

Aggregation and Utilization of Metadata for Intangible Folk Cultural Properties Using Linked Open Data

Itsumi Sato and Masao Takaku

University of Tsukuba, Tsukuba, Ibaraki, Japan

s1921633@s.tsukuba.ac.jp

masao@slis.tsukuba.ac.jp

<http://www.slis.tsukuba.ac.jp/>

Abstract. Intangible Folk Cultural Properties (IFCP) represent cultural customs or events related to transition in people’s lives, and need to be protected and passed on to future generations. The Japanese government adopted the revised Act on Protection of Cultural Properties in 2018, which aims to ensure comprehensive protection and utilization of cultural properties. IFCP are required to make it available and accessible to the public considering these features. This study proposes an IFCP data model using Linked Open Data (LOD). This model is based on the CIDOC Conceptual Reference Model (CRM) and other vocabularies. We constructed a dataset based on this model and published it on the web. The dataset contains 5,106 triples from 103 IFCP focusing on religious faiths festivals (RF) and annual observances (AO). To evaluate our data model, we defined functional requirements for the IFCP and implemented a prototype system to verify utilization feasibility of the IFCP. The prototype system shows the IFCP lists based on retrieval feature using the SPARQL query language.

Keywords: Intangible Folk Cultural Properties · Metadata · Linked Open Data · CIDOC CRM · SPARQL

1 Introduction

1.1 Background

In November 2018, “Raiho-shin, ritual visits of deities in masks and costumes”, an intangible Japanese cultural heritage consisting of ten different rituals such as Namahage, was inscribed on UNESCO’s Representative List of the Intangible Cultural Heritage of Humanity [1]. The news attracted people’s attention. Each Raiho-shin ritual is registered as Intangible Folk Cultural Properties (IFCP) in Japan. IFCP are intangible heritage created by people to express and celebrate their daily lives and major life transitions [2].

The Japanese government revised the “Act on Protection of Cultural Properties” in June 2018. This act emphasizes that cultural properties can be leveraged

to support town development, and local communities need to protect and work together to pass them on to future generations [3]. In particular, IFCP are regarded as living cultural objects, and very natural, as stability due to their characteristics cannot be stored in physical form and tend to change over time [4]. Thus, considering their importance and characteristics, it is extremely important to immortalize them by recording information on these IFCP for posterity.

Additionally, some museums have little knowledge of how to deal with intangible cultural heritage customs, such as IFCP, and little interest in them among museum staff. Sometimes, IFCP records are not maintained in an appropriate state by museum staff [5]. Furthermore, multimedia data such as images and videos about intangible cultural heritage customs, including IFCP, need to be easily accessible through the web to be widely utilized [6].

1.2 Research Objective

We propose a data model structuring IFCP information using LOD. In addition, we examine the model utilization by aggregating its structured information.

In this study, we focus on RF and AO, and design a data model to make information about them available using LOD. We adopted the CIDOC Conceptual Reference Model (CRM) [7] as the main vocabulary. Additionally, we implement a utilization application by aggregating the structured AO and RF information.

2 Related Works

2.1 Existing Platform for Cultural Heritages

Europeana: Europeana [8], launched in 2008 and managed by European cultural heritage organizations, is a representative portal for the retrieval and access of digitized cultural and academic heritage items. Europeana collects metadata of their digital contents and presents it as LOD using API. This data model introduces the Europeana Data Model (EDM) [9]. This EDM consists of OAI ORE, Dublin Core, Simple Knowledge Organization System (SKOS), and CIDOC CRM.

Japan Search: Japan Search [10] is a national, integrated, and cross-sectoral portal that contains various digital archives containing books, cultural objects, and art media, to comprehensively retrieve metadata for the country's patrimony. It provides aggregated metadata in accessible formats and promotes the utilization of this content. The official version of the portal was released in August 2020. Japan Search aims to reveal the whereabouts of contents and help the country's digital information resources be discoverable and utilized effectively.

2.2 Structured Data and Model for Intangible Cultural Properties

Tan et al. [11] conducted modeling for data related to the Dragon Boat Festival, a Chinese intangible cultural heritage. Their knowledge data model was structured

centered on intangible cultural heritage information's entities and properties based on CIDOC CRM. Then, an experimental prototype was implemented, and the validity of its ontology was verified. The results revealed that the data model is flexible and dynamic.

Giannoulakis et al. [12] constructed a metadata schema focusing on another intangible cultural heritage, folk dance. They examined their original metadata model to describe the complicated metadata of folk dance. This data model adopted Dublin Core, MovementXML, TEL, and VRA, and defined an original schema. As a result of this analysis, they came to the conclusion that encoding intangible cultural heritages is complex work, owing to the consideration for some contextual factors, such as the evolution of people's social environment and their emotions related to performance.

These studies focus on one intangible cultural heritage or one genre of it, and aim to create structured data on that target. However, intangible cultural heritage or IFCP have various genres and objects, depending on cultural backgrounds and regions. In this study, we aim to conduct structuring IFCP information beyond a single target and genre.

2.3 Implementation Using SPARQL

Dannells et al. [13] proposed a multilingual SPARQL-based retrieval interface for querying cultural heritage data in natural language. This system is based on grammatical framework, provides a grammar-based approach, supports natural language queries in 15 languages and it inquires as SPARQL queries and receives its language queries.

Mishima [14] proposed a metadata model and system requirement functions for the aggregation and integration of regional cultural heritage information. He aggregated its core information, targeting a cultural heritage list on which Tokyo municipalities open to the public through the web. The system, which supports the discovery of regional cultural heritage information adopting aggregated metadata of cultural heritage lists and LOD, was implemented using SPARQL's query functions.

While these studies focus on tangible cultural heritage subjects, this study focuses on intangible cultural heritage, especially IFCP in the country, and conducts structuring of the data model and implementation using this data model. Additionally, before implementing the application, functional requirements were defined based on users' needs.

3 IFCP

3.1 Definition

As described in the introduction, IFCP are intangible heritage customs, practices, artistic expressions, and values that people created in their daily lives to mark and celebrate transitions. IFCP are a part of cultural properties in Japan.

Cultural property is a term of cultural administration in Japan and refers to a cultural heritage that has been created, fostered, and passed down throughout the long history until the present day [15].

3.2 IFCP Categories and Scope

In this study, IFCP are categorized based on the Important Cultural Properties Database (Kunishitei Bunkazai-tou Database) [16], which consists of three main categories and 16 sub-categories (Table 1).

Table 1. Categories and IFCP count as of July 2019

Main category (number of properties)	Sub category (number of properties)
Manners and customs (130)	Production and regular vocation (9) Life and ceremony (6) Entertainment and games (10) Social life (folk knowledge) (2) Annual observances (34) Festivals (religious faiths) (69) Others (0)
Folk performing (159)	Kagura (35) Dengaku (25) Furyu (spectacular dancing) (35) Narrative and entertainment for blessing (5) Ennen and deed (7) Entertainment from abroad and stage performance (36) Others (16)
Folk techniques (16)	Production and regular vocation (14) Food, clothing and housing (2) Others (0)

We focus on Important IFCP (IIFCP), which are especially important cultural properties in Japan. There are 305 such IIFCP listed in Japan’s Agency for Cultural Affairs Important Cultural Properties Database as of July 2019. From this record, we extracted 103 properties including RF (69 properties) and AO (34 properties), for use in this study. RF are formed from people’s mental attitudes by repeating religious experiences or courtesies. AO are religious events or official ceremonies held repeatedly and traditionally every year.

In terms of the number of registered properties, RF is the largest sub-category, and AO is the second largest sub-category in the manners and customs main category. They can be seen as representative IFCP. Additionally, they share common characteristics regarding the periodicity with which they are held every year or at any other regular interval. For these reasons, we chose RF and AO as the focus of this study.

Table 2. Entity to describe RF and AO, and application with CIDOC CRM

Entity of RF and AO	Entity class from CIDOC CRM
RF and AO	E1 CRM Entity
Web site	E42 Identifier
Category	E55 Type
Holding event	E5 Event
Holding place	E53 Place
Holding period	E52 Time-Span
Origin event	E5 Event
Source	E42 Identifier
Registration event	E5 Event

Table 3. Property to describe RF and AO, and application with CIDOC CRM (excerpt)

Property	Property used vocabularies	Example of properties' values
Title	rdfs:label	Akita no Kanto (秋田の竿燈)
Category	P2 has type	ifcp:category/practice ifcp:category/p5
Overview	P3 has note	Overview of RF and AO (竿燈は、長い竹竿に9本の横竹を結び、…)
Web site	P67 has created	http://www.kantou.gr.jp/index.htm
Origin event	P94 has created	Expressing entities for the origin event as a blank node
Registration event	P141 assigned	ifcp:84/register
Publication date	P170 define time	Expressing as a blank node, which includes two properties below
Official publication date	P141 assigned	3rd to 6th, August every year (毎年8月3日～6日)
Supplemental explanation	P3 has note	Currently held in August

CIDOC CRM: CIDOC CRM [7] is a documentation standard defined by the International Committee for Documentation (CIDOC), an international committee of the International Council of Museums. CIDOC CRM provides an ontology to manage museum information, adopting an object-oriented conceptual model based on the concept of ontology and focusing on museum information, while also being available for other general purposes.

CIDOC CRM is regarded as an event-centric model, which is a model describing several events that are related to one object. It is possible to construct a network of enriched entities by representing these events. For instance, it would be possible for one IFCP to be described in terms of when and where it occurred, who participated in it as well as holding, origin, and registration of its IFCP.

4.2 Definition of Events

This study defines three types of IFCP events.

The first type is the holding event, which describes when an RF, AO, or other IFCP-related event is held. In the case of Akita no Kanto, the festival is held annually, so the holding event is represented as Akita Kanto Festival 2019, Akita Kanto Festival 2018, and Akita Kanto Festival 2017. Each holding event has its title, overview, website URL, and categories as describing information.

The second type is the origin event, which relates to information regarding IFCP's origin. Information about IFCP's origin would be different depending on the referred date or each content, and is described differently. To make this situation possible, event information about its origin is required. This event concretely includes origin place, original period, source of reference (URL), its editor, and its edit date.

The third type is the registration event, which relates to event information as inscribed in the IIFCP in Japan. This event has a registered date and registration number.

5 Building and Publishing the Dataset

This section describes the flow from data collection to the release of the structured RF and AO information datasets (Fig. 2). The core IFCP data, including RF and AO, are acquired from the Important Cultural Properties Database [16] as a CSV file. The acquired data includes basic information about IFCP and their categories. Event information such as holding, origin, and external resources about RF, AO, and their metadata are collected and added to enrich the data. The collection of this enrichment information was conducted manually by the first author. We only collected event information for the years 2018 and 2019, and converted the data to Excel format.

We then used Poorman's Toolkit [17] to convert an Excel file to RDF/Turtle data using LOD. This flow on Fig. 2 shows the construction of the RF and AO datasets.

In order to publish IFCP’s LOD data, the RDF/Turtle data is made available on the web. We used ttl2html [18] to create data for web release and convert the RF and AO RDF/Turtle data to HTML files. The website including these files is already released to the public [19] and is updated according to revisions to the data model. This web site aims to publish developed datasets to the public.

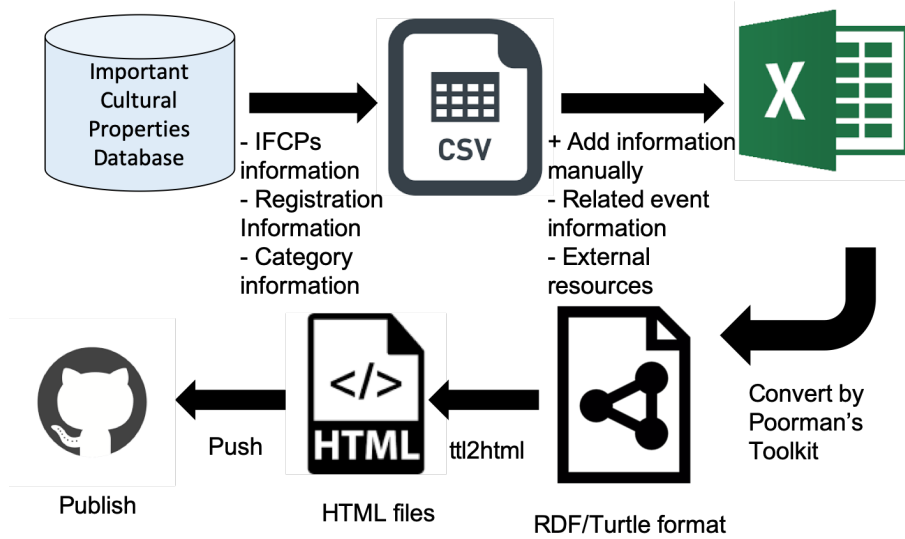


Fig. 2. Construction to release flow of datasets

6 Evaluation

We conducted an evaluation of our model and LOD dataset on RF and AO using the approaches described in the following section.

6.1 Functional Requirements and Prototype System

The first step was verification of the structured data model. The construction of the data model and datasets represented the minimum level of applicability. We further evaluated the functionality of the data model by implementing a prototype system to judge whether it performs as expected.

This prototype system aims to utilize RF and AO information. We supposed some use cases for this prototype system. For example, a user, who studies intangible cultural heritage, wants to find its related resources on the web, but it is difficult to locate credible information. We defined four functional requirements for utilization of the datasets based on the use cases as follows:

- I. Browsing RF and AO
This shows detailed information regarding all RFs and AOs recorded in the RDF store.
- II. Temporal information search
This shows RF and AO information for each period in which they were occurred.
- III. Spatial information search
It shows RF and AO information for each area in which they were held.
- IV. Holding information search
It shows RF and AO information for each occurrence year.

In implementing the prototype system, the RF and AO RDF/Turtle data are recorded in the RDF store (Apache Jena Fuseki [20]) along with regional data from the Statistical LOD of Japan. The RDF store makes it possible to retrieve recorded RDF data using the SPARQL query language. We integrated this retrieval result into the prototype system and implemented the above requirement functions.

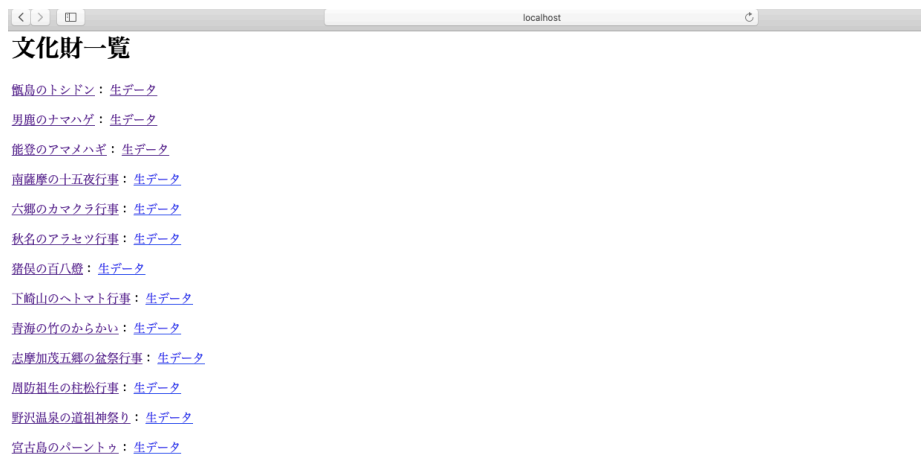


Fig. 3. Prototype system screenshot: RF and AO list example

Screenshots (Figs. 3 and 4) are from the prototype system based on the functional requirements. Figure 3 shows a list of RFs and AOs recorded in the RDF store. When one RF or AO is selected, the prototype system provides detailed information on it (Fig. 4). This flow allows users to browse every IFCP and fulfills the functional requirements.

6.2 Application Implementation

The second evaluation approach was the implementation of an application based on the implemented prototype system. This application verifies the feasibility of



Fig. 4. Prototype system screenshot: detailed information example

RF and AO utilization. This approach is still in progress. This application targets researchers and officers at local governments involving IFCP. We will conduct an evaluation experiment focused on this function and confirm its feasibility after the implementation of the prototype system.

7 Conclusion

This study aimed to structure a data model for building and publishing an IFCP dataset focused especially on RFs and AOs-related information and its release. For evaluation, we defined functional requirements and, based on these requirements, we implemented a prototype system using SPARQL based on the requirements.

We defined the data model based on basic elements, event entities, and CIDOC CRM vocabularies. We found that some properties do not fit perfectly with original definition of CIDOC CRM. We will extend the data model considering detailed descriptions.

In future works, we plan to implement, based on the prototype system, an application for the utilization of RF and AO information and then conduct a user experiment to verify its effectiveness.

References

1. UNESCO: Thirty-one new elements inscribed on the Representative List. <https://ich.unesco.org/en/news/thirty-one-new-elements-inscribed-on-the-representative-list-00327> Accessed 22 Sept

2. Agency for Cultural Affairs: Folk Cultural Properties in Japanese https://www.bunka.go.jp/english/policy/cultural_properties/introduction/folk/ Accessed 22 Sept
3. Cultural resources utilization section: Revision of Act on Protection of Cultural Properties (in Japanese). Monthly reports of the board of education **71**(5). 30–42 (2019)
4. Oshima, A.: Considering “Change” of Intangible Folk Cultural Properties with Emphasis on the Relation with the Designation of Cultural Properties (in Japanese). Research and Reports on Intangible Cultural Heritage (2), 228–214 (2008)
5. Tokyo National Research Institute for Cultural Properties Department of Intangible Cultural Heritage: 14th Conference on the Study of Intangible Folk Cultural Properties (in Japanese). <https://www.tobunken.go.jp/ich/wp-content/uploads/kyougikai2019.pdf> Accessed 24 Sept
6. Artese M. T. and Gagliardi I.: Cataloging Intangible Cultural Heritage on the Web, Progress in Cultural Heritage Preservation, vol. 7616, pp. 676–683 (2012)
7. CIDOC CRM. <http://www.cidoc-crm.org/>
8. EUROPEANA. <https://www.europeana.eu/en>
9. EUROPEANA DATA MODEL, <https://pro.europeana.eu/page/edm-documentation>
10. JAPAN SEARCH. <https://jpsearch.go.jp>
11. Tan, G., Hao, T., Zhong, Z.: A Knowledge Modeling for Intangible Cultural Heritage Based on Ontology. 2009 Second International Symposium on Knowledge Acquisition and Modeling, vol. 1, pp. 304–307 (2009)
12. Giannoulakis, S., Tsapatoulis, N., Grammalidis, N.: Metadata for Intangible Cultural Heritage - The Case of Folk Dance. Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Application, vol. 5, pp. 634–645 (2018)
13. Dannells, D., Enache, R., Damova, M.: A Multilingual SPARQL-Based Retrieval Interface for Cultural Heritage Objects. Proceedings of the ISWC 2014 Posters & Demonstrations Track a track within the 13th International Semantic Web Conference (ISWC 2014), vol. 1272, pp. 205–208 (2014)
14. T, Mishima.: Aggregation and Linking of Regional Cultural Heritage Information with the Use of Cultural Property Lists in Japan (in Japanese). Master Thesis, University of Tsukuba (2020)
15. Agency for Cultural Affairs: Cultural Properties https://www.bunka.go.jp/english/policy/cultural_properties/ Accessed 11 Jan
16. Important Cultural Properties Database (Kunishitei Bunkazai-tou Database). <https://kunishitei.bunka.go.jp/bsys/index>
17. Poorman’s Linked Data Toolkit. <https://github.com/jp-textbook/jp-textbook.github.io/wiki/Toolkit.en>
18. ttl2html. <https://github.com/masao/ttl2html/blob/master/README.md>
19. Intangible Folk Cultural Properties. <https://w3id.org/ifcp/>
20. Apache Jena Fuseki. <https://jena.apache.org/documentation/fuseki2/>