IFLA WLIC 2018
KUALA LUMPUR

Submitted on: 26.06.2018

A Development of the Metadata Model for Video Game Cataloging: For the Implementation of Media-Arts Database

Kazufumi Fukuda
Kinugasa Research Organization, Ritsumeikan University, Kyoto, Japan
E-mail address: fukudakz@gmail.com

Tetsuya Mihara
Faculty of Library, Information and Media Science, University of Tsukuba, Tsukuba, Japan
E-mail address: mihara@slis.tsukuba.ac.jp

Abstract:

A lot of studies for video games are progressing because of the expansion of research interest in video games. But it can be said that there is a lack of means to access video games as research materials. Therefore, cataloging or metadata model for video games has been discussed since 2010. Opinions are divided among researchers on the FRBR's potential of description for video game resources. In this research, we design application model for video games based on IFLA LRM for the purpose of implementation in Media-Arts Database. The database is the only one comprehensive database produced by Agency for Cultural Affairs including bibliographic records of video game in Japan. We are proceeding research activities through cataloging practice of Ritsumeikan Center for Game Studies (RCGS), Ritsumeikan University. At First, we discuss the functional requirements for video game cataloging. Secondly, For the design of the metadata model, definitions, interpretation and criteria for record distinction were specified for the four entities of Work, Expression, Manifestation, and Item. Furthermore, the existing list of bibliographic relationships is deficient in the recording relationships of the video game. We created a vocabulary of bibliographic relationships required for the description of the video game by analysing some examples of the records in Media-arts Database and interview data to the catalogers. Thirdly, the designed model was verified through application for two cases. It was confirmed that the metadata model has potential of description to satisfy the above functional requirements.

Keywords: video games, cataloging, metadata model, bibliographic relationship, IFLA LRM

1. Backgrounds

There are still few libraries holding video games. Therefore, production volume and distribution volume of bibliographic records are also limited. Numerous studies for video games are progressing because of the expansion of research interest in video games. But it can be said that there is a lack of means to access video games as research materials. Works of video game often have multiple editions. In addition, those editions have various relations representing unique characteristics of video game production, e.g., different versions, remake, same virtual space, porting, localization, bug fix and so on. These relationships need to be recorded in the catalog of video games. Nowadays, data collaboration of archive agencies is required, and it is worthwhile to consider appropriate catalogs and metadata describing the features of game materials that also take into consideration the reusability of data.

2. Previous Studies

The subject of designing a metadata model for video games has been discussed since 2010. Jerome McDonough argued traditional bibliographic descriptions are not suitable for interactive fiction including video games in the final report of Preserving Virtual Worlds Project (McDonough et al. 2010a). Furthermore, they have been investigated the potential of FRBR as a model to describe video games (McDonough et al. 2010b). This study evaluated that it was not perfect but reasonable to record the difference among video game versions. The specific tangled relationship, e.g. series crossover, was pointed out. They mentioned the need for the entity organizing some works which is quite similar to the concept of superwork.

A close study on metadata of video games is proceeding at Washington University. They are creating metadata schema that can describe characteristic information of video game and interactive media. Jin Ha Lee proceeds the construction of metadata schema collaborate with Seattle Interactive Museum (Lee et al. 2013). This paper adopted the approach of evaluating each element of the candidate list of metadata one by one from the viewpoint of six personas, player, player's parent, collector, researcher, game developer, and librarian. 16 core metadata elements (CORE 16) were selected in this research. Lee et al. (2015) developed the schema. They proposed 46 recommended elements through user data survey and facet analysis. Jett et al. (2016) proposed a conceptual model for video games and interactive media. They criticized the FRBR model as inappropriate for recording the specific descriptive attributes, the level of distribution to record, and the relationship between video games. This model is based on the hierarchy of the four entities of Game, Edition, Local Release, and Distribution Package. The entities of Series, Franchise, and Universe are defined for grouping “Game” entity. “Agent” entity is also defined for recording the responsible agent. Collection and Additional content entities are defined to disentangle compilations and multiple versions.

Stanford University Library is also listed as one of the research centers for video game cataloging. Greta de Groat discusses the confusing situation in the dawn of cataloging and the progress of its standardization about how the game was recorded in cataloging the library from records of each field in MARC in the United States (de Groat 2015). Video Games Best Practice Task Force she is chairing published the guide for video game cataloging (Online Audiovisual Catalogers, Inc. et al. 2015). This document guides cataloging by RDA (Resource Description and Access) that adopted the FRBR model. This is a proposal adopting MARC 21 to contribute standardization at libraries.

Various research of cataloging and metadata development for video games has been developed as described above. It can be said that the standard model, especially the FRBR model is the axis of these researches. For example, McDonough evaluated the rationality of the FRBR model. On the other hand, Lee sees it critically and is aiming at the development of a specific model. The issue of "standard model vs. specific model" is a point of controversy at considering the metadata model.

3. The Purpose of This Study (or Research Objective)

We are proceeding research activities through cataloging practice of Ritsumeikan Center for Game Studies (RCGS), Ritsumeikan University (Fukuda et al. 2017, Fukuda & Mihara 2018). The targets of cataloging are video game software and related materials such as magazines and reference of game studies held by RCGS. Cataloging based on the metadata model proposed in this research is proceeding. Part of it is published in RCGS-OPAC[1].

In addition, RCGS is in charge of construction of the game field of Media-Arts Database (development version) operated by Agency for Cultural Affairs. The database is the only one comprehensive database produced by Agency for Cultural Affairs including bibliographic records of video game in Japan. The number of video game collection by metadata creating institution is small in Japan. Therefore, the bibliographic data distribution amount is small. For that reason, we proceeded to create comprehensive bibliographic data on video games circulating in Japan, using secondary resources such as product catalogs and video game magazines.

The development of the official version scheduled to be released in 2020 is proceeding in progress. This version is developing focused on the following issues:

1. Developing a metadata model that can describe specifics of resources
2. Adopting Linked Open Data technologies for the reuse metadata and the connection with other databases
3. Implement the bibliographic data on the system in a user-friendly form

We are developing a model close to the standard one. Because Media Art database covers several art forms. Game owners often collect the relevant forms of materials such as magazines, strategy guides, and soundtracks. Standard models have not only high reusability but also the potential of description for various materials as general catalog. Specifically, we design application model for video game based on IFLA Library Reference Model (IFLA LRM).

IFLA LRM was released in 2017 (Riva et al. 2017). It defines high-level conceptual reference model developed within an entity-relationship modeling framework, to quote the document, “The conceptual model as declared in IFLA LRM is a high-level conceptual model and as such is intended as a guide or basis on which to formulate cataloging rules and implement bibliographic systems”. IFLA LRM designed as the basis for multiple literature systems. Therefore, the entities are defined with high abstraction. It is thought that clarification of

[1] Ritsumeikan Center for Game Studies. “RCGS-OPAC”. http://www.dh-jac.net/db/rcgs/search.phps (accessed 2018-04-20)

interpretation is necessary for implementation in video game catalog. In this research, we design application model for the purpose of implementation in Media-Arts Database[2].

This research mainly discusses about four entities, work, expression, manifestation & item[3]. At First, we discuss the functional requirements for video game cataloging. Secondly, For the design of the metadata model for game cataloging, definitions, interpretation and criteria for record distinction were specified for WEMI. Furthermore, the existing list of bibliographic relationships is deficient in the recording relationships of the video game. We created a vocabulary of bibliographic relationships required for the description of the video game by analysing some examples of the records in Media-arts Database and interview data to the catalogers. Thirdly, the model designed in this study was verified through application for multiple cases. The description potential of the model was evaluated through these applications. Finally, we discuss several issues for cataloging with this data model as future works.

4. Functional Requirements of the Metadata Model for Video Game Cataloging

What are the requirements for the metadata model of games? As mentioned above, several previous studies suggested on this issue. We have defined the following four requirements related to WEMI.

1) Works, versions and their relationships can be recorded.
2) Versions of different game contents can be distinguished.
3) Specific relationships between records of video game can be described.
4) The relationship of the series can be recorded.

The first requirement is to record multiple versions and their associations arising from a work. For example, Pac-man's work has many ported editions, in addition to the arcade game that was first published. It is required to be able to discover and select them collectively.

The second requirement is to recognize the identity of game content. Even minor differences in content among derivatives of the same work can cause the difference of the system or operating environment. On the other hand, there are different versions of the same content as exemplified by popular versions and low-priced versions. In other words, the same content and the different one can be included in the version of the same work. It is required to be able to distinguish them. Also, we must consider the tendency of increasing online function of video games in recent years. These contents are frequently changed by online update. It is necessary to have a function to record these finely. However, the number of versions of the game got popularity is enormous. As pointed out by McDonough et al. (2010), there is a possibility that problems may arise from the viewpoint of finding and identification if a huge number of records are related in parallel to one work. Therefore, it is necessary to be able to record them structurally. For example, the structure can be recorded as follows, minor update version “is part of” major update version.

The third one is bibliographic relationships among records. The relationships can be defined as an association, connection, or interaction between different bibliographic entities, or

[2] Agency for Cultural Affairs. “Media-Arts Database: Game Field”. https://mediaarts-db.bunka.go.jp/gm/ (accessed 2018-04-20)
[3] These entities are sometimes called “WEMI” from their initials. hereinafter, these four entities are called WEMI.

components of entities[4]. Bibliographic relationships are assumed to have a contribution to user tasks such as identification and exploring. Furthermore, the bibliographic relationships affect clarification of the criteria of record classification of abstract entities such as works and phenotypes in the whole system such as catalogs (Jepsen 2005, Noruzi 2012). As the previous research discussed, the catalogs of the video games require descriptions of the specific relations (Lee et al. 2014, Jett et al. 2016). The fourth one is to describe series composed by some works or versions. Most popular games are often produced sequels. In fact, the information about series seems to be important for publishers and users to identify the game. For example, recognizing works and versions such as Dragon Quest series or Call of Duty series is common in culture of video games.

To design a metadata model that meets these requirements, it is needed to define the criterion for record creation of each entity to describe works, versions, and so on. The vocabulary of entity relationships is also required. These are shown in next chapter.

5. The Basic Metadata Model for Video Game Cataloging Applying IFLA LRM WEMI Entities

In this chapter, we describe definitions, interpretation in this model, criteria of record distinction, subclass and bibliographic relation for each entity of WEMI. After that, we discuss about descriptions of structural entities for collective expressions such as compilations or collections. Metadata model is shown by these descriptions.

We mention about the vocabulary of bibliographical relationships in advance. They already have lists such as RDA and IFLA LRM. For example, RDA has various terms of it. However, it is for recording traditional library materials. Thus, the vocabulary of relationships in RDA is insufficient for the specific description of video games. It is necessary to have a vocabulary for presenting a pattern of association between comprehensive and concrete bibliographic records as possible. Jet pointed out the issue of dispersion at the description level (Jett et al. 2016). We think that this problem could be solved by standardizing record classification. It is also assumed that designing bibliographic relationships will further contribute it.

In this research, we mapped records of the media art database to WEMI of this model. And as much bibliographic relationships as possible were extracted. We used interview data from catalogers and notes on the catalog for data creation. The vocabulary of bibliographic relationships was created through such a process.

5.1.1 Work

IFLA LRM defines work as "Intellectual or artistic content of a distinct creation[5]". Work is an entity indicates game content based on individual creations in this model. Furthermore, related works in game contents have relations to each other. Criteria to distinguish works depend on activities of creation. A remake or an extended creation of a separate responsible agent is recorded as a new work. In this case, the work and new works have relations of “remade as” or “expanded as”. The series or the sub-series covering these works can be recorded as a new work which is the superclass of the works. These relationships of series are recorded with nesting of 1: n relation.

[4] Arsenault & Noruzi (2012), p. 641.
[5] Riva et al. (2017), p. 21.

The description of tangled relationships, e.g. series crossover, is an important issue in the specification of the work (McDonough et al. 2010). The relationships cannot be recorded in the nested structure. This should be described by an entity grouping works like superwork. Cross-series features such as franchises and characters are recorded at IFLA LRM res entity in this model. In other words, it is recorded as a kind of subject.

Bibliographic relationships among works are as follows: has part, has main series, has non-main series, in series, subseries of, precede, sequel, remade as, created using same system with, created using MOD as, supplement, expanded as, abridged as, absorbed by, inspiration for, paraphrased as, spin-off. These don't include inverse relationships.

5.1.2 Expression

Expression is defined as "A distinct combination of signs conveying intellectual or artistic content[6]" in IFLA LRM. The identity of the game content serves as a criterion for distinction of expression. It also applies to fine differences. Games with different operating environments (including platforms) with the same title are recorded in new expression records. Differences in the platform often bring about differences of multiple fine contents. For example, they appear in behaviours of characters and instructions of buttons on the controller. It is considerable that they are important differences for researchers, developers and core game players. Version differences also make new records. However, some popular games that have an online function make minor version upgrades frequently. A huge number of these may appear, so it is not appropriate to record in parallel. Thus, this model defines the function for expression to describe by grouping these as "has minor version".

Bibliographic relationships among expressions are as follows: mutual complement, simultaneous derivatives of same work, same virtual space, porting, emulated, localized version, trial version, expanded, abridgement, bugfix version, preceded software version, effect by save data, has minor version.

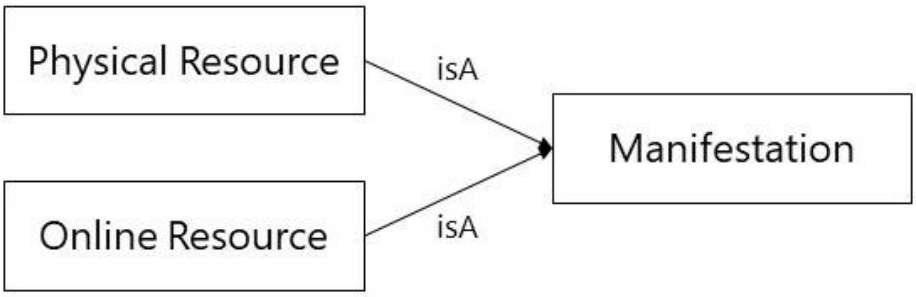
5.1.3 Manifestation

The definition of manifestation is "A set of all carriers that are assumed to share the same characteristics as to intellectual or artistic content and aspects of physical form[7]" in IFLA LRM. Differences in specifications and contents of distribution packages are the criteria for distinction of this entity in this model. Version differences such as bonus version, special version, and cheap version make new records of manifestation. The difference between carriers such as disks, cartridges and online resources also make new records.

The two subclasses are defined for recording different types of resources. They are "Physical resource" and "Online resource". For example, Attributes such as size, material, and some type of identifier, e.g. ISBN or Article Number, are not required for online resources. Figure 1 indicates relationships with manifestation and subclasses.

[6] Ibid, p. 23.
[7] Ibid, p. 25.

Figure 1. Subclasses of Manifestation



Bibliographic relationships among manifestations are as follows: deluxe edition, online resource edition, reproduction, preservation facsimile, insert.

5.1.4 Item

Item is defined as "An object or objects carrying signs intended to convey intellectual or artistic content[8]". It is an entity that records individual actual objects. It is the actual material itself. Item is recorded based on the format managed by the collector.

5.2 Other Relationships

The relationship between an entity and another entity is the same as FRBR and IFLA LRM. A work is realized by an expression. An expression is embodied by a manifestation. A manifestation is exemplified by an item. Sangokushi 13 has a power up kit which is supplementary to the game. The power up kit cannot play the game by itself.

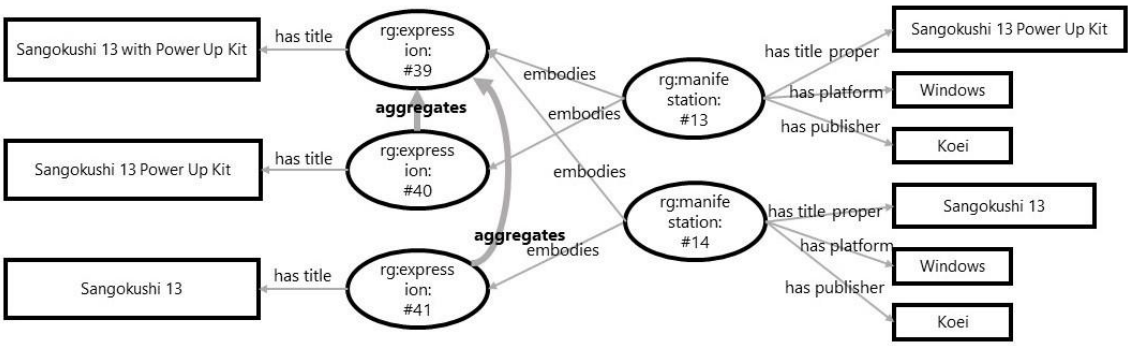
5.3 Description of Structural Entities

Examples of the video game compilation of distribution packages are not enumerated. It is necessary to mention a method for recording these appropriately. Manifestations embodying multiple expressions such as collections and compilations are recorded using "Aggregate Collection of Expressions" of "Modelling of Aggregates" that defined in IFLA LRM[9]. The collective expression "aggregates" expressions created by individual creations ("aggregates" is one of the defined relationship in IFLA LRM). Compilation is recorded in the description of this structural entity.

Augmentation is one of the forms characteristic of video games. For example, Sangokushi 13 (Romance of the Three Kingdoms 13) published by Koei corresponds. An augmented version can be played by adding a power up kit to Sangokushi 13. Such relationships often appear in video games. They are described using "Aggregates Resulting from Augmentation" of "Modelling of Aggregates"[10]. Figure 2 indicates above aggregated relationships.

[8] Ibid, p. 27.
[9] Ibid, p. 93-94.
[10] Ibid.

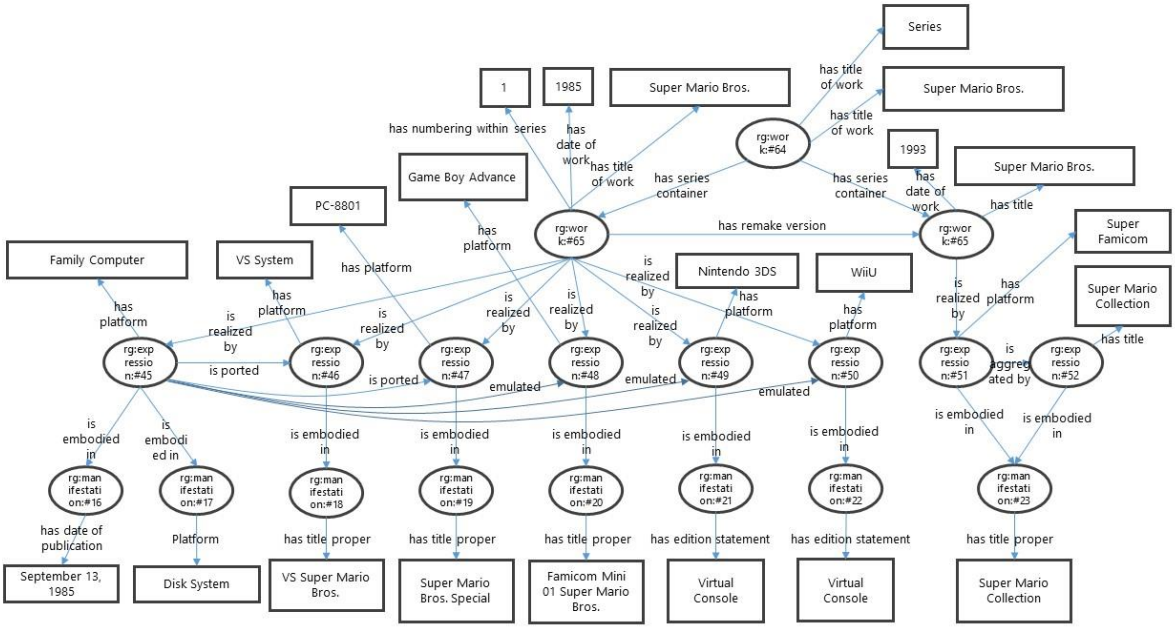
Figure 2. Example of Aggregates Resulting from Augmentation (Sangokushi 13)[11]



6. The Description Example of the Proposed Model

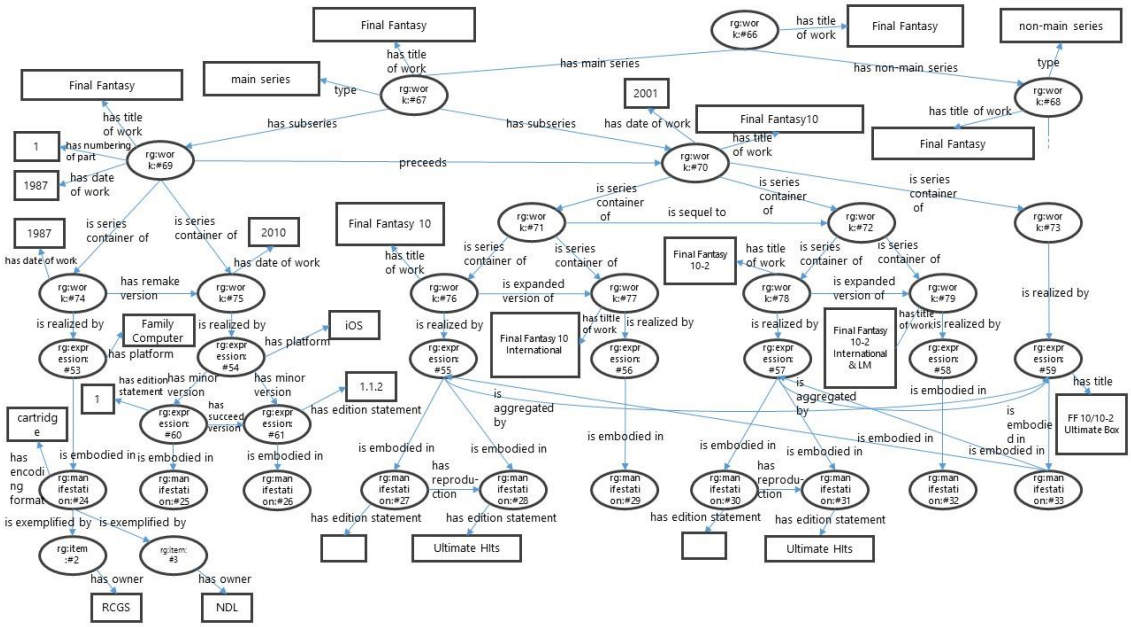
Figure 3 & 4 indicate Applying a case with the metadata model described in the previous chapter. The first case is All of the first released version of Super Mario Bros in Japan. The Second case is a part of Final Fantasy series. A potential of description is confirmed by these cases.

Figure 3. Applying Super Mario Bros. on the proposed metadata model



[11] Hereinafter, name space “http://www.rcgs.jp” is abbreviated as “rg”.

Figure 4. Applying a Part of Final Fantasy Series on the Proposed Metadata Model



7. Conclusion and Future Works

It was confirmed that the metadata model proposed in this research has potential of description to satisfy the above functional requirements. The problem of variance of the description level for WEMI could be arranged. In other words, works, expressions and manifestations correspond to creations, specifications of contents and distribution packages in this model. These were clarified by the definitions, the interpretation and the overall data with bibliographic relationships. We think that This application policy is able to propose new perspective more effective way in considering video game cataloging in the future.

However, because of the lack of space, we could not discuss attributes of classes yet. Furthermore, explanation and case verification can be mentioned only summarily in our short study. We would like to complement a topic through the more research publication or the publishing the document of specification. There are also many research topics of practice of cataloging video games with this model.

In the first place, it is necessary to evaluate the metadata model through cataloging practice. We have been preparing manuals for such cataloging since last year. We would like to advance the cataloging resources we owned by upgrading the manual to a version adapted to this metadata model. The model will be more useful by applying practical feedback. Specifically, We would like to proceed with verification of writing potential and supplement of vocabulary. In addition, it is not easy to distinguish records of work and expression that are defined as highly abstract entities in cataloging. Data created by the fan community that contributes to this subject already has been published on the Internet. They can be found in Wikipedia[12], MobyGames[13], for example. We would like to connect these data with catalogs created by archive institutions. With this connection, the discussion of data comparison

[12] "Wikipedia". https://www.wikipedia.org. (accessed 2018-04-30)
[13] "MobyGames". http://www.mobygames.com/ (accessed 2018-04-30)

progresses. Such data analysis seems to progress the classifying in patterns of works and expressions as well. It is assumed that the design of subclasses of these entities is also required. These activities will promote contributions for cataloging and associating data.

In the second place, controlled vocabulary is an important issue as a utility to support cataloging. Several previous studies are discussed related the issue (Donovan et al. 2013, Rossi et al. 2014, Clarke et al. 2015, Kaltman et al. 2016). It is assumed that these developments and research will progress through research exchange.

In the third place, there was no discussion on entities other than WEMI. It is also essential to build authority data such as agent and res. We must also discuss creating the vocabulary of specific relationships among these entities and WEMI.

Furthermore, there are some topics related to data utilization and collaboration. It is a kind of technical issue such as the formal description of data oriented to Linked Open Data and improvement of metadata schema.

This research and related practices generated several new research themes. Video game as a new type resource in libraries also has multiple research needs in perspective of cataloging. It is thought that this is extremely social applicability. We would like to continue these researches in the future.

References

Arsenault, C. & Noruzi, A., 2012. Analysis of work-to-work bibliographic relationships through FRBR: A Canadian perspective. Cataloging and Classification Quarterly, 50(5–7), pp.641–652.

Catalogers Online Audiovisual et al., 2015. Best Practices for Cataloging Video Games - Using RDA and MARC21. (October 2014), p.89. Available at: http://olacinc.org/sites/capc_files/GameBestPractices.pdf. (accessed 2018-04-20)

Clarke, R.I., Lee, J.H. & Clark, N., 2015. Why Video Game Genres Fail: A Classificatory Analysis. School of Information Studies: Faculty Scholarship, 167.

de Groat, G., 2015. A History of Video Game Cataloging in U.S. Libraries. Cataloging & Classification Quarterly, 53, pp.135–156.

Donovan, A., Cho, H., Magnifico, C., & Lee, J.H., 2013. Pretty as a pixel: Issues and challenges in developing a controlled vocabulary for video game visual styles. In Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries. pp. 413–414.

Fukuda, K. & Tetsuya, M. 2018. A Development of the Metadata Model for Video Game Cataloging –Focusing on Bibliographic Relationships-. Proceedings of DiGRA Japan 8th Annual Conference, pp. 88-91.

Fukuda, K. et al. 2017. A Study of Data Model for Game Database –Application of LOD Technology-. Proceedings of DiGRA Japan 2016 Annual Conference, pp. 22-25.

Jepsen, Erik Thorlund. 2005. Structuring the catalogue according to FRBR user task: Utilizing bibliographic relationships through collocation and linking. In: Bibliotheca Universalis: How to organize chaos? 11-12 August 2005. Finnish Library Association, 2005.

Jett, J. et al., 2016. A Conceptual Model for Video Games and Interactive Media. Journal of the Association for Information Science and Technology, 67(3), pp.505–517.

Kaltman, E. et al., 2016. Implementing Controlled Vocabularies for Computer Game Platforms and Media Formats in SKOS. Journal of Library Metadata, 16(1), pp.1–22.

Lee, J. H. et al., 2014. Relationships among video games: Existing standards and new definitions. Proceedings of the ASIST Annual Meeting, 51(1).

Lee, J. H., Clarke, R. I. & Perti, A., 2015. Empirical evaluation of metadata for video games and interactive media. Journal of the Association for Information Science and Technology, 66(12), pp.2609–2625.

McDonough, J. et al., 2010. Twisty Little Passages Almost All Alike: Applying the FRBR Model to a Classic Computer Game. Digital Humanities Quarterly, 4(2). Available at: http://www.digitalhumanities.org/dhq/vol/4/2/000089/000089.html. (accessed 2018-04-20)

McDonough, J. P., Olendorf, R., Kirschenbaum, M., Kraus, K., Reside, D., Donahue, R., & Rojo, S., 2010. Preserving Virtual Worlds Final Report. Library, p.195. Available at: http://hdl.handle.net/2142/17097. (accessed 2018-04-20)

Noruzi, A., 2012. FRBR and Tillett's Taxonomy of Bibliographic Relationships. Knowledge Organization, 39(6), pp.409–416.

Riva, P., Le Boeuf, P. & Žumer, M., 2017. IFLA Library Reference Model: A Conceptual Model for Bibliographic Information, (August), p.101. https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017.pdf. (accessed 2018-04-20)

Rossi, S., Lee, J.H. & Clarke, R.I., 2014. Mood Metadata for Video Games and Interactive Media. In Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (JCDL 2014).