Heritage Object) and *edm:WebResource* (for the *edm:ProvidedCHO* digital representations). EDM also uses OAI-ORE classes [6] in order to support the aggregation activities made by the provider of the *edm:ProvidedCHO* and *edm:WebResource* objects, as well as to contextualise metadata descriptions. More specifically, an aggregation object is instantiated in order to bridge the providers’ cultural heritage object with all its web-accessible digital representations. The specialised properties *edm:hasView* and *edm:aggregatedCHO* of the *ore:aggregates* link the *ore:Aggregation* instance with the *edm:ProvidedCHO* and the *edm:WebResource* instances, respectively. The provider’s original metadata for the cultural heritage object are preserved as properties of the *edm:ProvidedCHO*. EDM uses the *ore:Proxy* class for the option of preserving a context for the provider’s metadata. In the context-specific approach, the

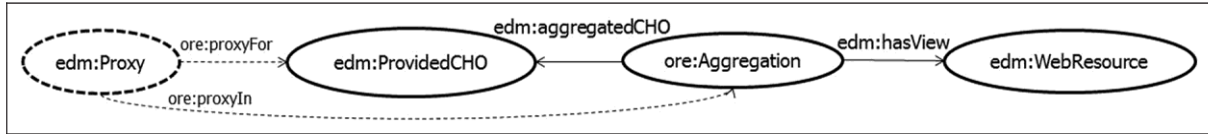


Figure 1. European Data Model: basic representation patterns.

provider's original metadata are assigned as properties of the *ore:Proxy* instance, which is linked with the *edm:ProvidedCHO* and the *ore:Aggregation* instances with the *ore:proxyFor* and *ore:proxyIn* properties, respectively. In Figure 1, use of *ore:Proxy* class is represented with dashes.

FRBR, FRBRoo and BIBFRAME, independently of the granularity level, converge to primarily separate the intellectual or artistic creation from its materialisation embodiments. Besides, EDM primarily separates the provided cultural heritage object from its digital representations.

In the context of the EDM, two actions were accomplished in order to smooth the conceptual discrepancies between EDM and the bibliographic models and to achieve interoperability. More specifically, a library metadata alignment report [40] and an EDM–FRBRoo application profile [16] have been published while there are current open issues, such as representation of serials and other library materials. In the first alignment report [40], FRBR concepts were considered, but compliance with EDM was not achieved. A new concept ‘edition’ was introduced instead to represent the union of the FRBR *Work*, *Expression* and *Manifestation* entities. ‘Edition’ information is expressed via the *edm:ProvidedCHO* class, while the web representation for the real world object is represented by the *edm:WebResource* class. The ‘Edition’ approach adopted in the library metadata alignment report is close to the current bibliographic records’ flat structure. The need of compliance with FRBR concepts was postponed and the report was designated to be reviewed in order to ‘integrate the FRBR entities in EDM using FRBRoo terms’.

This provision prompted the launch of the EDM–FRBRoo application profile Task Force in July 2012 which completed its activities in April 2013 announcing an application profile [16]. For the creation of the profile, real examples were selected and only three types of material were taken under consideration: monographs (*Don Quixote* case), plays (*Hamlet* case) and musical works (Brahms and Stravinsky cases). The EDM–FRBRoo application profile [16] accommodates FRBR concepts by using FRBRoo classes and properties. In spite of the semantic differences between EDM and FRBR, the profile did not recommend the introduction of new specialised classes in EDM. Instead, the profile used the EDM extension mechanism by assigning specific external types from other vocabularies to EDM resources. More specifically, FRBR classes are considered as *skos:Concept* instances and then specialised instances of the *edm:InformationResource* classes are generated by relating them to the SKOS FRBR vocabulary using the *edm:hasType* property. Details of the alignment between the EDM and the studied bibliographic models are presented in the *Representations* section.

The library metadata alignment report is closer to current bibliographic practices, while the EDM–FRBRoo application profile is closer to FRBR and FRBRoo. In fact, the application profile development advocates in favour of the assumption that mapping between the models is possible when it takes place within a specific context [17].

4. Representations

4.1. Single monographs

In order to explore the modelling capabilities of each conceptual model for a single-volume monograph, the first part of *Don Quixote*, published in 1605, is represented. When more than one modelling alternative exists, these cases are represented and commented.

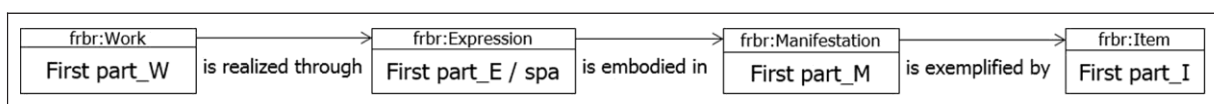


Figure 2. FRBR representation of *Don Quixote*'s First Part, single-volume monograph. Note that in this instance representation, the general FRBR relationships' cardinalities are not depicted.

Figure 2 depicts the representation of the case according to the FRBR modelling principles. The overall ideas that delimit a distinct intellectual creation are represented by a *Work* entity. As reported in the model specification, the *Work* entity is an abstract entity with no clear boundaries for its specification. All this entirety of ideas represented by the *Work* entity is realised through a textual form in a Spanish language *Expression*. *Work* and *Expression* entities are related by the one-to-many *is realised through* relationship (property). It is worth mentioning that due to the abstract essence of the *Work* entity and the absence of a single material pointing to it, a *Work* is mainly recognised through its individual realisations. Figure 2 depicts the representation of the First Part, an instantiation of the FRBR model describing only the entities with their relationships of the First Part. It is worth mentioning that Figure 2 does not depict the relationships' cardinalities supported by the general FRBR model.

The publication of *El ingenioso hidalgo don Quixote de la Mancha*, first part of *Don Quixote* in 1605, which represents all the physical objects that bear the same content and physical characteristics, is the physical embodiment of the Spanish language *Expression*. This individual materialisation embodiment defines the *Manifestation* entity. *Expression* and *Manifestation* entities are related through the many-to-many *is embodied in* relationship. The many-to-many nature of the *is embodied in* relationship implies that an *Expression* could be realised in many *Manifestations*, and a *Manifestation* could incorporate more than one *Expression*. Finally, the specific copy of this publication held in the National Library of Spain is represented by the *Item* entity. The first record (see description of records in the 'Approach' section) explicitly specifies the information related to the *Manifestation* and *Item* entities, while the information for the *Work* and *Expression* entities are implicitly inferred.

FRBR enables the representation of four distinct moments in the creation timeline of the *El ingenioso hidalgo don Quixote de la Mancha*: the set of ideas that form the *Work* later entitled *El ingenioso hidalgo don Quixote de la Mancha* (*Work* entity); the original text in Spanish (*Expression* entity); the publication product (*Manifestation* entity); and a single exemplar of that publication product held at a Library.

FRBRoo provides more than one alternative representation approach, depending on the representation granularity as well as on the modelling pattern. Therefore, single monographs may be expressed following either the static or dynamic view. In Figure 3, the static view is presented with a representation that is close to the FRBR WEMI one; specialised classes that extend the FRBR entities' semantics are used though. The First Part, entitled *El ingenioso hidalgo don Quixote de la Mancha*, is a distinct intellectual creation of Cervantes and does not have other works as parts. It is represented by an *F14 Individual Work* instance, which is realised by one and only one expression created at a particular point in time and is regarded as a complete whole (*F22 Self-Contained Expression*). It is worth mentioning that *F14 Individual Work* is a specialisation of the *F1 Work* class and represents the exact set of ideas incorporated into the *F22 Self-Contained Expression* instance. This *F22 Self-Contained Expression* is related through the one-to-many *P165 incorporates* property with an *F24 Publication Expression* instance. FRBRoo separates the intellectual creation from the publication production contexts by the representation of the publisher's contribution with the *F24 Publication Expression* class. This *F24 Publication Expression* class instance incorporates the Cervantes' *F22 Self-Contained Expression* of *El ingenioso hidalgo don Quixote de la Mancha* along with the complete set of signs, textual and visual, that the publisher had decided to use for the creation of the publication product. An *F24 Publication Expression* may incorporate many *F22 Self-Contained Expressions*. In Figure 3 there are no other incorporated expressions, even though they may exist. The *F24 Publication Expression* should be carried by an *F3 Manifestation Product Type* class instance that represents the publication product. *FRBRoo* further decomposes the concept for the FRBR *Manifestation* with the aim of differentiating between the industrial publication product created by a publisher (*F3 Manifestation Product Type* class) and the

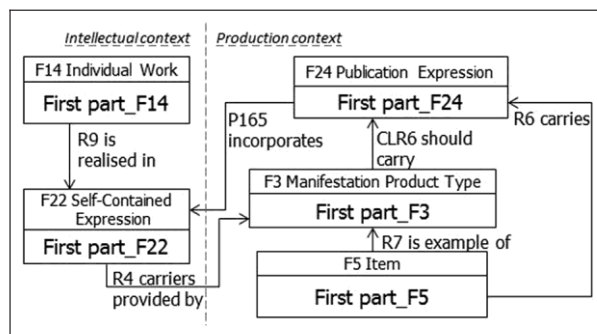


Figure 3. FRBRoo representation (static view) of a single monograph.

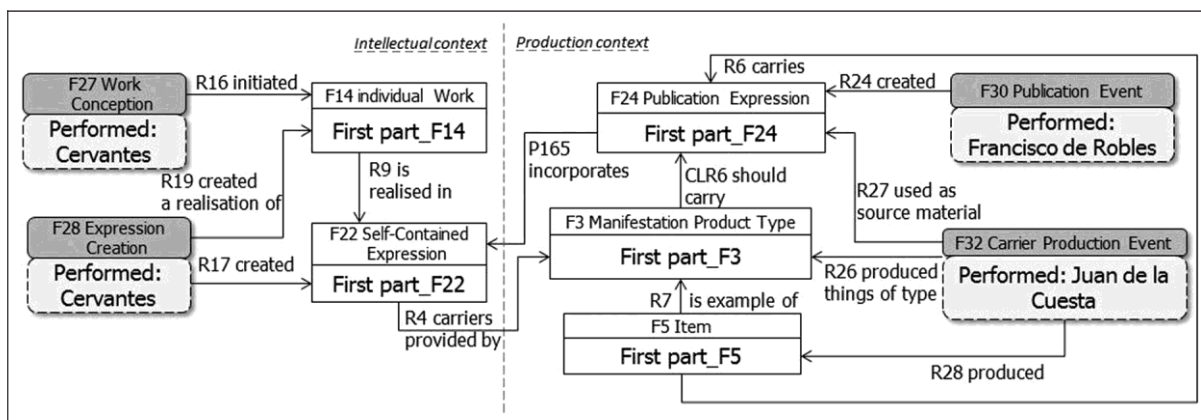


Figure 4. FRBRoo representation (dynamic view) of a single monograph.

final manuscript, or the digital material, sent by an author to a publisher (*F4 Manifestation Singleton*). A physical object produced by an industrial publication process involving an *F3 Manifestation Product Type* is held at the National Library of Spain and is represented with an *F5 Item* instance.

FRBRoo static view enables the representation of more distinct moments in the creation timeline of *El ingenioso hidalgo don Quixote de la Mancha* than FRBR. It does this by using classes that refine FRBR semantics: the set of ideas that form an individual *Work* with no other parts (*F14 Individual Work* class), the original text in Spanish that is a whole not incorporating other text-expressions (*F22 Self Contained Expression* class), the final text and layout decided for the publication product by the publisher (*F24 Publication Expression* class), the publication product (*F3 Manifestation Product Type* class) and a single exemplar of that publication product held at a Library (*F5 Item* class). It is worth mentioning that while a *F3 Manifestation Product Type* encapsulates the properties of the carrier and should carry the *F24 Publication Expression*, the actual carrier of the *F24 Publication Expression* is the *F5 Item*. In this representation the *F14 Individual Work* instance represents only the ideas realised into the *F22 Self Contained Expression* of the First Part. Alternatively, an instance of the more general class *F1 Work* could be used in order to represent the set of the union of the ideas realised in more than one expressions of the First Part.

The FRBoo dynamic view introduces ‘temporal entities’, e.g. events, activities, to associate the concepts declared in the static view to time-spans, locations and agents. For instance, as Figure 4 presents in its left and upper part, the dynamic view adds to the instance of the *F14 Individual Work* class called ‘First Part_F14’ the information that this work was initiated during an instance of the class *F27 Work Conception* event and the specific event, was performed by an instance of the *E39 Agent* class named ‘Cervantes’. Due to space limitations, Figure 4 does not present analytically that the ‘Cervantes’ node is an instance of the *E39 Agent* class and that the property *performed* associates the *E39 Agent* class (which is the domain class of the property), with the class *F27 Work Conception* (which is the range class of the property). For readability reasons, in this text, events are represented with different shapes. The upper part represents the event class; the lower part does not represent the instance of the event class, but the instance of the class *E39 Agent* who performed that event along with the property *performed*. As an example in the case of the *F27 Work Conception* event, the lower part of the shape denotes the statements: *Cervantes – rdf:type – F10 Person – P14i performed (inverse of P14 carried out by) – F27 Work Conception* event. Thus Figure 4 presents in an abbreviated form the semantic enrichment provided by the FRBRoo dynamic view.

Focusing on our case study, the First Part, *El ingenioso hidalgo don Quixote de la Mancha*, denoted as First Part_F14 in Figure 4, is an instance of the class *F14 Individual Work*, a distinct intellectual creation of Cervantes that has no other works as parts; as mentioned First Part_F14 was conceived by Cervantes in an event that is an instance of *F27 Work Conception* event class. Furthermore, an instance of the *F22 Self-Contained Expression* class, i.e. a complete expression that contains no other ones, named First part_F22, was externalised through an instance of the *F28 Expression Creation* event, performed also by Cervantes at a particular time. Then, an instance of the *F30 Publication Event* class, performed by Francisco de Robles, created an instance of the *F24 Publication Expression* class, named First Part_F24, which incorporates First Part_F22, the instance of the *F22 Self-Contained Expression* class. All the activities regarding the publication of *El ingenioso hidalgo don Quixote de la Mancha* are included in the instance of the *F30 Publication Event*.

Moreover, another *E39 Agent* named Juan de la Cuesta participated to an *F32 Carrier Production Event* and produced the instance of the *F3 Manifestation Product Type*, named First Part_F3 that carries First Part_F24, the instance

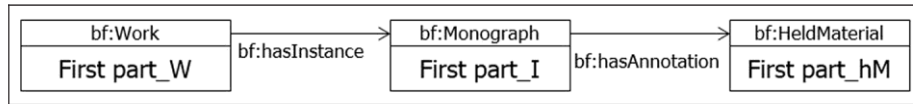


Figure 5. BIBFRAME representation of a single monograph.



Figure 6. EDM representation of a single monograph.

of the *F24 Publication Expression* class. Additionally, Juan de la Cuesta during the *F32 Carrier Production Event* produced also *First Part_F5* that is an instance of the *F5 Item* class and exemplifies *First Part_F3*. Thus the instance of the *F32 Carrier Production Event* incorporates all activities for the materialisation of the publication expression as well as the production of copies (instances of the class *F5 Item*) in a specific carrier. Thus, the dynamic view pattern adds information about the events for the generation of the instances of the classes *F14 Individual Work*, *F22 Self-Contained Expression*, *F24 Publication Expression* and *F3 Manifestation Product Type* and *F5 Item*, as well as for the participants of these events. Therefore, this results to a semantically richer representation that reveals the process and the participants of the creation of the mentioned instances.

While FRBRoo provides more granular representations of the bibliographic universe, on the contrary, BIBFRAME could be considered less detailed than FRBR in consideration to the Group 1 entities. As shown in Figure 5, single monographs are expressed in BIBFRAME relating the conceptual essence of the cataloguing resource (*bf:Work*) with its material embodiment (*bf:Monograph* subclass of *bf:Instance*) with the *bf:hasInstance* property. A copy of the *bf:Instance* held at a library is expressed as an annotation (*bf:HeldMaterial* class is a subclass of *bf:Annotation*) and related with the *bf:Instance* by the *bf:hasAnnotation* property. A *bf:Monograph* class instance is used at the *bf:Instance* level, since the First Part is a ‘single unit cataloguing resource’. BIBFRAME does not specify different conceptual levels for the idea of an intellectual creation (*Work* in FRBR terms) and its realisation in a specific intellectual form (*Expression* in FRBR terms). Therefore, the class *Creative Work* (or *Work*) ‘seems to be semantically closer to the FRBR *Work* and *Expression* entities’ [17, 38].

Europeana focuses only on and collects descriptions about European cultural heritage objects available in digital form and which are web-accessible. Therefore, according to the first record of our selected set, a First Part *Item* has been digitised by the National Library of Spain and thus may be represented in EDM. In Figure 6, a single monograph (First Part) is represented as an instance of the *edm:ProvidedCHO* class; its metadata descriptions will be assigned as properties to this instance and its digital representation as an instance of the *edm:WebResource* class (Figure 6). The *ore:Aggregation* class is used to ‘group together all important elements of cultural heritage objects contributed by the content providers’ [41]. All information regarding the First Part is included at the *edm:ProvidedCHO* class level. Therefore, the representation of single monographs in EDM presents a flat structure similar to legacy bibliographic standards.

An alternative modelling approach in EDM uses the *ore:Proxy* class [4] in the framework of preserving the provenance and the context of the *edm:ProvidedCHO* instances. A representation of single monograph in EDM using the *ore:Proxy* class is shown in Figure 7. The provider’s description (metadata) regarding Don Quixote’s First Part is at the *ore:Proxy* class instance. For each European Heritage object, use of *ore:Proxy* class shall enable in a single *edm:ProvidedCHO* class instance the contextualisation of multiple assertions made by different providers. As depicted

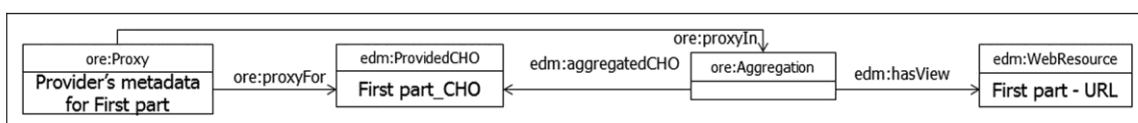


Figure 7. EDM representation of a single monograph using the *ore:Proxy* class.

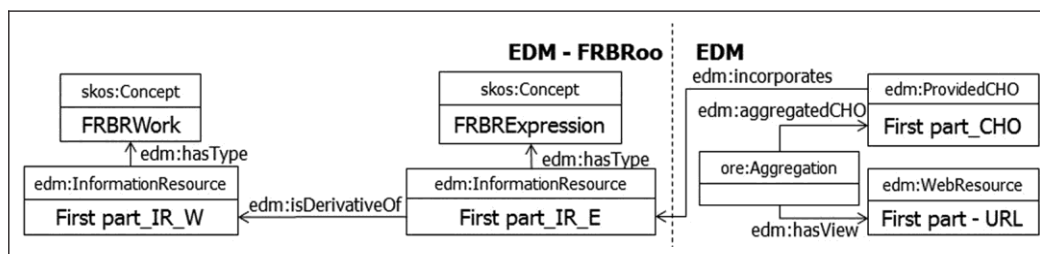


Figure 8. EDM representation of a single monograph using specialised EDM classes according to the EDM-FRBRoo application profile.

in Figure 7, for each provider providing a description for a given *edm:ProvidedCHO* an *ore:Proxy* pattern will be added.

The EDM-FRBRoo application profile [16] enables another alternative representation that extends the semantic interoperability between FRBR and EDM models. Use of EDM-FRBRoo may represent additional information on the occasion of an *edm:ProvidedCHO*, as depicted in Figure 8. Instead of using *edm:ProvidedCHO* properties to describe a library material, EDM-FRBRoo specialises EDM classes that are mapped to FRBRoo classes and properties. Thus EDM-FRBRoo accommodates FRBR semantics and all FRBR WEMI entities may be expressed.

As shown in Figure 8, a specialised *edm:InformationResource* is at the FRBR *Work* level (use of a new *skos:Concept* class instance – ‘FRBRWork’) and another specialised *edm:InformationResource* is at the FRBR *Expression* level (*edm:hasType skos:Concept* ‘FRBRExpression’). The second *edm:InformationResource* class instance derives (*edm:isDerivativeOf*) from the first *edm:InformationResource* class instance. The type of derivation – in contrast to FRBRoo – cannot be stated due to absence of a more specific property. The second *edm:InformationResource*, typed as ‘FRBRExpression’ is incorporated in the existing *edm:ProvidedCHO*. It must be noted that ‘FRBRWork’ and ‘FRBRExpression’ are *skos:Concept* instances introduced by the EDM-FRBRoo task force and not the actual FRBR *Work* and *Expression* entities.

4.2. Content relationships

According to Tillett [42], ‘A bibliographic relationship is an association between two or more bibliographic items or works’. In the same study [42], seven bibliographic relationships are recognised: equivalent, derivative, descriptive, whole-part, accompanying, sequential and shared characteristic ones. Equivalent, derivative and descriptive relationships of the content are thought as *content relationships* that ‘can be viewed as a continuum from works/expressions/manifestations/items’ [27]. In the linked data environment, content relationships could be used to provide links to *Works* that somehow derive from the same predecessor *Work*, helping users ‘to navigate and interpret complex descriptive metadata regarding bibliographic sources’ [43].

According to studies [43, 44], there is some evidence that popular works are likely to have many derivations and complex bibliographic families. These relationships may be expressed with many ways depending on the selected data model or a specific library’s cataloguing policy. In this section, we investigate the ability of the models to represent the content relationships between Group 1 bibliographic entities, in terms of changes in the intellectual content (derivative bibliographic relationships), in physical form (equivalent bibliographic relationships) and in terms of whole-part relationships.

4.2.1. Derivative relationships: translations, adaptations, free translations, etc. The definition of derivative relationships has evolved and been refined by researchers and expert groups. Tillett [42], in her cataloguing codes-based analysis of bibliographic relationships, described four types of derivative relationships: ‘(a) variations or versions of another work, such as editions, revisions, translations, summaries, abstracts, digests; (b) adaptations or modifications that become new works but are based on earlier works; (c) changes of genre, as with dramatisations and novelisations; and (d) new works based on the style or thematic content of other works, as with free translations, paraphrases, imitations, and parodies’. Smiraglia [44] focused on relationships between works and further refined derivative relationships identifying eight categories of derivation: simultaneous derivations, successive derivations, translations, amplifications, extractions, adaptations, performances and predecessor derivations. The extent of derivation between the predecessor work, also known as progenitor, and its derivatives determines if the derivative work consists a new work or not. Explicit definition of

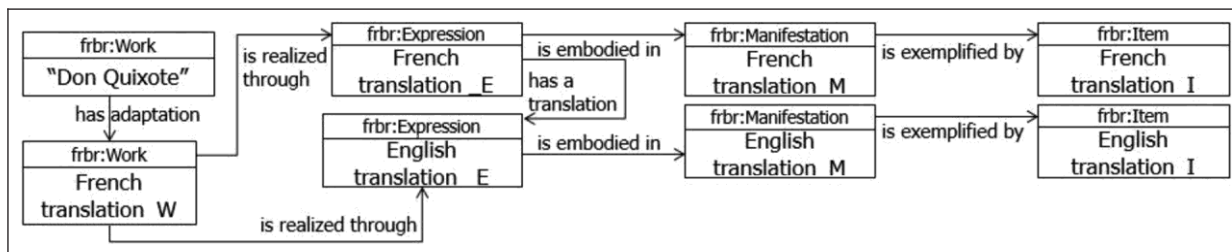


Figure 9. FRBR representation of derivative (translation-adaptation) relationships of Don Quixote.

derivative relationships was not made in the original FRBR report [45] and derivative relationships were discussed in various sections as part of relationships between *Works*, between *Expressions*, or between *Expression* and *Work* entities. The *Expression* entity was later revised [46] to better clarify when a change in the intellectual content of a given *Expression* results in a new *Expression* or in a variation within the same *Expression*. Therefore, major changes, such as revisions and modifications, result in a new *Expression*, while minor changes, such as spelling, are considered as ‘variations within the same expression’ [46]. Translations are often considered as new realisations (*Expressions*) of the same work, while adaptations and free translations are considered as new *Works* derived from the progenitor. Though the specification for the *Work* and the *Expression* entities is clear, defining two disjoint types (e.g. an instance of a *Work* must not be an instance of an *Expression* or vice versa), the decision whether the intellectual content of a new translation has been substantially changed in order to constitute a new *Work*, or otherwise another *Expression* of the original *Work*, depends on the translator’s contribution.

Don Quixote is a work with many translations and adaptations. The two parts of *Don Quixote* have been translated into many languages as independent volumes, as well as in a single volume. In our selected records, there are two translations of *Don Quixote*. Filleau de Saint-Martin has freely translated the original *Don Quixote* (First and Second Part in a single volume) in French, also changing the end of the novel. According to the FRBR report [1], his free translation (adaptation), entitled *Histoire de l’admirable don Quichotte de la Manche*, is considered a new *Work*. John Phillips later translated Filleau de Saint-Martin’s French text in English. The Phillips English translation is close to the French text and consequently is considered as a new *Expression* not to the original Cervantes work, but to the Filleau de Saint-Martin’s derived translation.

As already mentioned in section 2, the *Work Don Quixote* refers to the whole set of concepts that appeared during the coherent evolution of the initial ideas expressed in the 1605 first publication of *El ingenioso hidalgo don Quixote de la Mancha*. Therefore, following the above considerations in FRBR representation (Figure 9), the free French translation *Histoire de l’admirable don Quichotte de la Manche* is considered a new *Work* that is related to the original *Don Quixote Work*¹ with the *has adaptation* relationship. The Phillips English translation, as depicted in Figure 9, is considered as a new *Expression* that realises the Filleau de Saint-Martin’s *Work*. The English translation *Expression* is related to the French translation *Expression* with the *has a translation* relationship.

In FRBRoo representation (Figure 10), an instance of the general *F1 Work* class, referred to as *Don Quixote*, could represent the whole set of concepts appeared during the evolution of the First Part’s initial ideas. The popular French translation was published many times and had been translated into many languages. Therefore, to express the dominance of the concept of Filleau de Saint-Martin’s French text, the abstract content of the French translation is modelled as an *F15 Complex Work*, related to the *Don Quixote* instance with the *R2 is derivative of* (type: adaptation) property, also linking the derived *Works* from the French translation. Moreover, the set of ideas completely and uniquely expressed in the French translation is represented by an *F14 Individual Work* instance, which is a member of the *F15 Complex Work*, while their complete set of signs are realised by an *F22 Self-Contained Expression* instance. Respectively, the ideas completely and uniquely expressed in the single set of signs of the English translation are modelled as an *F14 Individual Work*, which is also a member of the *F15 Complex Work*, derivative (the type of derivation is translation) of the French translation *F14 Individual Work*. The English translation is considered as a distinct intellectual creation with no other works as parts, externalised by J. Phillips through an *F28 Expression Creation* event and resulted at a particular time in the creation of a complete expression that contains no other ones (*F22 Self-Contained Expression* instance).

FRBRoo *F22 Self-Contained Expressions* are incorporated into others depending on the text that each contributor has used for his own expression – variation of *Don Quixote*. Consequently, as depicted in Figure 10, an ‘*R2 is derivative of* (type translation)’ relationship relates the English translation work to the original French translation one. Therefore,

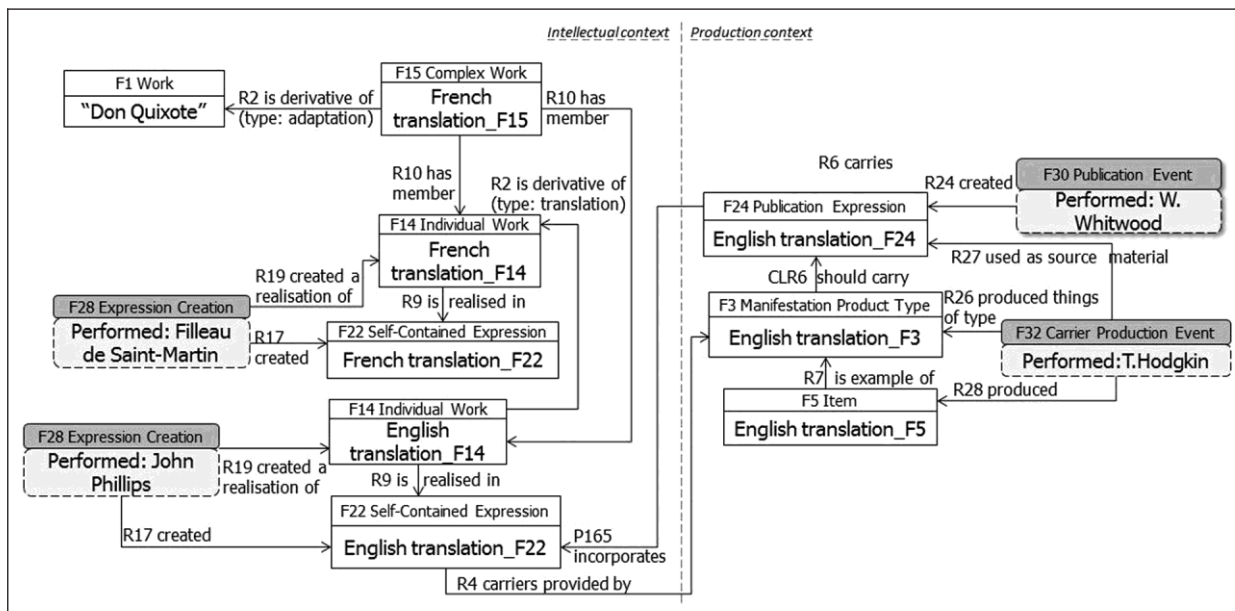


Figure 10. FRBRoo representation of derivative (translation, adaptation) relationships by specialising the *R2 is derivative of* property with the respective property type.

derivations may be explicitly expressed at *Work* level by the *R2 is derivative of* property and indicate the type of the derivation using the *type* property of the *R2 is derivative of* property.

As it has already been stated, a cataloguing agency may select its policy and implement specific patterns in cataloguing works and their translations. In the supplementary document² of the EDM–FRBRoo profile [16], two different modelling patterns that a cataloguing agency may implement in regard to translations are identified: the derivation-based pattern and the realisation-based pattern. The derivation-based pattern, which is the pattern already presented above, is detailed enabling the representation of each intellectual contribution as a new work with its own expression, while the corresponding derivative relationships are declared between works. Alternatively, the realisation-based pattern is a rather simple one, since translations are considered as realisations (expressions) of the same work. It is worth mentioning that the decision for the derivation or the realisation pattern refers to translations, making another realisation to the original intellectual content. Undoubtedly, free translations, as well as translations that incorporate additions and/or modifications to the intellectual content, clearly justify new works forming derivation patterns. Adopting either the realisation or the derivation approach is a policy decision for a uniform representation of all vague cases. Both patterns may be applied in both FRBRoo views (static or dynamic).

Alternatively, following the realisation pattern (Figure 11) there is the *Don Quixote F1 Work* that has as derivation (adaptation) the French translation. Since the French translation is a free translation with an ending other than the original *Don Quixote*, it is considered as an *F1 Work* externalised by Filleau de Saint-Martin through an *F28 Expression Creation* event and resulted at a particular time in the creation of an instance of *F22 Self-Contained Expression*. John Phillips used the French translation for his English one. Similarly, as presented in the FRBR representation in Figure 11, adopting the realisation pattern, the English translation is not considered as a new *Work* but as another complete expression that contains no other ones (*F22 Self-Contained Expression* instance). This *F22 Self-Contained Expression* instance was created by J. Phillips through an *F28 Expression Creation* event. A drawback of this simpler representation is that the translation is not explicitly represented using a specialised property, such as an *R2 is derivative of* property typed as translation. The more generic *R3 is realised in* property is used instead.

BIBFRAME utilises only one class to represent the intellectual idea as well as its realisations. Any derivation in the intellectual content of a *bf:Work* is considered as a new *bf:Work*. Depending on the type of derivation, the progenitor *bf:Work* is related to new *bf:Works* with the corresponding property. Derivation is represented using the *bf:hasDerivative* property, translation is represented using the *bf:hasTranslation* property, while realisations are represented by the *bf:hasExpression*. The free French translation is represented as another *bf:Work* derivative of the original *Don Quixote Creative Work* (Figure 12). Due to the absence of an expression entity, the English translation is a new *bf:Work* related to the French translation *Creative Work* with the *bf:hasTranslation* property. In BIBFRAME, implementation of the

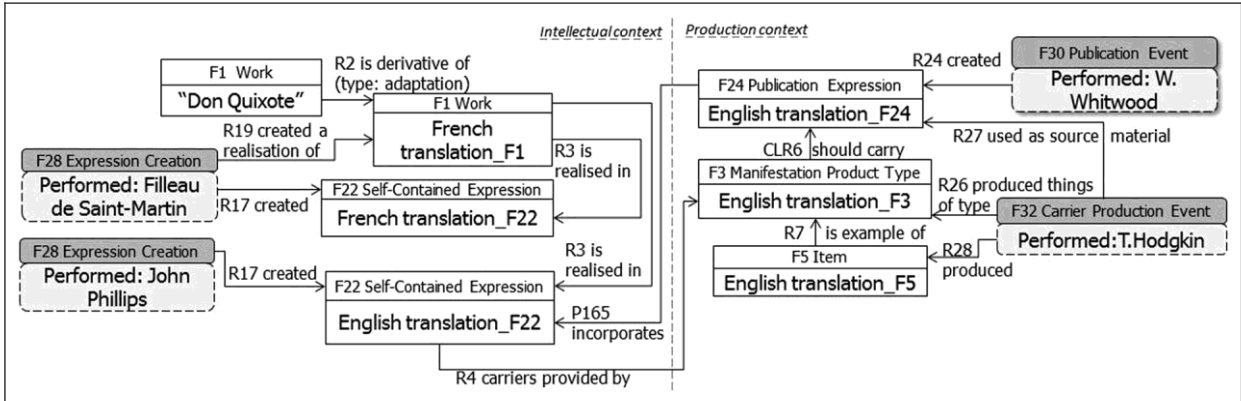


Figure 11. FRBRoo alternative representation of derivative (translation) relationships using the realisation pattern.

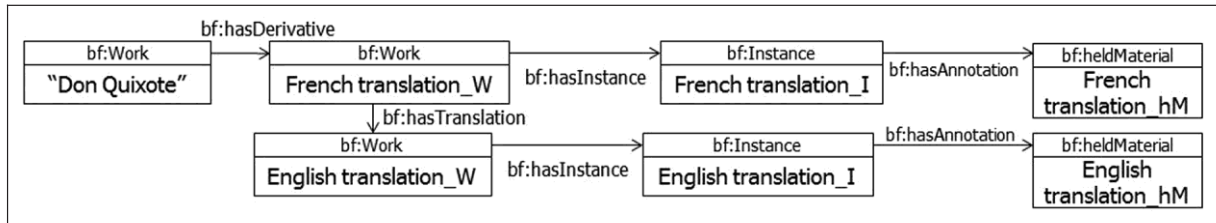


Figure 12. BIBFRAME representation of derivative (translation, adaptation) relationships.

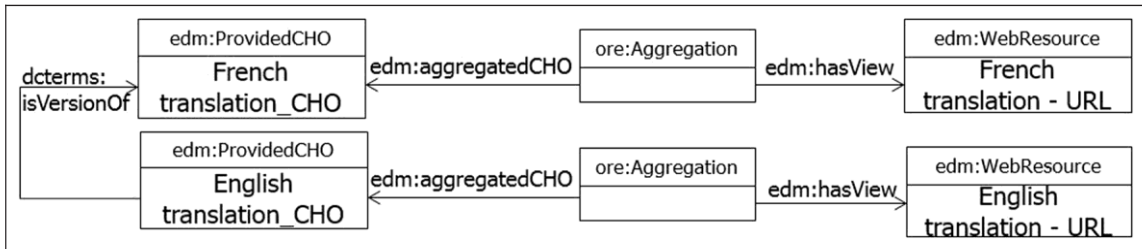


Figure 13. EDM representation of derivative (translation, adaptation) relationships.

realisation pattern is not possible because there is no distinction of the intellectual content from its realisations. Moreover, expression relationships between different realisations of the same work could be expressed as relations between *Works* using the *bf:hasExpression* property.

Due to EDM's wider scope, compared to the other studied models, it does not intend to model the abstract concepts of work and expression, therefore the intellectual concept of *Don Quixote* could not be represented directly. While, every provided object is represented as an instance of the *edm:ProvidedCHO*, derivations may be represented using the *dcterms:isVersionOf* and *dcterms:hasVersion* properties. According to the official DCMI metadata terms vocabulary [47], these two properties are used in cases of versions, editions or adaptations. Therefore, both the French translation and the English translation are represented as instances of the *edm:ProvidedCHO* class (Figure 13), linked with the *dcterms:isVersionOf* property.

As already stated, the EDM–FRBRoo profile supports the semantic interoperability between FRBR and EDM models enabling representation of more information on the occasion of an *edm:ProvidedCHO* class instance. Therefore, the relationship of the French translation with *Don Quixote* may be represented, even when there might not be a *Don Quixote* cultural heritage object description in Europeana. The relationship is expressed at the *Work* level (Figure 14) with the *edm:isDerivativeOf* property, which states that the French translation information resource (typed as an 'FRBRWork')

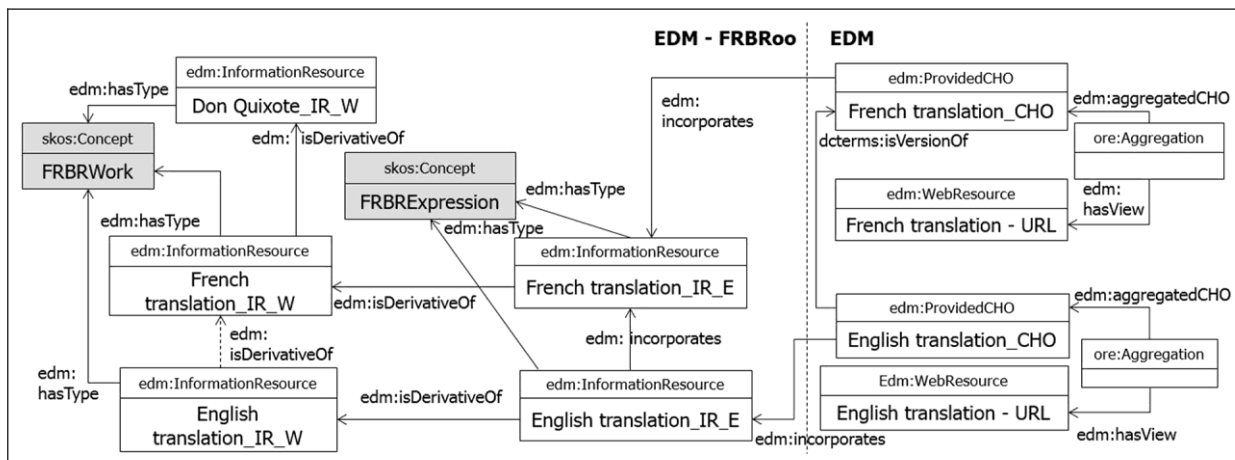


Figure 14. EDM alternative representation of derivative (translation, adaptation) relationships using specialised EDM classes according to the EDM-FRBRoo application profile.

derives from the *Don Quixote* information resource (also typed as an ‘FRBRWork’). The EDM-FRBRoo application profile does not explicitly state on which level derivations should be expressed. Since the *edm:ProvidedCHO* is the key class in Europeana representing the cultural heritage objects hosted by European cultural institutions, in Figure 14 derivations are expressed at the *edm:ProvidedCHO* level with the *dcterms:isVersionOf* property. An alternative approach of representing derivations in EDM-FRBRoo would be at the *edm:InformationResource* (*edm:hasType* ‘FRBRWork’) level using the *edm:isDerivativeOf* property. This alternative representation for the derivation relationship (translation) between the French translation and the English translation is expressed with a dashed arrow. Use of the *edm:isDerivativeOf* property does not provide any information regarding the type of derivation that relates the two *edm:InformationResource* class instances.

4.2.2. Equivalence relationships: reproduction, digitisation. Equivalence relationships are involved between copies of the same manifestation or reproductions of the same item, provided that there are no changes in the intellectual content and the authorship [42]. Facsimile copies, reproductions, reprints and so on all realise the original work through the same mode of expression. Differences may exist only at the physical format of the new reproduction of the *Item*. A reprint may exhibit ‘many of the same physical characteristics’ [1] as the *Items* of the original *Manifestation*, even though the initial intent of the reprint process is to preserve the intellectual or artistic content of the original *Manifestation*. Digitisation preserves the intellectual or artistic content of the original *Manifestation* by creating a new digital embodiment (*Manifestation*). Facsimiles not only preserve the intellectual or artistic content, but also the appearance and the physical characteristics of all *Items* exemplifying the original *Manifestation* [1].

Don Quixote is a classic literary work that has fallen out of copyright. Many libraries have reproduced rare publications for preservation reasons or have digitised them for online access. Most *Items* (in FRBR terms) described in our selected records have been digitised and are available online. MARC records inherently describe more than one entities and a common problem is to distinguish the descriptions of originals and reproductions in various analogue and digital formats [48]. The English translation has been selected as an example.

As already stated, differences in reproductions occur in the individual material embodiments. Therefore, for the digitised version of the English translation of *Don Quixote*, in FRBR (Figure 15) the action of reproduction results in the creation of another *Manifestation* for the digitised version of the *Manifestation*. Even though digitisation has occurred based on a specific *Item* at hand, a new *Manifestation* is created. The *Manifestation* from which the digitised *Item* was reproduced is related to the new *Manifestation* for the digitised *Item* through the *has a reproduction* relationship. This happens when the reproduced *Item* is representative of all the exemplars of the *Manifestation*, as well as when the reproduction results to an *Item* with different physical characteristics. It is worth mentioning that the alternative *Item-to-Item* way for specifying reproduction relationships in FRBR is only implemented when the reproduction results to an *Item* with the same type of the carrier, which is not the case for digitised materials. If the reproduced *Item* is not representative of the *Manifestation*, e.g. there is something characteristic or special about this *Item*, a note, a signature, etc., the *has a reproduction* relationship is placed between the special *Item* and the new *Manifestation* which exhibits the digital

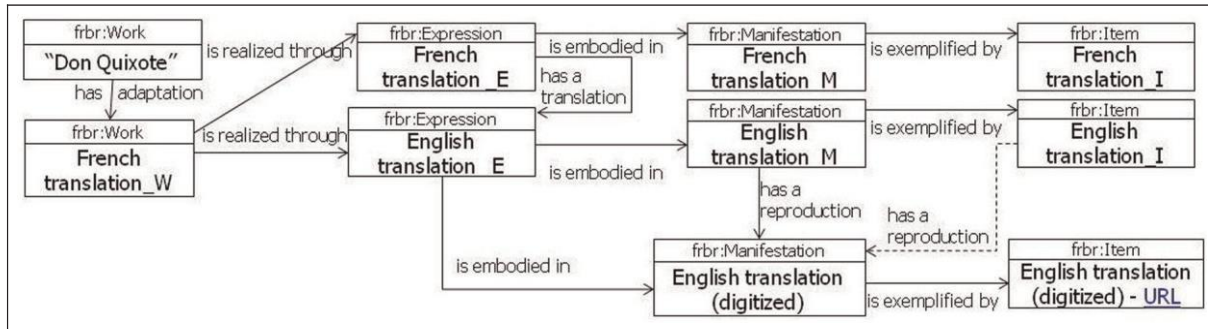


Figure 15. FRBR representation of equivalence relationship expressed at *Manifestation* level. The dashed arrow depicts an alternative representation for the reproduction of a characteristic *Item*, expressed between the digitised *Item* and the new *Manifestation* exhibiting the reproduction characteristics.

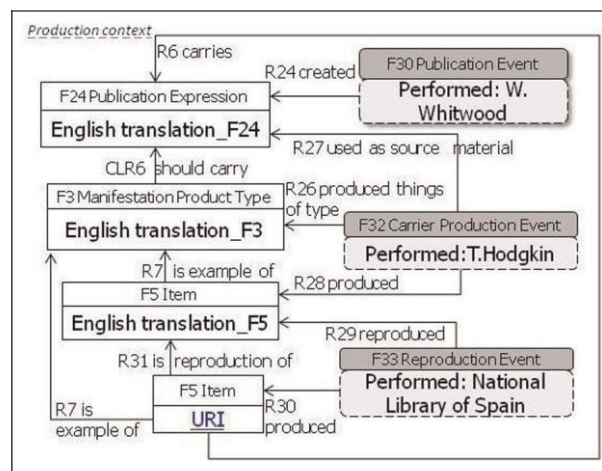


Figure 16. FRBRoo representation of equivalence relationship.

characteristics of the reproduction (dashed arrow in Figure 15). Moreover, a URL from where the digitised copy (*Item*) could be accessed can be used as an attribute of the respective *Item* instance.

FRBRoo represents reproduction at the *Item* level and gives more information about the reproduction event through the *F33 Reproduction Event* class. In Figure 16, the production context is only depicted for readability reasons. The *F24 Publication Expression* incorporates the *F22 Self-Contained Expression* of the English translation (not depicted in Figure 16) and is used as source material in an *F32 Carrier Production Event* when T. Hodgkin printed the book. This *F32 Carrier Production Event* produced an *F3 Manifestation Product Type*, examples of which are an *F5 Item* held in the National Library of Spain and another reproduced *F5 Item*, digitised copy available online through the Biblioteca Digital Hispanica website. The digitised *F5 Item* was produced by an *F33 Reproduction Event* that reproduced the printed *F5 Item*. It is worth mentioning that while a *F3 Manifestation Product Type* encapsulates properties of the carrier and should carry the *F24 Publication Expression*, the actual carrier of the *F24 Publication Expression* is the *F5 Item*.

In BIBFRAME, media embodiment is expressed at the *Instance* class level; hence, reproduction is expressed at the *Instance* class level using the *bf:reproduction* property (Figure 17). In case of a digitisation produced from a unique item, the domain and the range restrictions of the *bf:reproduction* property do not permit the representation of reproductions at the *bf:heldMaterial* level. On the contrary, FRBR permits such representation at the *Item* level. Moreover, while FRBRoo represents reproductions at the information carrier level focusing on the different attributes between them, BIBFRAME considers a reproduction as another material embodiment.

In EDM, reproductions are considered as different resources represented by their own *edm:ProvidedCHO* instances. Different versions of a *Cultural Heritage Object* may be represented and related to each other with equivalence

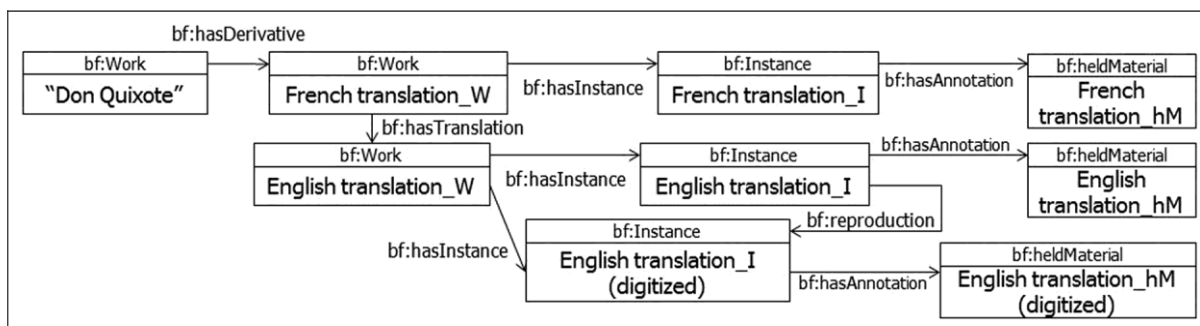


Figure 17. BIBFRAME representation of equivalence relationship (*bf:reproduction* property at the *bf:Instance* level).

relationships expressed by *dcterms:hasFormat* and *dcterms:isFormatOf* properties. These two properties relate resources that are ‘substantially the same _ , but in another format’ [41].

4.2.3. *Whole-part relationships.* Representation of aggregate and component constructs, along with their relationships between FRBR Group 1 entities, is a matter of study and deliberation. Whole-part relationships are one type of bibliographic relationships identified by Tillett [42]. Whole-part relationships exist between bibliographic resources and their parts. They include multipart monographs, monographic series, components, aggregates, etc. Components are dependent parts of a *Work* intended to be used only in its context [1]. Aggregates are collections of works and include anthologies, monographic series, collection of private papers, a website consisting of text and images, two items bound in one volume and so forth. Aggregates are a significant category of monographs as estimated by researchers using WorldCat data [19, 20]. Moreover, aggregates are important in the electronic environment where different resources are used together to form complex resources such as an individual web page and a whole website.

In the original FRBR report, aggregates and components are considered as *Works* having the same relationships as if they were ‘integral units’ [1]. Therefore, *has Part* relationships might exist at any one of the *Work–Expression–Manifestation–Item* entities or even between them, e.g. *Work–Expression*. The decision of using *has Part* relationships between WEMI instances was a cataloguing policy matter and was depended on the available information providing details about the aggregation process. As it has already been stated, the FRBR report allows recursive relationships; hence, representation of aggregates according to the FRBR model may become really complex and sometimes inconsistent. IFLA in order to study and tackle the complexity and inconsistencies in the application of the FRBR model to aggregates established a Working Group (WG) on Aggregates in 2005.

The IFLA FRBR WG on Aggregates, after a long period of study and deliberations, provided a report with modelling pattern recommendations, as well as the proper amendments for the FRBR report. This report focused on specific categories of aggregates and considered different components, e.g. a book chapter, than aggregates. More specifically, the WG studied three distinct types of aggregates of Group 1 entities: (1) aggregate collection of expressions; (2) aggregate resulting from augmentation; and (3) aggregate of parallel expressions. For these three types of aggregates, the WG on Aggregates report [28] suggested that aggregation should be represented at the *Manifestation* level, while component parts would be modelled by whole-part relationships, as already considered by the FRBR report. The aggregate *Manifestation* may embody multiple *Expressions*. An interesting feature of this approach is the ‘Aggregating Work’. The Aggregating Work is used to represent the effort of the aggregator for selecting different *Works’ Expressions* and arranging them in the aggregate [21]. It is worth mentioning that another approach discussed in this working group, presented in Appendix B of the report, was handling aggregates as *Works* enabling cataloguing agencies to either treat ‘an aggregate entity as an integral unit’ and ignore its components or describe both the aggregate entity and some or all its components. Moreover, Appendix B stated that ‘_ this Appendix is a reminder that more aggregates and components exist in FRBR and in the bibliographic universe that FRBR represents’. The Aggregates report was approved in 2011 and is expected to be incorporated in the consolidated FRBR LRM model [26]. Thereafter, aggregates are represented at the *Manifestation* level, while components (e.g. a table of a report) are considered and related to one another as entities at the integral unit level.

Cervantes published the story of *Don Quixote* in two separate monographs in 1605 and 1615. While both parts were firstly published in a single volume by Bautista Sorita and Sebastian Matevat in Barcelona in 1617, our example for aggregates is a latter edition. The third record in our set of records, denoted as Annotated, describes a latter single volume

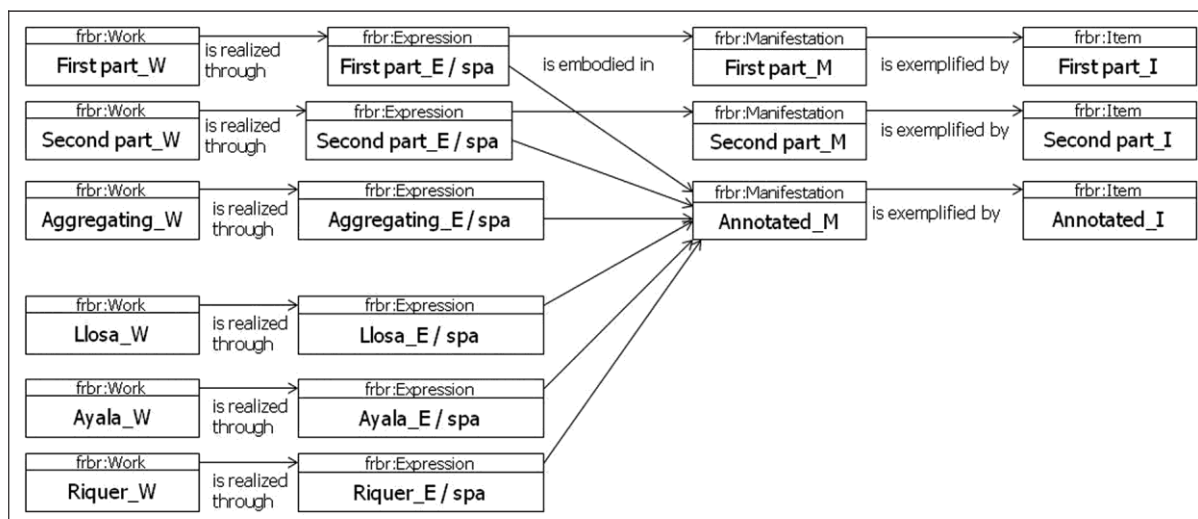


Figure 18. FRBR representation of aggregates following the FRBR WG on Aggregates report. Unlabelled arrows represent *is embodied in* relationship.

publication entitled *Don Quijote de la Mancha*, published by Alfaguara for the Real Academia Española and the Asociación de Academias de la Lengua Española in Madrid in 2004. Following the FRBR WG on Aggregates report [28], the First and Second Part *Expressions* are embodied in the Annotated aggregate *Manifestation* (see Figure 18). In this Annotated aggregate *Manifestation* the Annotated *Expression* of the Annotated *Work* is also embodied. The Annotated *Work* represents the idea of Francisco Rico to aggregate and annotate the First and Second Part of *Don Quijote*, as well as to include *Novela*, *Invención*, *Lengua* and the other *Works* in one aggregate *Manifestation*. The *Expression* of this Annotated *Work* represents only the signs related to Rico's contribution and not the signs related to the whole aggregate. Each aggregated *Work*'s signs are realised by its own *Expression*. As depicted in Figure 18, the Annotated aggregate *Manifestation* embodies the *Novela*, *Invención* and *Lengua Expressions* of the *Novela*, *Invención* and *Lengua Works*. It is worth mentioning that an *Expression* is not necessary to be embodied into its own *Manifestation*, but could be embodied to a common *Manifestation* embodying many other *Expressions*. For reasons of clarity, the property *is embodied in* is written only once in Figure 18.

Aggregates and components had also been considered extensively during the development of the FRBRoo model. For the representation of any idea for adding value to a set of already existed works, with respect to the related process, the *F16 Container Work* class or its specialisations could be instantiated. More specifically, the *F16 Container Work* class comprises works whose essence is adding value to expressions of other works either by selecting and arranging them (*F17 Aggregation Work*) or by adding features such as layout (*F19 Publication Work*) or by adding instrumentation to musical scores (*F20 Performance Work*). According to FRBRoo, the *F17 Aggregation Work* is used to represent the concept of aggregation, namely the 'selection and/or arrangement of expressions of one or more works' [2]. Consequently, an instance of the *F17 Aggregation Work* class may represent the intellectual effort of the person or group of people who conceived the aggregation and not the resulting aggregate work. The *F17 Aggregation Work* class as a subclass of *F14 Individual Work* class can be only realised in a *F22 Self-Contained Expression* class instance which may incorporate other *F2 Expression* class instances. In case this *F22 Self-Contained Expression* does not incorporate other expressions, then it represents only the concept of aggregation as the *F17 Aggregation Work* realises, otherwise it represents the whole of all aggregated expressions including the aggregation work expression. Due to the fact that in most cases a publisher has the idea of aggregation, an *F24 Publication Expression* may incorporate all distinct *F22 Self-Contained Expressions*, including the one realising the *F17 Aggregation Work*.

Different bibliographic records of the Annotated edition and its introduction were used. For the 400th anniversary of the First Part publication in 1605, Francisco Rico had the idea of aggregating the whole text of *Don Quijote* (First Part and Second Part), along with texts from notable authors and Spanish philologists into one *Manifestation*. His contribution for selecting and arranging the volume is also significant. Therefore, the Francisco Rico's idea and effort of aggregation are expressed with an *F17 Aggregation Work* instance, which is realised in *F22 Self-Contained Expression*. This *F22 Self-Contained Expression* is incorporated along with the *F22 Self-Contained Expression* instances of the First Part,

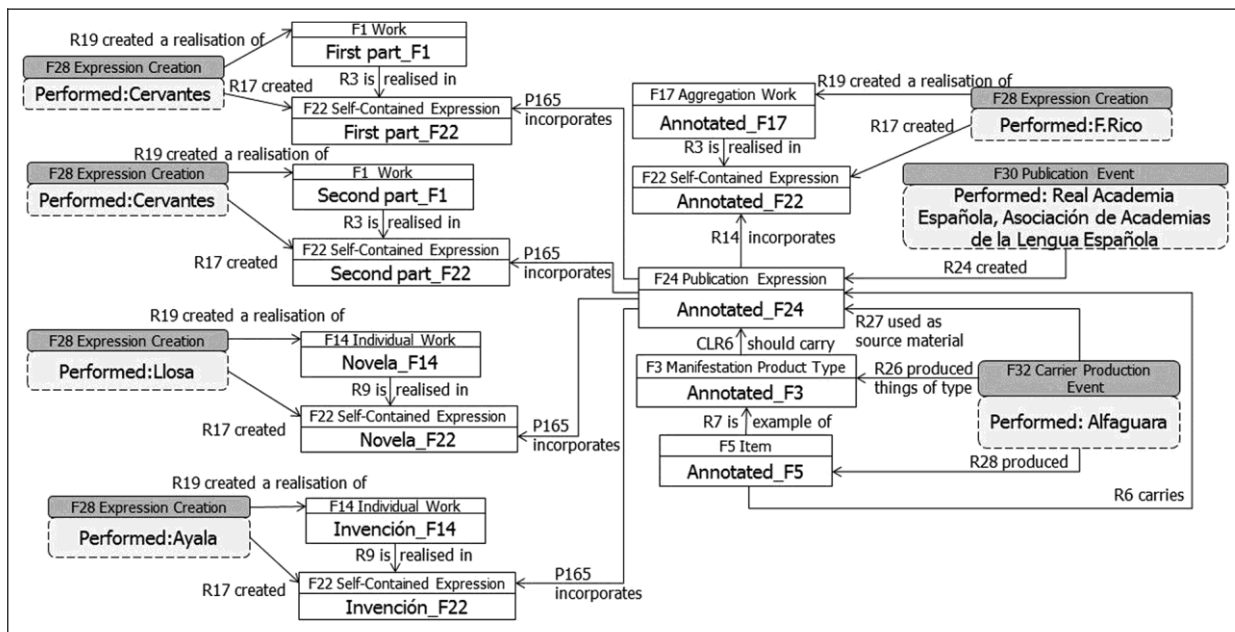


Figure 19. FRBRoo representation of whole-part relationships.

Second Part, Novela, Invención and Lengua in an *F24 Publication Expression*. For reasons of clarity, Figure 19 does not include the *F3 Manifestation Product Type* and the *F5 Item* instances for the First Part, Second Part and the other aggregated *Works*.

Representation of aggregates and components in FRBRoo is at the *Expression* level where more than one *Expressions* may be incorporated into one. The *F17 Aggregation Work* class is similar semantically to what the WG on Aggregates report names as Aggregating Work. Despite this similarity between FRBRoo and the WG on Aggregates report [28], there is a significant difference. *F22 Self-Contained Expression* instances that realise *F17 Aggregation Work* instances may represent both aggregator's effort and all the aggregated expressions. On the other hand, the WG on Aggregates adopts the 'aggregates-as-manifestations' approach, where the aggregation is expressed at the *Manifestation*. As a result, there is one *Expression* that realises the Aggregating *Work* only, along with a number of individual *Expressions* realising the aggregated *Works*. All these *Expressions* are then incorporated into the same *Manifestation* embodiment.

The BIBFRAME model does not describe in a special document how aggregates should be represented within its framework. Yet there are many properties for aggregates and components described in the BIBFRAME model [3] and in the draft specification on BIBFRAME relationships [37]. According to modelling primitives, representation of aggregates and components in BIBFRAME seems to be very close to the FRBR report, before the acceptance of the WG on Aggregates report, which treats aggregates as 'integral units'. Whole-part relationships for multipart monographs, as well as aggregates and components, may be represented using the *bf:hasPart* and the *bf:unionOf* properties. The *bf:hasPart* property describes the fact that a resource constitutes an either physical or logical part of another resource. While the *bf:hasPart* property is sub-property of the *bf:relatedTo* without any additional domain or range restriction, it may relate instances between *bf:Works* or *bf:Instances* or even between them. The *bf:unionOf* property may be used only at the *bf:Work* level representing that two or more *bf:Works* came together to form a new *bf:Work*.

Since the *bf:Work* class is considered equivalent to the union of the FRBR *Work* and *Expression* entities [38], in Figure 20 aggregation is represented at the *bf:Work* level. The *bf:hasPart* property is used to relate the Annotated *bf:Work* with the two other *bf:Work* instances that Cervantes wrote. The Novela and the Invención *Works* have not been published and are not represented in Figure 20. The alternative representation of aggregation at the *bf:Instance* level is represented with the dashed *bf:hasPart* relationships at the *bf:Instance* level, requiring the prior knowledge of the *bf:Instances* incorporated into the Annotated single volume.

Due to cardinality restrictions of the BIBFRAME model, BIBFRAME is not able to represent aggregates following the FRBR WG on Aggregates report [28]. According to the BIBFRAME model [3], a *bf:Instance* reflects the materialisation embodiment of one and only one *bf:Work*. This one-to-many relationship constraint between the *bf:Work* and the *bf:Instance* classes does not allow the embodiment (*bf:hasInstance*) of more than one *bf:Works* into one *bf:Instance*.

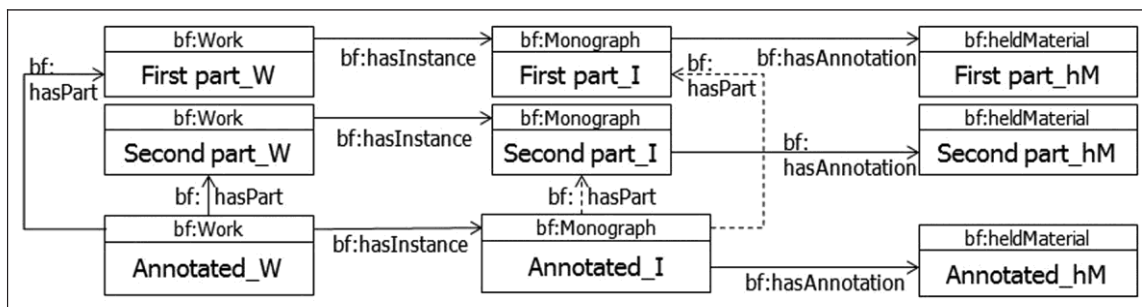


Figure 20. BIBFRAME representation of whole-part relationships.

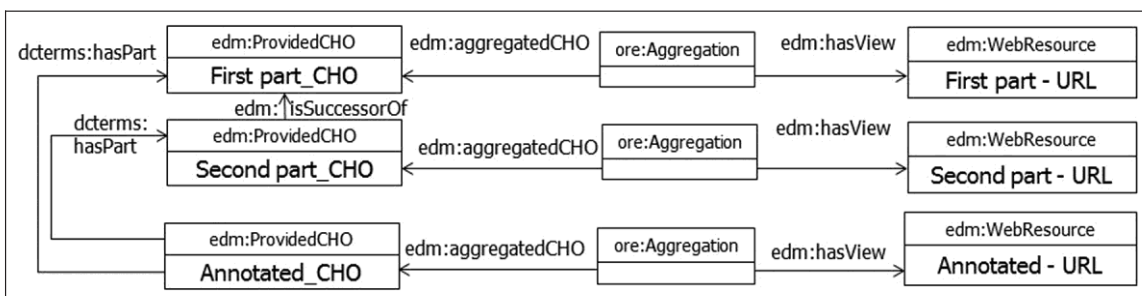


Figure 21. EDM representation of whole-part relationships.

Therefore, an aggregating *bf:Instance* shall always be an instance of one and only one *bf:Work* that represents the whole, that is the concept of aggregation along with all other concepts from the incorporated *bf:Work* instances.

Moreover, depending on the type of aggregation, other properties that could also be used are *bf:absorbed* (*bf:absorbedBy*) and *bf:absorbedInPart* (*bf:absorbedInPartBy*), *bf:accompanies* (*bf:accompaniedBy*) and *bf:supplement* (*bf:supplementTo*). The properties *bf:absorbed* (*bf:absorbedBy*) and *bf:absorbedInPart* (*bf:absorbedInPartBy*) are used for representing *bf:Work* instances that have been wholly or partially incorporated by another *bf:Work*. The *bf:accompanies* (*bf:accompaniedBy*) property is used for the representation of resources accompanied by other resources. The *bf:supplement* (*bf:supplementTo*) property, sub-property of *bf:accompanies* (*bf:accompaniedBy*), is used at the *bf:Work* level representing *bf:Work* class instances that supplement other *bf:Work* class instances. Components may be represented at the *bf:heldItem* level using the *bf:componentOf* property.

Whole-part relationships may be represented in EDM using *dcterms:hasPart* or *dcterms:isPartOf* properties between *edm:ProvidedCHO* class instances. The Annotated record of our set does not provide information regarding a web-accessible equivalent resource. Assuming that there is a web-accessible digital resource for the single volume of the Annotated edition, an *edm:ProvidedCHO* class instance could be justified. This Annotated *edm:ProvidedCHO* instance would then be related to the First Part, Second Part *edm:ProvidedCHO* instances with *dcterms:hasPart* property instances (Figure 21). Moreover, it is possible to designate the publication order of the various aggregate parts using the *edm:isSuccessorOf* property. The aggregated *Works* have not been published separately and these are not represented in Figure 21. This approach is compatible with the EDM library alignment report's [40] suggestions regarding multipart works. It is worth mentioning that the EDM library alignment report does not introduce bibliographic concepts into the EDM representation as the EDM-FRBRoo profile does.

The representation of the core bibliographic entities into the EDM by the EDM-FRBRoo application profile provides more flexibility as well as more accurate semantics for modelling aggregates and components. It is worth mentioning that the EDM-FRBRoo application profile, using the extension mechanism of the EDM, represents FRBR core concepts in EDM by specialising the *edm:InformationResource* class. Therefore, representation of whole-part relationships following the FRBR report [1] is shown in Figure 22, as well as an incorporation of multiple expressions in Figure 23. According to this approach, aggregations may be expressed at the *edm:InformationResource* (*edm:hasType* 'FRBRWork') using *dcterms:hasPart* or *dcterms:isPartOf* properties. Alternatively, aggregation may be expressed at

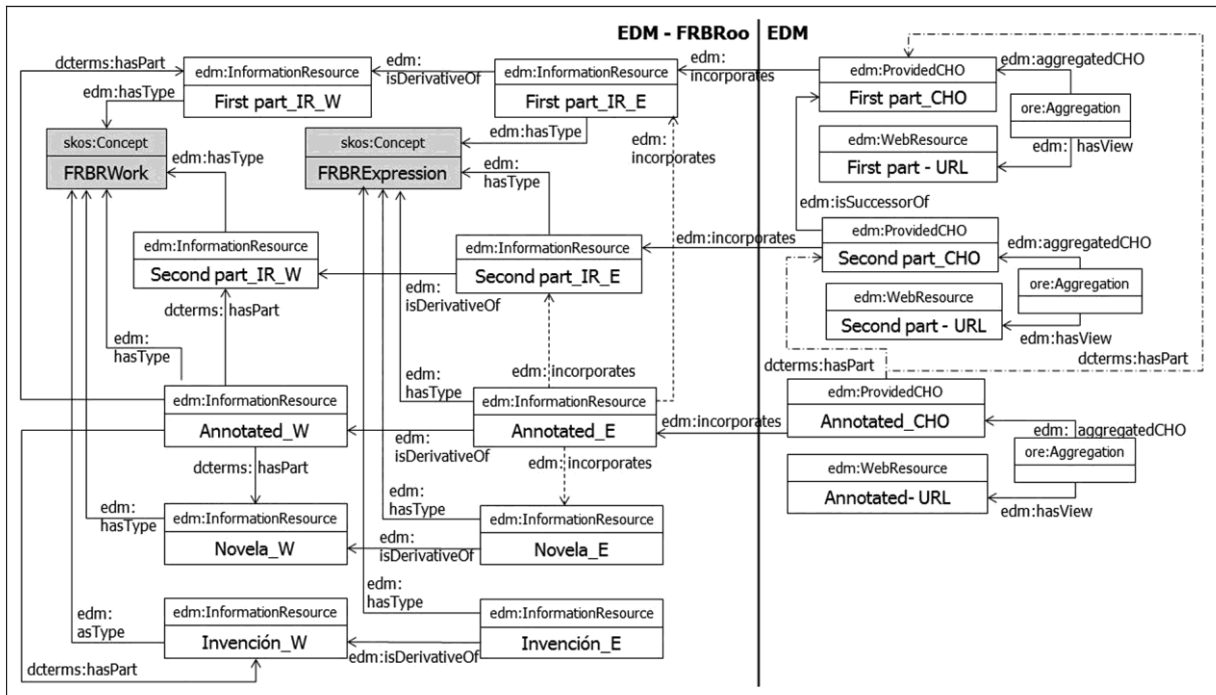


Figure 22. EDM representation of whole-part relationships based on the EDM-FRBRoo application profile. Representation is at the *Work* level (*edm:InformationResource* typed as *FRBRWork*). Dashed lines express alternative representation at the *Expression* (*edm:InformationResource* typed as *FRBRExpression*) and publication level (*edm:ProvidedCHO*).

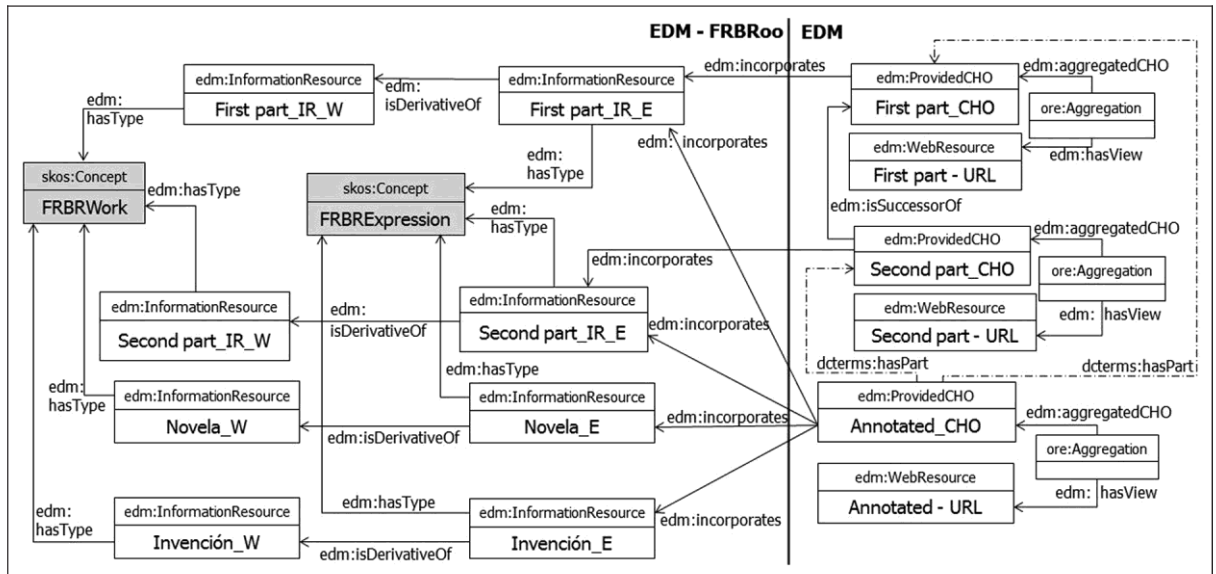


Figure 23. Alternative EDM representation of whole-part relationships based on the EDM-FRBRoo application profile. In this alternative representation, whole-part relationships are represented as multiple incorporated expressions in *edm:ProvidedCHO* instance following the *Aggregates-as-Manifestations* approach adopted by the FRBR WG on *Aggregates*.

the *edm:InformationResource* (*edm:hasType* 'FRBRExpression') level using the *edm:incorporates* property, depending on the available information about how the aggregation took place. This alternative representation is expressed with dashed arrows. Representation of whole-part relationships may also be additionally expressed as in the previous paragraph at the *edm:ProvidedCHO* level and is expressed with a long dash-dot line in Figure 22.

Figure 23 depicts a representation pattern where the aggregate is an incorporation of multiple expressions in *edm:ProvidedCHO* instance. More specifically, the Annotated aggregate instance of the *edm:ProvidedCHO*, which is considered as a publication expression by the EDM–FRBRoo profile, incorporates the expressions instances (*edm:InformationResource* typed as ‘FRBRExpression’) of the First Part, Second Part and all the other aggregated works (Novela and Invención are the only ones represented in Figure 23). This representation is semantically close to the FRBRoo representation of aggregates, as well as to the ‘aggregates-as-manifestations’ approach of the FRBR WG on Aggregates report [28]. Complimentary representation of whole-part relationships at the *edm:ProvidedCHO* level is again possible.

5. Discussion

While the four studied models (FRBR, FRBRoo, BIBFRAME and EDM) are intended for the representation of domain-specific resources from the library and the wider cultural heritage community, they are also considered as the reference endpoints in order to contextualise and extract respective entities, being implicitly, from the currently used metadata descriptions. The study of library and other cultural heritage domain-specific resources’ representations shall enable the integration of bibliographic information into the Semantic Web. The comparative study held by this paper provides an insight on how these models represent monographs, as well as of content relationships (derivative and equivalent bibliographic relationships) and of whole-part relationships between them, aiming to identify common characteristics and differences that might drive to interoperability between them. Although our study focuses on conceptual models and not on metadata schemas, the categorisation of the existing techniques for achieving metadata interoperability by Haslhofer and Klas in [49] could be utilised to present *semantic* and *structural* heterogeneities between the models.

Semantic heterogeneities occur in the meaning of the models’ classes and properties due to different interpretations of real-world entities to model elements. All studied models build upon, or at least were influenced by, ideas and principles expressed previously by librarians within the scope of defining the purpose of catalogues and meeting user needs. The FRBR model is an endpoint of discussions regarding library catalogues [50] and is used as a point of reference by the rest of the studied models. FRBRoo extends FRBR model, BIBFRAME incorporates FRBR concepts, while EDM may accommodate FRBR semantics. A central point of convergence between the models is the differentiation between intellectual content and its physical embodiments as articulated in the FRBR model. This differentiation was expressed in the Simonton report [51] to provide guidelines regarding the description of reproductions. Simonton identified two theories: the facsimile theory and the edition theory. Facsimile theory focuses on the intellectual content considering reproductions ‘of previously existing works as copies of the original’ [52], while edition theory focuses on the description of the physical item, product of a reproduction process. The first edition of the Anglo-American Cataloguing Rules (AACR) supported the facsimile theory [53]. Later, though, the second edition of AACR adopted the edition theory [54], which is still used by libraries worldwide. FRBR uses an entity-relationship approach delineating bibliographic entities and the functions that bibliographic entities’ descriptions may serve. Hence FRBR, as well as the other bibliographic models FRBRoo and BIBFRAME, present a semantic homogeneity; they all separate the intellectual or artistic content from its embodiments. EDM focuses on the wider cultural heritage domain and represents different semantics separating cultural heritage objects from their web representations. Table 1 presents each model’s core entities/classes, as well as the level they belong to (intellectual or material embodiment), providing a core mapping among them.

All models specify different primitives and present different granularity levels. FRBR and FRBRoo separate the concepts (ideas) from the signs realising the concepts, while BIBFRAME and EDM do not (Table 1). More specifically, FRBR uses the *Work–Expression–Manifestation–Item* entities representing four distinct moments in the creation timeline of a monograph: the ideas that form the *Work*, the *Expression* of these ideas as text, etc., the publication product containing also some characteristics of the exemplar (*Manifestation*) and an *Item* held by a library. FRBRoo extends the WEMI model, adds more specialised classes (e.g. *F3 Manifestation Product Type*, *F4 Manifestation Singleton*) and uses specialised events to represent activities resulting in the creation of instances of the classes (e.g. *F24 Publication Expression* class), thus enabling representation of more distinct and accurate moments in the creation timeline. BIBFRAME uses three main classes (*Creative Work – Instance – Annotation*) where *bf:Work* may be considered as the union of the disjoint FRBR *Work* and *Expression* entities. BIBFRAME’s viewpoint does not separate the intellectual content from its realisation, hence its *bf:Work* class shares similar considerations to that presented in the museums community and the CIDOC-CRM *E73 Information Object* class [55]. Moreover, BIBFRAME in order to group different realisations sharing the same intellectual content (i.e. different *Expressions* of the same *Work* in FRBR terminology) provides the *bf:hasExpression* property relating two *bf:Works* sharing the same intellectual concepts in which one is an *Expression* of the other. In EDM, regardless of its different conceptual considerations with the other bibliographic models, the *edm:ProvidedCHO* class could be considered as the union of the *Work–Expression–Manifestation* entities and therefore integrates properties

Table 1. Core entities/classes clustered according to intellectual and material embodiment characteristics, expressing semantic and structural commonalities and heterogeneities.

Level \ Models	FRBR	FRBRoo	BIBFRAME	EDM	
Intellectual	Concepts	Work	F1 Work and subclasses	Creative Work and subclasses	Provided Cultural Heritage Object*
	Signs	Expression	F2 Expression and subclasses		
Material embodiment		Manifestation	F3 Manifestation Product Type / F4 Manifestation Singleton	Instance	
		Item	F5 Item	Held Material	Web Resource*

*Provided Cultural Heritage Object instances are described in EDM, only if there is at least one digital copy (born digital or digitised) of it. A Web Resource instance provides the URL pointing to the digital copy of a given Provided Cultural Heritage Object. Note that the edm:ProvidedCHO class, as equivalent to the union of the FRBR *Work-Expression-Manifestation* entities is expanded semantically to both intellectual and material embodiment levels.

referring to them. An alternative approach is using the EDM extension mechanism to accommodate FRBR concepts as specified in the EDM-FRBRoo application profile. Without doubt, the richer a model is, the more details may be given regarding monographs' existence moments and there is always the possibility that some pieces of information would be lost when mapping bibliographic representations from a richer model to a simpler one.

At this point it must be noted that FRBR Group 1 *Work*, *Expression*, *Manifestation* and *Item* entities are often referred to as 'WEMI hierarchy'. A hierarchy presupposes that there are classes in *IsA* relationships, i.e. subclasses or superclasses, where all instances of a subclass are also instances of its superclass. This is not the case for these four FRBR Group 1 entities as specified by the model's relationships. *Work*, *Expression*, *Manifestation* and *Item* entities, as well as their respective classes in the other models, are disjoint and are related to each other with specialised non-*IsA* relationships, ruled by specific cardinality constraints. Therefore, they do not hold a hierarchy, while an *Item* instance could not be a *Manifestation* instance, a *Manifestation* could not be an *Expression*, and an *Expression* could not be a *Work* and the term hierarchy is mistakenly used. It is more likely that the intention of using the term hierarchy for these four classes is to refer to the sequence of their instantiation. An instance of an *Item* presupposes and becomes from a *Manifestation* instance, which in turn presupposes the embodiment of an *Expression* instance, which presupposes and realises the concepts of a *Work* instance.

Another semantic difference between the models is the representation of the event concept. The FRBR *Event* entity is used only as a subject in the Group 3 entities, even though Tom Delsey, as quoted in [55], accepted that the FRBR report does not explicitly discourage other uses. FRBRoo, as a harmonisation effort between CIDOC-CRM and FRBR, introduces subclasses to the CIDOC-CRM *E5 Event* class that express the creation events of bibliographic moments, e.g. *F27 Work Conception*, *F28 Expression Creation*, *F30 Publication Event* and *F32 Carrier Production Event*. In BIBFRAME there is the intention of using the concept of events, but event-related concepts are expressed without a clear hierarchy. The instances of *bf:Event* class may associate only *bf:Work* class instances with *bf:Agent* and *bf:Place* instances and a specific date or period of time. What *bf:Event* expresses is that a *bf:Agent* or *bf:Place* or a specific period of time is somehow associated with a *bf:Work*. The type of a *bf:Event* cannot be further specialised, since BIBFRAME does not provide extension mechanism as FRBRoo and EDM support. Events related to the production, publication, distribution and manufacture of *bf:Instance* class instances may be expressed using its *bf:provider* property or its sub-properties, while the range of the *bf:provider* property is the *bf:Provider* class which is subclass of the *bf:Resource* and not of the *bf:Agent* class, as it was expected. Finally, in the EDM specification [41] the *edm:Event* class is defined as equivalent to the CIDOC-CRM *E4 Period* class, a superclass of the *E5 Event* class and also to the FRBR *Event* entity resulting to an incompatibility. The *edm:Event* is not used in the representation of bibliographic information and has no specialisations defined.

Structural heterogeneities are 'caused by the distinct structural properties of models' [49]. Despite the common background, there are different interpretations of bibliographic concepts that end up to *multilateral correspondences*, *abstraction-level incompatibilities* and *meta-level discrepancies*. In our case such structural heterogeneities can be identified by observing the structural differences between the graphs generated by each model. *Multilateral correspondences* occur when a class in one model can correspond to multiple classes in another model and vice versa [49]. Examples of such multilateral correspondences can be observed in Table 1. The FRBR *Work* entity corresponds to FRBRoo *F1 Work* class and its subclasses. The BIBFRAME *Creative Work* class corresponds to both FRBR *Work* and *Expression* entities and to both FRBRoo *F1 Work* and *F2 Expression* classes and their subclasses. The EDM *ProvidedCHO* class

Table 2. Content relationships in each model* exhibiting structural heterogeneities between the models.

Content Relationships	Models	FRBR	FRBRoo	BIBFRAME	EDM
Derivative	Adaptation	Work – has adaptation – Work	F1 Work – R2i has derivative (type:adaptation) – F1 Work	Creative Work – has derivative – Creative Work	ProvidedCHO – dcterms:hasVersion – ProvidedCHO
	Translation	Expression – has a translation – Expression	F14 Individual Work – R2i has derivative (type:translation) – F14 Individual Work	Creative Work – has translation – Creative Work	
Equivalence	Reproduction	Manifestation – has a reproduction – Manifestation	F3 Manifestation Product Type – P130i features are also found on (type of similarity: reproduction) [†] – F3 Manifestation Product Type or F5 Item – R31i has reproduction – F5 Item	Instance – reproduction – Instance	Not applicable
Whole-part	Aggregates	Expression – is embodied in – Manifestation	F2 Expression – P165i is incorporated in – F24 Publication Expression	Creative Work – has part – Creative Work	ProvidedCHO – dcterms:hasPart – ProvidedCHO

*Content relationships are represented as domain-property-range statements [56].

[†]Reproduction is expressed as type of similarity, namely P130i features are also found on {P130.I kind of similarity: E55 Type = 'Reproduction'}.

corresponds to FRBR *Work*, *Expression* and *Manifestation* entities, to FRBRoo *F1 Work*, *F2 Expression*, *F3 Manifestation Product Type* and *F4 Manifestation Singleton* classes and their subclasses.

In general, there is some common ground as well as important differences between the models. As already stated, all selected models share a common background [50] and FRBR is part of this common background. Yet they were all developed to meet different needs in a common infrastructure as specified by the Semantic Web principles. These differences may be rationalised by the approach each model follows and by the time each model has been developed. FRBR is an entity-relationship model developed to support users' tasks (find, identify, select, obtain). FRBRoo is an ontology developed to serve cooperation between museum and library communities as reflected by CIDOC-CRM and FRBR models. The classes of FRBRoo are organised to a taxonomy and hence the model provides class and properties hierarchies and follows the inheritance principle. BIBFRAME is aspired to enable transition of MARC21 bibliographic data to the Semantic Web and generally to provide a steady foundation for the future of bibliographic descriptions in the broader networked world. FRBR semantics were taken under consideration during BIBFRAME's development process. EDM is an ontology focused on cultural heritage and on the Europeana aggregation context and thus it provides a class taxonomy, following the principles of CIDOC-CRM and FRBRoo. Therefore *abstraction-level incompatibilities* occur between the selected models; they all arrange same real-world entities (e.g. book, author) or concepts (e.g. intellectual content) in different hierarchies or aggregated differently into model paths [49]. A path is defined by a sequence of the triples 'domain class – property – range class' [56]. Table 2 depicts the paths for content relationships, namely derivative (adaptation, translation), equivalence and whole-part (aggregates) relationships, resulted from the representations in section 4.2., as examples of abstract-level incompatibilities. More specifically, the derivative relationship of translation is represented in FRBR at the *Expression* level (signs), while in FRBRoo is expressed at the *Work* level (concepts) or *Expression* level (signs) depending on the cataloguing pattern selected by the cataloguing agency. Translation is represented at the *Creative Work* class in BIBFRAME, while in EDM it is represented at the *edm:providedCHO* level. As shown in Table 2 there is no such heterogeneity when representation of adaptation is the case. Mapping of adaptation relationship may be prove easier, since there is some consensus between the models; adaptation is represented at the *Work* level (FRBR, FRBRoo, BIBFRAME) and at the *edm:providedCHO* level (EDM). FRBR and FRBRoo *Work* concepts are semantically close, while the BIBFRAME *Creative Work* class is somehow equivalent to the union of the FRBR *Work* and *Expression* entities.

Aggregates (whole-part relationships) are represented at the physical embodiment level (*Manifestation* entity) in FRBR, while FRBRoo represents aggregates at the signs level (*F24 Publication Expression* class). In FRBR, aggregation

is represented at the *Manifestation* level on the basis of the assumption that aggregation may not be constructed unless there is a specific externalisation of a work. Respectively, in FRBRoo, aggregation is represented at the *Expression* level (*F22 Self-Contained Expression* or *F24 Publication Expression*) where other *Works' Expressions* are incorporated. Due to the fact that aggregation is most likely to be formed by some publication activity, in this case a *F24 Publication Expression* may incorporate all distinct *F22 Self-Contained Expressions*. Moreover, for the representation of any intellectual content adding value to a set of already existed works the *F17 Aggregation Work* class could be instantiated. There are similarities between FRBRoo and the WG on Aggregates document. More specifically, the *Aggregating Work* introduced by the WG on Aggregates semantically expresses only the aggregator's effort, while the same effort in FRBRoo is expressed by the *F17 Aggregation Work* class. The *Expression* of the *Aggregating Work* expresses only the aggregator's effort, while in FRBRoo the realisation of the *F17 Aggregation Work* may represent both the aggregated content, as well as aggregator's effort. Both FRBRoo and WG on Aggregates do not express aggregates on *Work* level. Furthermore, the difference in expressing aggregates in the *Expression* and *Manifestation* levels more likely hides a common trend to the representation than a different opinion. The difference between the two approaches lies in the content that in FRBRoo aggregation is realised at the *Expression* level. FRBR *Expressions* may not incorporate *Expressions* from other integral units as well as due to the one-to-many cardinality constraint of the *is realised through* relationship. Therefore, WG on aggregates semantically is able to enable aggregation at the *Manifestation* level while the many-to-many cardinality of the *is embodied in* relationship permits the representation. Aggregation in BIBFRAME may be expressed at both *bf:Work* and *bf:Instance* classes with a number of specialised properties. The aggregations-as-manifestations approach that both FRBR and FRBRoo have adopted cannot be implemented in BIBFRAME due to the one-to-many relationship constraint between the *bf:Work* and the *bf:Instance* classes. This is an incompatibility that may hinder interoperability. Finally, EDM enables representation of aggregates at the *edm:ProvidedCHO* class. Interoperability between the models is a harder issue as far the aggregates are concerned.

Meta-level discrepancies transpire due to primitive representation conflicts among the models. While this kind of discrepancies does not imply that same semantics is not possible to be captured from the compared models, it mainly occurs when the same real-world entities or concepts could be modelled differently. The studied models do not have any direct correspondences in each other, although they can capture the same information about real-world entities (e.g. book, author) or concepts (e.g. intellectual content) [49]. Such discrepancies between the models are also presented in Table 2 in the representation of derivative (adaptation-translation) and equivalence (reproduction) relationships. Specific properties are used in FRBR (e.g. *has translation*) and BIBFRAME (e.g. *has translation*) to represent different types of derivative relationships. In FRBRoo, though, the type of derivation is not represented with a property but as a property of a property (*R2 is derivative of – R2.1. has type – translation*). Another example of meta-level discrepancy where a property of one model is represented as content value in another is the representation of equivalence. In FRBR and BIBFRAME equivalence is represented at the *Manifestation* and *Instance* levels correspondingly with use of a relationship/property, namely *has a reproduction* in FRBR and *bf:reproduction* in BIBFRAME. In FRBRoo reproduction is expressed as a type of similarity between two *F3 Manifestation Product Type* instances. In FRBRoo reproduction at the *F3 Manifestation Product Type* level is chosen when there is a commercial production scope, while reproduction for 'fair use' purposes is represented at the *F5 Item* level using a specific property (*R31 is reproduction of*).

Achieving interoperability is a challenging issue aiming to ensure that library data will be exchanged and used by different systems and applications. In general, interoperability can be achieved by bridging the structural and semantic heterogeneities that exist among different models. There have been three strategies to enable interoperability: (1) agreement on a common standard model; (2) creation of a new meta-model to which the models will be integrated; and (3) mappings between the models to resolve structural and semantic heterogeneities [49]. In previous years, AACR2r cataloguing rules [54] and MARC21 format were dominating the bibliographic universe and their adoption was world-wide. This study has shown that even in the framework of close or similar domains, namely libraries, archives and museums, it is not expected that one common model shall be used as the desired meta-model. The second strategy, which recommends the introduction of a meta-model agreement, adds one more level of complexity and is more likely to have unfavourable effects in terms of efficiency and semantics. Moreover, meta-model agreement is not applicable since the object of study is not mapping different metadata schemas to one meta-model, but the interoperability between four conceptual models. Application profile, a type of meta-model agreement, is an interesting case for achieving interoperability. First, it enables explicit definition of cataloguing policies. As shown in this paper, each model offers alternative representations to serve different needs. Each institution may select a specific way (application profile) for describing (representing) bibliographic resources to address its needs. A well-defined cataloguing policy as expressed in an application profile shall contribute to future mappings. Second, it reuses elements from existing metadata schemes which also enables interoperability. The case of the EDM–FRBRoo application profile advocates the cooperation between different communities (libraries and other cultural heritage institutions, like archives and museums) in a specific context, that of the Europeana. EDM–

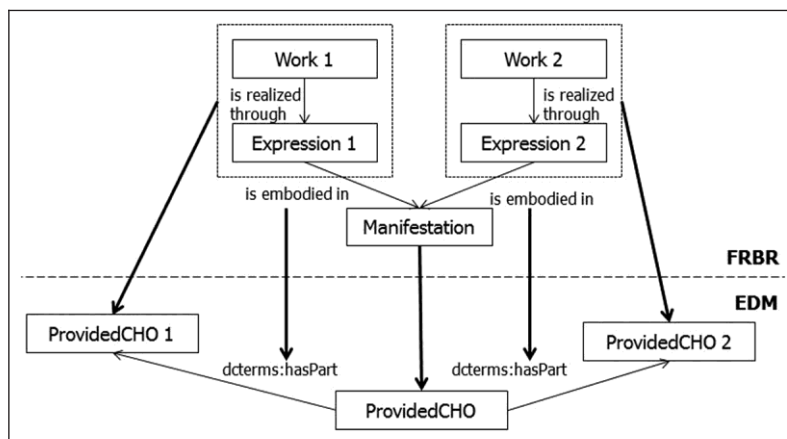


Figure 24. Mapping aggregates from FRBR to EDM using Tables 1 and 2. Bold arrows represent mappings between FRBR entities/relationships and EDM classes/properties. The mappings from FRBR relationships to EDM properties are context dependent and are valid only in the context of aggregates.

FRBRoo application profile accommodates bibliographic concepts, as realised by FRBR semantics through the FRBRoo classes and properties, in EDM by utilising the extension mechanism of the EDM model. Application profiles shall enable quality representations and ‘preservation of local meaning’ [57], while future mappings are anticipated to enable re-contextualisation of local meaning and knowledge in other domains [57].

It is common sense that mapping is easier from richer (e.g. FRBRoo) to simpler (e.g. EDM) models with loss of information, while mappings from simpler to richer models may end up using high-level classes. Therefore, it is more likely that the expressive power of the richer model will not be fully exploited. Reconciliation of semantic and structural heterogeneities may be achieved through mappings, which enable instance transformation. In general, mapping includes four stages: (1) mapping discovery; (2) mapping representation; (3) mapping execution; and (4) mapping maintenance [49]. Our findings, concisely expressed in Tables 1 and 2, present semantic and structural relationships between the studied models and consequently may serve to the first mapping stage. As an indicative example for the mapping discovery stage, consider the FRBR to EDM mappings for aggregate patterns, as demonstrated in Figure 24. FRBR represents aggregates (whole-part relationship) as a set of paths *Expression – is embodied in – Manifestation* ending to the same aggregate *Manifestation* instance (Table 2). Respectively, in EDM aggregates are represented by a set of paths *edm:ProvidedCHO – dcterms:isPartOf – edm:ProvidedCHO* ending to the same aggregated *edm:ProvidedCHO*. According to Table 1, the aggregate *Manifestation* is mapped to *edm:ProvidedCHO*. The aggregate *Manifestation* embodies two *Expression* instances. Concretely, both *Work* instances along with their realizations through the *Expression* instances are mapped to two *edm:ProvidedCHO* instances. Finally, to complete the mapping the relationship *is embodied in* is mapped to the property *dcterms:hasPart*. While in this example the *is embodied in* relationship is mapped to the *dcterms:isPartOf*, in general these relationships could not be considered as semantically equivalent. Therefore, from the aggregate mapping example it is also notable that some mappings are not context-free and could be dependent on their context.

Regarding the mapping representation stage, there exist mapping languages that can be utilised to specify machine-readable mappings enabling automated mapping executions [56, 58]. Since, paths are mapped from the source to the destination model, these languages should adopt a path-oriented approach as those presented in [56, 59] and in [60, 61]. To provide feedback that will support all mapping stages, further studies are needed to discover more semantic and structural similarities/differences between the models regarding representations of more types of bibliographic materials and more bibliographic relationships.

6. Conclusions

The present study focuses on the representation of products of intellectual or artistic endeavour, Group 1 entities in FRBR terms, in four conceptual models, namely FRBR, FRBRoo, BIBFRAME and EDM. The study involves the examination of each model’s expressiveness by exploring various representation patterns on real cases using a set of *Don Quixote* bibliographic records. These records were selected as representative of specific monograph types that may be found in great numbers in libraries, as reported by studies in WorldCat, while covering a wide range of representation

cases. Therefore, single monographs, as well as content (derivative and equivalence) and whole-part relationships (aggregates and components) were studied by creating their representations in each model. Many alternative representations and different modelling patterns were studied to investigate their influence in the expressiveness and interoperability of the models, while following each model's semantics.

Our key findings indicate that, with regard to core classes, there can be semantic similarities between FRBR Group 1 entities and the other models. As expected, FRBRoo presents similarities to FRBR. EDM presents no similarities with other models but the extension mechanism adopted by the EDM–FRBRoo application profile task force to accommodate FRBR concepts seems to provide a viable solution. BIBFRAME's *Creative Work* class is semantically dissimilar to the FRBR *Work* entity. New BIBFRAME proposals added recently to the BIBFRAME official site present a tendency to use concepts that are semantically close to FRBR/RDA ones, e.g. the *bf:HeldItem* and *bf:HeldMaterial* classes are proposed to be replaced by a new *bf:Item* class. As far as representations of certain cases of monographs are concerned, simple monographs seem to be easy to map between the selected models despite the granular differences between them. Content relationships are anticipated to be more challenging especially in the case of translations. There are incompatibilities between the models in the cases of representation of the event concept which affects the representation of equivalence relationship in FRBRoo, as well as in the representation of aggregates where BIBFRAME does not permit representations according to the aggregations-as-manifestations approach.

The evolution and the wide spread of Semantic Web impose semantically-rich representations of descriptions, while heavily affected by the underlying conceptual models and their interoperability. Future investigations are to be performed for other types of material too, while also taking under consideration anticipated development of the models, such as the consolidated version of the FRBR family of models (FRBR- Library Reference Model by IFLA [26]), the BIBFRAME 2.0 model and the BIBFRAME core by the National Library of Medicine [62]. Moreover, initiatives regarding bibliographic models are still under development, such as the Linked Data for Libraries project (<https://www.ld4l.org/>), which are likely to influence the representation of bibliographic data in the Semantic Web environment.

Acknowledgements

Thanks to our reviewers for their valuable comments.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Notes

1. *Don Quixote* is represented here as a single monograph, despite the fact that it is an aggregate work consisting of two parts. Aggregates are studied in section 4.2.3. Whole-part relationships.
2. The *Don Quixote* test case is available at: http://pro.europeana.eu/files/Europeana_Professional/EuropeanaTech/EuropeanaTech_taskforces/EDM_FRBRoo/DonQuixote.Summary.pdf.

References

- [1] IFLA Study Group on the Functional Requirements for Bibliographic Records. *Functional Requirements for Bibliographic Records Final Report*, http://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf (2009).
- [2] Bekiari C, Doerr M, Le Boeuf P et al. *FRBR Object-Oriented Definition and Mapping from FRBRer, FRAD and FRISAD (version 2.4)*. Paris, http://www.cidoc-crm.org/docs/frbr_oo/frbr_docs/FRBRoo_V2.4.pdf (2015, accessed 1 December 2015).
- [3] Miller E, Ogbuji U, Mueller V et al. *Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services*. Washington, DC: Library of Congress, 2012, <http://www.loc.gov/marc/transition/pdf/marclid-report-11-21-2012.pdf> (2012, accessed 16 October 2015).
- [4] Isaac A. *Europeana Data Model Primer*, <http://pro.europeana.eu/edm-documentation> (2014).
- [5] Digital Public Library of America. *Metadata application profile, version 4.0*, <http://dp.la/info/wp-content/uploads/2015/03/MApv4.pdf> (2015).
- [6] Lagoze C, Sompel H Van De, Johnston P et al. *Open Archives Initiative Object Reuse and Exchange ORE Specification - Abstract Data Model*, <http://www.openarchives.org/ore/1.0/datamodel> (2008).
- [7] Dublin Core Metadata Initiative. *Dublin Core Metadata Element Set, Version 1.1*, <http://dublincore.org/documents/dces/> (2012, accessed 16 October 2015).
- [8] Miles A and Bechhofer S. *SKOS Simple Knowledge Organization System Namespace Document - HTML Variant: 18 August 2009 Recommendation Edition*, <http://www.w3.org/2009/08/skos-reference/skos.html> (2009, accessed 16 October 2015).

- [9] Le Bœuf P, Doerr M, Ore CE et al. *Definition of the CIDOC Conceptual Reference Model, version 6.1*. Paris, http://www.cidoc-crm.org/docs/cidoc_crm_version_6.1.pdf (2015, accessed 1 July 2015).
- [10] Hickey TB and O'Neill ET. FRBRizing OCLC's WorldCat. *Cataloging & Classification Quarterly* 2009; 39: 239–251.
- [11] Library of Congress Network Development and MARC Standards Office. FRBR Display Tool Version 2.0, www.loc.gov/marc/marc-functional-analysis/tool.html (2009, accessed 1 July 2015).
- [12] Aalberg T, Haugen F and Husby O. A Tool for Converting from MARC to FRBR. In: Gonzalo J, Thanos C, Verdejo MF et al. (eds) *Research and Advanced Technology for Digital Libraries SE - 41*. Berlin Heidelberg: Springer Berlin Heidelberg, 2006, pp. 453–456.
- [13] Freire N, Borbinha J and Calado P. Identification of FRBR Works Within Bibliographic Databases: An Experiment with UNIMARC and Duplicate Detection Techniques. In: Goh D-L, Cao T, Sølvsberg I et al. (eds) *Asian Digital Libraries. Looking Back 10 Years and Forging New Frontiers SE - 36*. Springer Berlin Heidelberg, 2007, pp. 267–276.
- [14] Sfakakis M and Kapidakis S. Eliminating query failures in a work-centric library meta-search environment. *Library Hi Tech* 2009; 27: 286–307.
- [15] Chen Y-N and Ke H-R. FRBRoo-based approach to heterogeneous metadata integration. *Journal of Documentation* 2013; 69: 623–637.
- [16] Doerr M, Gradmann S, Le Boeuf P et al. *Final Report on EDM – FRBRoo Application Profile Task Force*, http://pro.europeana.eu/files/Europeana_Professional/EuropeanaTech/EuropeanaTech_taskforces/EDM_FRBRoo/TaskfoApplicationProfileEDM-FRBRoo.pdf (2013, accessed 16 October 2015).
- [17] Zapounidou S, Sfakakis M and Papatheodorou C. Highlights of Library Data Models in the Era of Linked Open Data. In: Garoufallou E and Greenberg J (eds) *Metadata and Semantics Research SE - 38*. Berlin: Springer International Publishing, 2013, pp. 396–407.
- [18] Svensson LG. Are current bibliographic models suitable for integration with the web? *Information Standards Quarterly* 2013; 25: 7–13.
- [19] Bennett R, Lavoie BF and O'Neill ET. The Concept of a Work in WorldCat: An Application of FRBR. *Library Collections, Acquisitions, and Technical Services* 2003; 27: 45–59.
- [20] Žumer M and O'Neill ET. Modeling Aggregates in FRBR. *Cataloging & Classification Quarterly* 2012; 50: 456–472.
- [21] O'Neill E, Žumer M and Mixer J. FRBR Aggregates: Their Types and Frequency in Library Collections. *Library Resources & Technical Services* 2015; 59: 120–129.
- [22] Cervantes project: Cervantes collection, <http://cervantes.tamu.edu> (accessed 6 October 2015).
- [23] Stavans I. One Master, Many Cervantes: Don Quixote in translation. *Humanities: The Magazine of the National Endowment for the Humanities* 2008; 29, <http://www.neh.gov/humanities/2008/septemberoctober/feature/one-master-many-cervantes> (2008).
- [24] International Federation of Library Associations and Institutions. *Functional Requirements for Authority Data: A conceptual model*. München: KG Saur, 2009.
- [25] IFLA Working Group on the Functional Requirements for Subject Authority Records. *Functional Requirements for Subject Authority Data (FRSAD): A Conceptual Model*. The Hague, <http://www.ifla.org/files/assets/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf> (2010, accessed 16 October 2015).
- [26] Riva P and Žumer M. Introducing the FRBR Library Reference Model. In: *IFLA WLIC 2015*. Cape Town, South Africa, <http://library.ifla.org/1084/1/207-riva-en.pdf> (2015, accessed 7 July 2015).
- [27] Tillett B. *What is FRBR?: A conceptual model for the bibliographic universe*. Washington, DC: Library of Congress, 2004, www.loc.gov/cds/downloads/FRBR.PDF (2004).
- [28] O'Neill E, Cato A, Goossens P et al. *Final Report of the Working Group on Aggregates*. Oslo, <http://www.ifla.org/files/assets/cataloguing/frbragggregates/FinalReport.pdf> (2011, accessed 1 July 2015).
- [29] IFLA FRBR Working group on Aggregates. *Differences between the works-of-works (Tillett) and the manifestation-of-expressions (O'Neill-Žumer) definitions of aggregates*. Milan, <http://www.ifla.org/files/assets/cataloguing/frbragggregates/definition-differences.pdf> (2009, accessed 16 October 2015).
- [30] IFLA. Working Group on Aggregates, <http://www.ifla.org/node/923> (2014, accessed 16 October 2015).
- [31] Tillett B. *Definition of Aggregates as Works: Tillett Proposal*. Milan, <http://www.ifla.org/files/assets/cataloguing/frbragggregates-as-works.pdf> (2009, accessed 22 June 2014).
- [32] O'Neill E and Žumer M. *Aggregates as Manifestations: O'Neill & Žumer Proposal*. Milan, <http://www.ifla.org/files/assets/cataloguing/frbragggregates-as-manifestations.pdf> (2009, accessed 16 October 2015).
- [33] Library of Congress. *A Bibliographic Framework for the Digital Age*. Washington, DC: Library of Congress, 2011, <http://www.loc.gov/bibframe/news/framework-103111.html> (2011, accessed 1 July 2013).
- [34] Denenberg R, Ashton J, Boughida K et al. *BIBFRAME Annotation Model: BIBFRAME Community Draft, 26 August 2013*, <http://bibframe.org/documentation/annotations/> (2013).
- [35] BIBFRAME – Bibliographic Framework Initiative. *BIBFRAME Authorities: Draft Specification – 28 April 2014*, <http://www.loc.gov/bibframe/docs/bibframe-authorities.html> (2014, accessed 16 October 2015).
- [36] BIBFRAME.org. *On BIBFRAME Authority 15 August 2013*, <http://bibframe.org/documentation/bibframe-authority/> (2013, accessed 13 July 2015).

- [37] BIBFRAME - Bibliographic Framework Initiative. BIBFRAME Relationships: Draft Specification, 25 April 2014, <http://www.loc.gov/bibframe/docs/bibframe-relationships.html> (2014, accessed 16 October 2015).
- [38] BIBFRAME – Bibliographic Framework Initiative. *BIBFRAME Profiles: Introduction and Specification (Draft – 5 May 2014)*, <http://www.loc.gov/bibframe/docs/bibframe-profiles.html> (2014, accessed 16 October 2015).
- [39] Europeana. *Europeana Semantic Elements Specification and Guidelines*, <http://pro.europeana.eu/share-your-data/data-guidelines/ese-documentation> (2013, accessed 1 July 2015).
- [40] Angjeli A, Bayerische M, Chambers S et al. *D5.1 Report on the alignment of library metadata with the European Data Model (EDM) Version 2.0*, http://www.theeuropeanlibrary.org/confluence/download/attachments/12091395/D5.1_EDM_for_libraries_v2.0.pdf (2012, accessed 16 October 2015).
- [41] Europeana. Definition of the Europeana Data Model v5.2.6, <http://pro.europeana.eu/page/edm-documentation> (2014, accessed 1 July 2015).
- [42] Tillett B. A taxonomy of bibliographic relationships. *Libr Resour Tech Serv* 1991; 35: 150–158.
- [43] Smiraglia R and Leazer G. Derivative bibliographic relationships: The work relationship in a global bibliographic database. *Journal of the American Society for Information Science* 1999; 50: 493–504, [http://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1097-4571\(1999\)50:6%3C493::AID-ASI4%3E3.0.CO;2-U/abstract](http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-4571(1999)50:6%3C493::AID-ASI4%3E3.0.CO;2-U/abstract) (1999, accessed 19 February 2015).
- [44] Smiraglia R. *Authority control and the extent of derivative bibliographic relationships* [Thesis]. Chicago, IL: The University of Chicago, 1992.
- [45] IFLA Study Group on the Functional Requirements for Bibliographic Records. *Functional Requirements for Bibliographic Records Final Report*. The Hague: Saur, 1998, <http://www.ifla.org/files/assets/cataloguing/frbr/frbr.pdf> (1998, accessed 16 October 2015).
- [46] IFLA Working Group on the Expression Entity. *Functional Requirements for Bibliographic Records Chapter 3: Entities Changes approved to the FRBR text Clean Version*, <http://www.ifla.org/files/assets/cataloguing/frbr/amend-1998-1-clean.pdf> (2007).
- [47] Dublin Core Metadata Initiative. *DCMI metadata terms*, <http://dublincore.org/documents/dcmi-terms> (2012, accessed 16 October 2015).
- [48] Urban RJ. The 1:1 Principle in the Age of Linked Data. In: *International Conference on Dublin Core and Metadata Applications; DC-2014, Austin Texas, U.S.A.*, <http://dcevents.dublincore.org/IntConf/dc-2014/paper/view/263> (2014).
- [49] Haslhofer B and Klas W. A Survey of Techniques for Achieving Metadata Interoperability. *ACM Computing Surveys* 2010; 42: 7, <http://doi.acm.org/10.1145/1667062.1667064> (2010).
- [50] Denton W. FRBR and the History of Cataloging. In: Taylor AG (ed) *Understanding FRBR: What It Is and How It Will Affect Our Retrieval*. Westport, CT: Libraries Unlimited, 2007, pp. 35–58, <http://hdl.handle.net/10315/1250> (2007).
- [51] Simonton W. The bibliographical control of microforms. *Library Resources & Technical Services* 1962; 6: 29–40.
- [52] Telerski RM. *Bibliographic Control of Preservation Photocopies*. Kent State University, <http://eric.ed.gov/?id=ED390399> (1995).
- [53] Spalding CS, American Library Association, Library of Congress, et al. *Anglo-American cataloguing rules*. London: Library Association, 1967.
- [54] Gorman M, Winkler PW and American Library Association. *Anglo-American cataloguing rules*. Chicago, IL: ALA, 1978.
- [55] Le Boeuf P. An example of library and museum cooperation: the FRBROO conceptual model. In: *CIDOC Annual Conference, Enriching Cultural Heritage, Helsinki, 10.-14.06.2012*. Helsinki, http://network.icom.museum/fileadmin/user_upload/minisites/cidoc/ConferencePapers/2012/le-boeuf-keynote.pdf (2012).
- [56] Kondylakis H, Doerr M and Plexousakis D. *Mapping Language for Information Integration. Technical report 385*. Heraklion, Crete, Greece, http://www.ics.forth.gr/tech-reports/2006/2006.TR385_Mapping_Language_Information_Integration.pdf (2006).
- [57] Oldman D, de Doerr M, de Jong G et al. Realizing Lessons of the Last 20 Years: A Manifesto for Data Provisioning and Aggregation Services for the Digital Humanities (A Position Paper) System. *D-Lib Magazine* 2014; 20, <http://doi.org/10.1045/july2014-oldman> (2014, accessed 15 December 2015).
- [58] Predoiu L, Martin-Recuerda F, Polleres A et al. *D1.5 A framework for representing ontologies consisting of several thousand concepts definitions*, http://dip.semanticweb.org/documents/D1.5_FrameworkforRepresentingOntologyNetworks.pdf (2004, accessed 15 January 2016).
- [59] Zapounidou S, Sfakakis M and Papatheodorou C. Library Data Integration: Towards BIBFRAME Mapping to EDM. *Communications in Computer Science* 2014; 478: 262–273, http://dx.doi.org/10.1007/978-3-319-13674-5_25 (2014).
- [60] Lourdi I, Papatheodorou C and Doerr M. Semantic integration of collection description. *D-Lib Magazine* 2009; 15: 1082–9873, <http://dx.doi.org/10.1045/july2009-papatheodorou> (2009).
- [61] Gergatsoulis M, Bountouri L, Gaitanou P et al Query transformation in a CIDOC CRM based cultural metadata integration environment. In: Lalmas M, Jose J, Rauber A et al. (eds) *Research and Advanced Technology for Digital Libraries: 14th European Conference, ECDL 2010, Glasgow, UK, September 6–10, 2010. Proceedings*. Berlin, Heidelberg: Springer, pp. 38–45, http://dx.doi.org/10.1007/978-3-642-15464-5_6 (2010).
- [62] National Library of Medicine. Experimentation with BIBFRAME at the National Library of Medicine – June 24, 2015, Update, <https://github.com/fallgrennj/BIBFRAME-NLM> (2015, accessed 7 July 2015).