

# Metadata Model for Organizing Digital Archives of Tangible and Intangible Cultural Heritage, and Linking Cultural Heritage Information in Digital Space

**Chiranthi Wijesundara**

Graduate School of Library, Information and Media Studies,  
University of Tsukuba, Japan  
Library, University of Colombo, Sri Lanka  
*chiranthis@gmail.com*

**Shigeo Sugimoto**

Faculty of Library, Information and Media Science,  
University of Tsukuba, Japan  
*sugimoto@slis.tsukuba.ac.jp*

## ABSTRACT

*Background.* There are many digital archives in cultural domains, but there is no well-established metadata model which covers both tangible and intangible cultural heritage. Neither is there a well-established metadata model applicable to building digital archives by aggregating existing cultural heritage information.

*Objectives.* The objective of this study is to develop a metadata model for digital archives of diverse cultural resources and dispersed communities, where metadata aggregation from different sources is required. This paper presents a model called Cultural Heritage in Digital Environment (CHDE) for organizing various digital CHI organized as a digital archive.

*Methods.* The CHDE model is created as a generalized metadata model. It clearly distinguishes an intangible cultural heritage entity and its instantiation to model digital archives for intangible cultural heritage like those for tangible cultural heritage. The CHDE model is defined using the One-to-One Principle of metadata, and is used to clearly identify the relationships between metadata and the entities described by the metadata. Finally, this paper shows the application of the CHDE with use cases of cultural heritage objects.

*Results.* CHDE enables the organization of digital cultural heritage information related to tangible and intangible cultural heritage. It identifies the physical and digital information environment of cultural heritage, and further supports the modeling of digital archives built by aggregating cultural heritage information on the Web.

*Contributions.* This novel approach will benefit memory institutions which have insufficient resources to create digital resources such as those in South and Southeast Asia. Also, intangible cultural heritage organization through instantiation is a useful approach.

## INTRODUCTION

The past few decades have seen a boom in Web-based services of heritage related information. Large quantities of Cultural Heritage Information (CHI) are continuously produced and added to the Web by memory institutions and other public and private sector institutions. Memory institutions such as Galleries, Libraries, Archives and Museums (GLAM) are the main players that collect, organize, disseminate and preserve CHI in various cultural domains. A set of CHIs collected and organized for use is called a *digital archive of cultural heritage* in this paper. In addition to many individual digital archives in various cultural domains, there are portals which collect CHI from many memory institutions and provide integrated functions to find and access the collected CHI, for example, Europeana.<sup>1</sup>

Memory institutions use metadata standards to organize information about their holdings in accordance with their demands, for instance, heritage related metadata schemas and ontologies such as CDWA (Categories for the Description of Works of Art)<sup>2</sup>, LIDO (Lightweight Information Describing Objects)<sup>3</sup>, and CIDOC-CRM (Center for Intercultural Documentation-Conceptual Reference Model)<sup>4</sup>, information aggregation models like EDM (Europeana Data Model)<sup>5</sup>, and authority control standards such as Getty thesauri.<sup>6</sup> This variety arose from the heterogeneity of memory institutions and their requirements. This study proposes a metadata model which is neutral to any of these standards, and is designed as a comprehensive model to organize digital archives by using existing CHI.

Digital archives at memory institutions act as key mediators to disseminate cultural resources, and users get access to these digital resources via the digital archive systems. Digital archives at memory institutions usually have metadata about each item, similar to conventional catalog databases of memory institutions. On the other hand, there is a vast amount of information resources about cultural heritage provided by individuals and organizations of various types, for example, Wikipedia and tourism websites. These sites often have their own schemas to describe Cultural Heritage Objects (CHOs) and do not follow the metadata standards used by memory institutions. The digital archives provided by memory institutions tend to exhibit individual object information (i.e. item-centric information), whereas non-memory institutions provide CHI in a large granule, such as descriptions about historical and cultural contexts (e.g., the historical background of a CHO). Therefore, it is natural to consider that linking CHI provided by these two camps will be useful for users.

One motivation for this research is that many memory institutions in South and Southeast Asia are still not well-developed, and there are not many digital cultural archives built by memory institutions in the region. On the other hand, we can find rich CHI of South and Southeast Asia at digital archives provided by developed countries as well as Wikipedia and other cultural heritage websites. Thus, linking CHI in these sites seems beneficial for memory institutions in South and Southeast Asia (Wijesundara, Sugimoto, Narayan, & Tuamsuk, 2016). There are well-known efforts such as Europeana and Getty Research

---

<sup>1</sup> <https://www.europeana.eu/portal/en>

<sup>2</sup> [http://www.getty.edu/research/publications/electronic\\_publications/cdwa/](http://www.getty.edu/research/publications/electronic_publications/cdwa/)

<sup>3</sup> <http://network.icom.museum/cidoc/working-groups/lido/>

<sup>4</sup> <http://www.cidoc-crm.org/>

<sup>5</sup> <https://pro.europeana.eu/resources/standardization-tools/edm-documentation>

<sup>6</sup> <http://www.getty.edu/research/tools/vocabularies/index.html>

Institute that have tried to provide CHI in Linked Open Data (LOD) formats. Thus, LOD is a natural platform for linking information on the Web.

South and Southeast Asia have rich intangible cultural heritage, with some common heritage across the region, such as “Ramayana”.<sup>7</sup> Therefore, linking intangible cultural heritage through these commonalities is an essential process that should be done. CHI for intangible cultural heritage has some complex features compared with CHI for tangible cultural heritage. This is because intangible cultural heritage is very diverse across countries. Creating a common standard to capture these aspects is challenging. In addition, language barriers and socio-economic factors directly influence intangible CHI creation and dissemination. Unlike museum Web portals, there are only a handful of intangible cultural heritage portals, and they also do not adhere to a common standard.

This paper proposes a metadata model for building digital archives to solve the issues mentioned above. Although the study was motivated by the situation in South and Southeast Asia, the model is designed to be generally applicable and not region-specific. The model is designed to help connect heterogeneous CHI in separated data silos, enabling more accessibility and enrichment to the CHI. The authors propose the CHDE model which collects and organizes CHI associated with tangible and intangible cultural heritage. The model is designed to collect digital CHI on the Web related to the region, and aggregate them into a single entity called *Curated Digital Instance (CDI)*. The One-to-One Principle of metadata is used as a basis for aggregating metadata. The CHDE model is proposed to clearly model digitally curated objects of intangible as well as tangible CHOs. The model proposes an entity named *Instantiation* to model intangible cultural heritage entities which are to be digitally curated into a digital archive.

The CHDE model was first reported in by Wijesundara, Monika, and Sugimoto (2017). This paper offers a more elaborate discussion of the CHDE. It also explicates the CHDE model using a few use case scenarios, and presents some semantic relationships between CHDE and existing schema classes. Metadata aggregation is a core issue for the CHDE model. Actual metadata aggregation technology is not discussed in this paper, which focuses on the theoretical aspects of the CHDE model.

## **ADDING CONTEXT TO CULTURAL HERITAGE OBJECTS**

Table 1 defines the cultural heritage-related terms used in this paper. This section further discusses the characteristics of Web-based Cultural Heritage Information (CHI), issues of CHI aggregation, and implications for metadata models. The objectives and contributions of this study are also elaborated.

### **The Nature of Web-based Cultural Heritage Information**

Cultural heritage information (CHI), which may be institutional or non-institutional, can exist as a digital or non-digital instance. This research was focused on digital information. An institution may have CHI in non-digital or analog forms, for example, an artifact and a cultural performance recorded on a VHS tape and a printed/filmed photograph. Later, the memory institution may digitize these analog materials and organize them as a digital archive. Parallel to this, there can be born-digital recourses such as a VR image of cultural heritage. Lor and Britz (2012) discusses the same idea and explained that both the content that was created and disseminated in analog format and subsequently digitized, and content that was

---

<sup>7</sup> <https://www.britannica.com/topic/Ramayana-Indian-epic>

**Table 1. Definitions of terms**

<b>Cultural Heritage</b>	Defined by UNESCO as “the legacy of physical artifacts and intangible attributes of a group or society that are inherited from past generations, maintained in the present and bestowed for the benefit of future generations” [1]. Cultural heritage has many sub-facets and variations. This study uses the term <i>Cultural Heritage Object (CHO)</i> to denote the entire cultural heritage domain including tangible and intangible heritage assets.
<b>Tangible Cultural Heritage</b>	Defined by UNESCO as “objects significant to the archaeology, architecture, science or technology of a specific culture” [1]. Tangible cultural heritage can be movable or immovable such as a statue, a monument, a painting, or a pot. The term <i>tangible CHO</i> is used to refer to a tangible object in the physical world.
<b>Intangible Cultural Heritage</b>	This denotes the “traditions or living expressions inherited from our ancestors and passed on to our descendants”[2]. Intangible cultural heritage cannot be touched. It occurs at a given space (i.e. in the physical world) during a given time, enabling a physical existence for a while. Traditional dance performance, culinary art and handicraft comes under this category.
<b>Cultural Heritage Information (CHI)</b>	Defined in this study as “a description or representation of a certain heritage object which can be tangible or intangible heritage”. Catalog records at memory institutions, Wikipedia articles about cultural heritage, and bibliographic data provided by digital archives are typical CHI. In this study, digital surrogates of original CHOs are included in CHI (e.g., a digital image of Mona Lisa and a virtual reality (VR) image of Angkor Wat). CHI includes metadata such as administrative data and external/additional information about the CHO. CHI may or may not be digital. However, this research focuses on digital CHI on the Web. As a digital surrogate is created from the original/primary CHO, it is a secondary instance and a sort of metadata. On the other hand, digital surrogates which are born digital or converted into digital can be recognized as primary CHOs in a digital archive. For digital archives of intangible cultural heritage, there exists no original object for a digital resource. Therefore, records created for the performances of intangible cultural heritage are digital surrogates of <i>Instantiations</i> of an intangible cultural heritage entity.
<b>Metadata</b>	Generally defined as “data about data”. Any CHI is a metadata. Digital surrogates such as a photograph of Mona Lisa and a VR image of Angkor Wat are a kind of metadata based on this definition. This paper, however, uses metadata in a slightly narrower sense—textual description of a CHO, that is, CHI expressed in a textual form such as plain text, XML texts, Excel sheets, relational databases, etc. Can be divided into institutional and non-institutional metadata.
<b>Institutional Metadata</b>	Defined as “metadata descriptions stored in the form of catalog records which are created, maintained and hosted by memory institutions can be considered Institutional Metadata”. Generally created for a single item collected by a memory institution (i.e. <i>Item-Centric</i> metadata), to support search and access.
<b>Non-Institutional Metadata</b>	Defined as metadata not created by a memory institution, for example Wikipedia articles and tourism websites. Can exist at both item-centric and thematic (subject) levels. For instance, a Wikipedia article can describe a single heritage object or a specific theme. Non-institutional metadata are mostly not based on well-standardized metadata schemas. However, some non-institutional metadata follow their own data standards (e.g., UNESCO Intangible Heritage Lists [3]).
<b>Digital Archive</b>	Defined as “a collection of digital resources selected, collected, organized and maintained for long-term use” (Sugimoto, 2014). Acts as a portal to disseminate information (mainly item-centric information) provided by a memory institution. The content of digital archives may be born-digital or digitized resources. In this paper, metadata of digital archive resources fall into the category of institutional metadata.

**Notes:** [1] <http://www.unesco.org/new/en/cairo/culture/tangible-cultural-heritage/>  
 [2] <https://ich.unesco.org/en/what-is-intangible-heritage-00003>  
 [3] <https://ich.unesco.org/en/lists>

created and disseminated digitally (born-digital) can be identified as digitized content.

CHI on the Web means the digital resources which are published on the Web. They can exist as websites, databases or portals. They can have different formats, yet all belong to the digital form. According to the Digital Curation Centre, a *digital object* can be a *simple digital object* (e.g., text or sound file along with some metadata) or a *complex digital object* (e.g., discrete digital objects made by combining a number of other digital objects, such as websites or databases) (Higgins, 2008). Ultimately, all these digital objects are published on the Web through digital archives or websites.

Memory institutions play the main role of disseminating Web-based CHI to users through their Web-accessible digital archives. Their catalog records are typical institutional metadata. According to Zeng and Qin (2016, p. 93), “when discussing levels of granularity, the term “item-level” (in contrast to “collection-level”) is often used to refer to the individual objects that, together, constitute a whole collection”. Somehow, this item-level concept came from the long tradition of management of museum holdings. The item-centric perspective is advantageous for knowledge organization, allowing easy search for and access to information relating to individual items. However, user needs are not always centered on an individual resource: users may need contextual information is not item-centric, such as found in a Wikipedia article. This research tries to connect item-centric institutional metadata and non-institutional metadata, aiming to provide users with more linked and enriched form of CHI.

On the other hand, “hybrid records” (Woodley, 2016) of CHI provided by digital archives are confusing as to the purpose of the objects and their metadata. As a result, users sometimes cannot distinguish between the original objects and their digital surrogate information in a digital archive. Therefore, identification of objects, objectives and their metadata are very important to collect metadata from different sources and aggregate them into a single database.

### **Metadata Models and CHI Aggregation**

This study aims to define a conceptual metadata model for building digital archives of heterogeneous CHOs by aggregating CHI collected from various sources. Therefore, metadata aggregation is a key aspect of this research. Metadata schemas of digital archives are often defined based on well-known standards and best practices in order to ensure the quality, consistency, and interoperability of data which is an important factor when it comes to metadata aggregation (Gilliland, 2008).

There are well-known models for metadata aggregation such as Open Archives Initiative-Object Reuse and Exchange (OAI-ORE)<sup>8</sup> and Europeana Data Model (EDM) (Isaac, 2013). However, there are several fundamental facets which have to be taken into account for metadata aggregation: identification of objects described by metadata (i.e., objectives of metadata description); identification of purposes of metadata description such as administrative, technical, and descriptive; underlying data models and metadata schemas of metadata instances to be aggregated which define syntactic, structural and semantic features; and metadata interoperability schemes to make metadata aggregable such as vocabulary mapping. This section discusses these facets primarily from the viewpoint of building digital cultural archives using CHI resources available on the Web.

A simple example of identification of objectives is identifier assignment to CHOs. On one hand, memory institutions have identification schemes for physical items curated as

---

<sup>8</sup> <https://www.openarchives.org/ore/>



cultural heritage. Instead, digital archives need to identify every digital surrogate created from a particular CHO. Compared with tangible objects, identification of intangible cultural heritage may not be so simple. Identification of objectives is related to the identification of purposes. For example, descriptive metadata about a CHO is given to find and access the object, descriptions about stakeholders of the object may be used to know the roles of the stakeholders, and so forth. Thus, identification of objectives and purposes are crucial. However, in many cases of conventional metadata schemas, a single metadata record describes more than one objective, and it is not obvious which part of metadata describes what objective. From this point, this study considers that One-to-One Principle of metadata is a useful underlying concept to utilize in this model. The One-to-One Principle is a “principle whereby related but conceptually different entities for example, a painting and a digital image of the painting, are described by separate metadata records”.<sup>9</sup>

The Dublin Core Application Profile (DCAP) (Coyle & Baker, 2009) provides a framework for metadata interoperability, which is also useful for metadata aggregation. According to DCAP, an application profile defines a metadata schema where a metadata instance consists of one or more *Descriptions*, and each *Description* is a set of *Statements*. Metadata terms used in the *Statements* should be defined separately from an application profile. Metadata of digital CHO is usually complex, consisting of *Descriptions* of different objects such as an original CHO, its digital surrogates, stakeholders, access conditions, and so forth. Hence, DCAP provides a framework to identify the objectives of descriptions. Clear identification of the objects described by metadata is the fundamental basis for aggregation.

Metadata aggregation is a well-known function to build digital archives. For example, OAI-ORE defines a generalized model to collect and aggregate metadata. In addition, EDM defines a data model for Europeana to organize metadata about cultural heritage resources collected from participating institutions. Both models are defined primarily for digital objects which are organized with conventional institutional metadata. These models define data models of aggregated metadata, but not the metadata creation and aggregation process. The metadata creation process depends on factors such as tangible or intangible cultural heritage, while born-digital or digitized objects may affect the metadata aggregation process.

The CHDE is defined using the Entity-Relationship Model which is used as the base model in the Resource Description Framework (RDF). Realizing the CHDE in RDF is described later. This paper proposes a conceptual metadata model as a foundation for CHI organization and aggregation of heterogeneous data sources. According to scholars, information aggregation enables, data interoperability, global view to diverse information contents, semantic searching, link and share content, data enrichment, data reuse and longevity of information, etc. (Hyvönen, 2012). The proposed model can be utilized to create such aggregation, enabling many functionalities required for digital archives of tangible and intangible cultural heritage of South and Southeast Asia.

### **Requirements for Metadata Model—Objectives and Novelty of the Research**

Organizing and connecting dispersed CHI into a single platform using a metadata model is the main goal of this research. This study proposes a metadata model to collect diverse CHI on the Web which we term the CHDE model. CHDE explicitly distinguishes the physical and digital spaces/environments of CHOs which cannot be clearly seen in the existing CHI aggregation or organization models.

---

<sup>9</sup> <http://www.dublincore.org/documents/usageguide/glossary/>

Connecting and disseminating information on the Web is not a brand-new idea. However, a metadata model that adheres to the One-to-One Principle provides a clear foundation for connecting and aggregating heterogeneous CHI on the Web, which will help memory institutions in South and Southeast Asia to build digital archives of their regional cultural heritage.

This study was started with two fundamental research questions which arose from discussions of digital archives of cultural heritage for South and Southeast Asia:

1. How can we describe intangible cultural heritage for digital archives? Conventional digital archives provide digital records of intangible cultural heritage such as dance performance and music performance. However, a single performance is not a particular intangible cultural heritage entity, but we have to use performance which is physically shown to record intangible cultural heritage.
2. How can we model metadata for digital object to be created by aggregation of fragments extracted from existing digital archives and other Web resources?

The main aim of this paper is to present the CHDE model and to discuss it based on these research questions. The paragraphs below present the basic requirements for the CHDE model:

1. Identification of “object” of intangible cultural heritage: What objects of intangible cultural heritage archived in a digital archive has to be identifiable? This is a fundamental requirement for digital archives and databases which have to record temporal objects such as events, performance, installations, and so on. This requirement is crucial even in the case of tangible cultural heritage because they may be lost by natural and/or man-made disasters.
2. Identification of objectives of metadata aggregation, using the One-to-One Principle of metadata as a foundation: It is essential to identify metadata description objectives individually, that is, identifying the original CHO and its surrogates as separate entities. However, current digital archives and other websites tend to provide mixed descriptions of more than one object. Clear identification of the relationships between metadata description and its objective is important for metadata aggregation. The One-to-One Principle of metadata concept is a reasonable foundation to satisfy this requirement. We shall clarify how the One-to-One Principle can be applied to cultural heritage resources on the Web.
3. Identification of facets required for aggregation of Institutional and Non-institutional CHI: It is important to provide information about a CHO together with contextual information about the object for users who want to learn about the object and its cultural contexts. This linkage between the two types of information can be realized by linking item-centric institutional CHI and general-description-oriented non-institutional CHI. Basic requirements for linking this very different CHI need to be clarified.
4. Identification of “Entities” in the Process of Organizing CHI: It is essential to identify entities which work in the process of organizing CHI into a digital archive. The process models (Figure 2) help us identify those entities and understand how we can create, collect and aggregate CHI in accordance with those entities.

## LITERATURE REVIEW: CONCEPTS, MODELS AND RELATED STUDIES

### Related Models and Concepts

There are a few existing data models related to CHI aggregation. First, the EDM acts as a typical CHI aggregator in the cultural heritage domain. EDM aggregates metadata over 3,000 cultural heritage institutions all over the European Union and enriches them further for better accessibility. The aggregated content is disseminated via the Europeana Collections which hosts nearly 6 million artifacts, books, films, music information, etc. as digital content. According to the EDM primer, “EDM is not built on any particular community standard but rather adopts an open, cross-domain Semantic Web-based framework that can accommodate the range and richness of particular community standards such as LIDO for museums, EAD for archives or METS for digital libraries” (Isaac, 2013, p. 5). The model is created using RDF and uses classes (e.g., *edm:ProvidedCHO* for provided CHOs) and properties (e.g., *edm:hasView* for one or more resources that are digital representations of the provided object), plus RDF syntax in Turtle to describe their model systematically.

The second prominent model is CIDOC-CRM by the International Council of Museums (ICOM). CIDOC-CRM “provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation”.<sup>10</sup> This model takes an ontological approach, and it consists of a comprehensive set of classes and properties to describe artifacts, monuments and any form of cultural heritage entity (Boeuf, Doerr, Ore, & Stead, 2015).

IFLA based Functional Requirements for Bibliographical Records object-oriented (FRBR<sub>OO</sub>) which is known as an extension of CIDOC-CRM, is another important model in the CHI domain. FRBR<sub>OO</sub> basically uses the Work, Expression, Manifestation, and Item (WEMI) entities of FRBR and adds some more classes and properties aligned with CIDOC-CRM. The main aim of FRBR<sub>OO</sub> is to integrate bibliographical information related to cultural heritage and facilitate library and museum information harmonization (Chryssoula, Doerr, Le Bœuf, & Riva, 2015).

The One-to-One Principle of metadata by the Dublin Core Metadata Initiatives (DCMI) is used as a foundation concept in the CHDE Model. The idea that related but conceptually different entities should be described by separate metadata records is very important when identifying certain CHI entities like original heritage object or their digital surrogate metadata. Miller (2010) discussed the benefits and challenges of using the One-to-One Principle in the cultural heritage sector. Although we assume that the information provided by institutions are One-to-One, it is not happening inconsistently. According to Woodley (2016, p. 44), “although some metadata standards (e.g., Dublin Core) follow the principle of a one-to-one relationship between a metadata record and an “item”—be it analog or digital—in practice many memory institutions use a single metadata record to record information about an original object as well as its digital surrogate, thus creating a sort of “hybrid” record. When migrating and harvesting data, this may pose problems if the harvester cannot distinguish between the elements that describe the original item and those that describe the digital surrogate”. The proposed CHDE provides a solution to overcome this “hybrid” record issue through the One-to-One Principle of metadata.

---

<sup>10</sup> <http://www.cidoc-crm.org/>



## Related Studies

Various studies have been conducted on data models for CHI organization and aggregation. One study examined collection level CHI aggregation based on Europeana Data Model (EDM). Since collection level data aggregation is not very prominent in the current practice, the authors proposed a set of general requirements for the representation of collections in digital aggregation systems (Wickett, 2014).

Smiraglia (2005) tried to model artifacts in museums using *Work* concept in the FRBR-WEMI model. He used Etruscan artifacts from the University of Pennsylvania Museum of Archaeology and Anthropology to demonstrate the connection of the *Work* concept in non-documentary artifacts. The model he presented, called the Content Genealogy Model, tries to conceptualize the representations of museum artifacts and their metadata. Both of these studies are focused on work level CHI aggregation, but used different approaches.

Carboni and Luca (2016) analyzed the dichotomy between tangible and intangible heritage, and proposed a way to document the same. They used CIDOC-CRM as the base, and modeled information using a use case to show that a cultural object has multiple facets and dimensions that incorporate both tangible and intangible elements.

Some scholars have proposed their own models for organizing and describing heritage resources. Amin, Baker, Deraman, and Yatim (2012) proposed one such model called a knowledge repository model for intangible cultural heritage as a framework and guideline to archive Malay Intangible Culture Heritage in Malaysia. The proposal was to digitize intangible to tangible heritage without losing their originality, and archive them using the proposed model.

The above related studies followed different approaches to aggregate CHI, some of which are item-centric information aggregation, and some are focused on specific kinds of CHOs. The proposed CHDE model tries to clearly identify the physical and digital objects via the One-to-One Principle of metadata. In addition, CHDE proposes a novel information organization approach specifically for intangible cultural heritage.

## Discussion of Existing Models

This section discusses some features related to the existing CHI models—EDM, CIDOC-CRM and FRBR<sub>OO</sub>—in contrast to the proposed CHDE model.

As mentioned earlier, digital archives of memory institutions organize CHI records as individual items. Data models such as EDM aggregate digital surrogates related to a heritage entity. This aggregation is item-to-item information aggregation. EDM is considered a top-down approach where data providers submit their CHI as packages to the Europeana in a form conforming to EDM. According to EDM, *Mona Lisa* by *Leonardo da Vinci* is a single item (known as *edm:ProvidedCHO*) and its digital images may be collected from more than one institution. These are collected using *edm:hasView* and *edm:aggregatedCHO* properties (Isaac, 2013). This method adheres to the One-to-One Principle of metadata concept in the CHDE model. However, the CHOs provided by these institutions are not 100% reliable. Although they represent individual heritage items, the provided information sometimes consists of a mixture of original heritage information and digital surrogate information in a single record. This mixed nature of CHI records and difficulty of understanding the objectives of metadata can be further clarified as follows. According to Orgel, Höffernig, Bailer, and Russegger (2015), EDM lacks providing provenance information related to the annotations they created for the digital cultural heritage in multiple views. Another paper states that the “distinction between *provided objects* (painting, book, movie, etc.) and their *digital*

*representations*” is a core principle in the EDM. However, “although the Europeana core classes stress the difference between the provided object (*edm:ProvidedCHO*), that is, the “real object”, and its digital representation (*edm:WebResource*), that is, its Web resource, sometimes this difference is not evident at all in the aggregated metadata exposed to the final user, generating confusion. Sometimes the description seems to be addressed to the electronic version, some other to the original work, without a clear distinction” (Peroni, Tomasi, & Vitali, 2012).

Untangling the mixed nature of the CHI records and distinguishing CHI according to their objectives are the focus of this paper. As South and Southeast Asian memory institutions do not have a common basis to share and aggregate CHI, CHDE uses a bottom-up approach for metadata aggregation which relies on the existing Web-based CHI provided by memory institutions and non-institutions. A fundamental issue in metadata aggregation is to identify every single object to be described by a metadata record or presented as a CHI.

In contrast, neither EDM nor CIDOC has specific entities designed for expressing intangible cultural heritage. Since an intangible cultural heritage entity is not an item physically collectable by memory institutions, the item-centric resource aggregation is not suitable. The CHDE model provides a solution via *Instantiation* as a bridge to aggregate those resources related to intangible cultural heritage. This *Instantiation* acts as a specific aggregator which forms a collection/set of resources related to a specific intangible cultural heritage entity.

We have previously mapped CHDE classes to those in CIDOC-CRM and FRBR<sub>OO</sub> to check the compatibility of the models. From those mappings, we realized that CIDOC-CRM and FRBR<sub>OO</sub> do not provide an underlying model to explicitly express the relationships between digital surrogates and their source objects which may be tangible or intangible. Besides, despite the vast ontology provided by CIDOC-CRM, it is difficult to find appropriate classes which clearly represent digital surrogates of a CHO. Similarly, these existing models have no definite classes to represent intangible cultural heritage entities (Wijesundara, Monika, & Sugimoto, 2017).

## **CHDE: AN APPROACH TO ORGANIZE CHI ON THE WEB**

As a solution to the problems discussed above, we designed the CHDE model as a conceptual model for CHI resource aggregation. This section details the CHDE model, which was previously presented in Wijesundara, Monika, and Sugimoto (2017).

### **CHDE Model**

Figure 1 represents the proposed CHDE (Cultural Heritage in Digital Environment) model. The model defines entities and their relationships between CHI resources, and explicitly differentiates the *Physical* and *Digital Space* of CHI. Memory institutions mainly collect resources that are realized in the physical environment, and further digitize them in the *Digital Space*. This situation is applicable to both tangible and intangible cultural heritage. CHDE defines separate metadata for each of these instances based on the One-to-One Principle of metadata.

Figure 1(a) and (b) present the CHDE model for tangible and intangible cultural heritage, respectively. The entities which constitute the CHDE model are defined below.

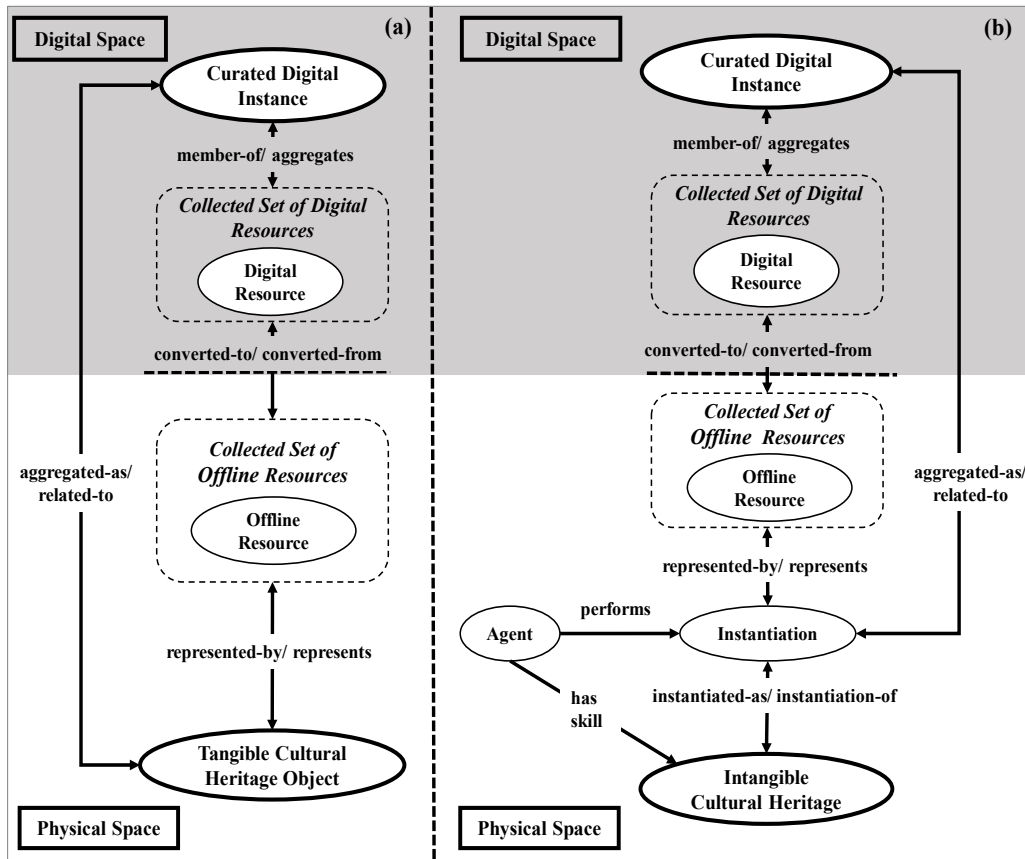


Figure 1. CHDE model describing tangible and intangible CHI aggregation

Entities that appears in both Figure 1(a) and (b) are as follows:

- *Offline Resources*: a tangible CHO or *Instantiation* of intangible cultural heritage can be recorded by the memory institutions into various non-digital resources such as image (printed photograph), sound (audio tape) or textual formats (printed book). In addition, there can be readily available digital resources such as an image file on a CD or a USB flash drive. Since both of these resources are not connected to the networked environment, they are named *Offline Resources* in this model.
- *Collected Set of Offline Resources*: One or more of these *Offline Resources* (corresponding to a tangible cultural heritage object or *Instantiation of an intangible cultural heritage*) can be identified as a *Collected Set of Offline Resources*
- *Digital Resources*: *Offline Resources* which can be converted into *Digital Resources* or utilized as they are (if they are already in a digital format) in the *Digital Space*. For instance, a printed photo or an audio record on a tape can be converted to JPEG or MP3 format in the *Digital Space*, while a JPEG image on a CD can be used as it is without any conversion. In addition, there can be purely born-digital materials such as games or animations, which are readily available as *Digital Resources* in the *Digital Space*.
- *Collected Set of Digital Resources*: Subsequently, one or more of these *Digital Resources* can be identified as the *Collected Set of Digital Resources*.
- *Curated Digital Instance*: The topmost circle denoted *Curated Digital Instance (CDI)* acts as the aggregated set of digital resources corresponding to the tangible cultural heritage object/ *Instantiation* of intangible cultural heritage entity at the bottom. This

entity may include one or more digital resources and their metadata descriptions corresponding to a particular tangible cultural heritage/*Instantiation*.

Entity that appears only in Figure 1(a):

- *Tangible Cultural Heritage Object*: This represents a touchable heritage object in the physical environment. These tangible objects can be housed in a museum or can be positioned as an immovable monument.

Entities that appear only in Figure 1(b):

- *Intangible Cultural Heritage*: Intangible cultural heritage which represents a conceptual entity might be a performance, ritual, skill, etc. This intangible cultural heritage has to be performed during a given time and location, and once it has occurred only the performance can be captured by any medium. This situation is different from tangible cultural heritage in the *Physical Space*.
- *Instantiation*: Unlike tangible cultural heritage, intangible cultural heritage does not exist as sole physical items and cannot be represented as items. However, intangible cultural heritage entity can have many *Instantiations*. For example, a traditional dance performance can be performed in many places. Therefore, each performance is an *Instantiation* of the corresponding intangible cultural heritage entity. A specific intangible cultural heritage is represented by a combination of *Instantiations* and their associated resources. These intangible cultural heritage *Instantiations* can be based on *temporal, location, category, agent, activity* or *concept* associated with the intangible cultural heritage entity (Wijesundara, Monika, & Sugimoto, 2017). In addition, once an *Instantiation* has been created, it can be used to connect tangible CHOs which are interrelated with the intangible cultural heritage.
- *Agent*: The *Agent* entity (e.g., person or group of people) is associated with an *intangible cultural heritage* entity, and also has a relationship with an *Instantiation*. In some situations, tangible CHOs too can have an *Agent*, but in this discussion, the authors apply it only to intangible cultural heritage and their *Instantiations*.

Furthermore, CHDE identifies two main resource environments, *Physical* and *Digital Spaces*.

- *Physical Space*: All the resources and entities that exist and occur in the physical environment (without any connection to the networked environment) belong to this category, including all *Offline Resources, Tangible CHOs*, and intangible cultural heritage and their *Instantiations*. Tangible cultural heritage exists as physical objects which humans can touch and see, and they can be recorded in physical mediums such as photographs and videos. An intangible cultural heritage entity has no physical existence as it is an abstract entity, and we can see intangible cultural heritage only through human activities in the physical space, such as dance performance and craftsmanship performance, which are called *Instantiations* of intangible cultural heritage. The CHDE model includes tangible CHOs, *Instantiations* and their recordings in the *Physical Space*.
- *Digital Space*: All the *Digital Resources* (converted from *Offline Resources* or born-digital) plus *CDI* (created from aggregated *Digital Resources* corresponds to a CHO on the networked environment comes under this category. *Digital Space* has no tangible-intangible cultural heritage differentiation. According to this paper, all CHIs belong to the *Digital Space*.

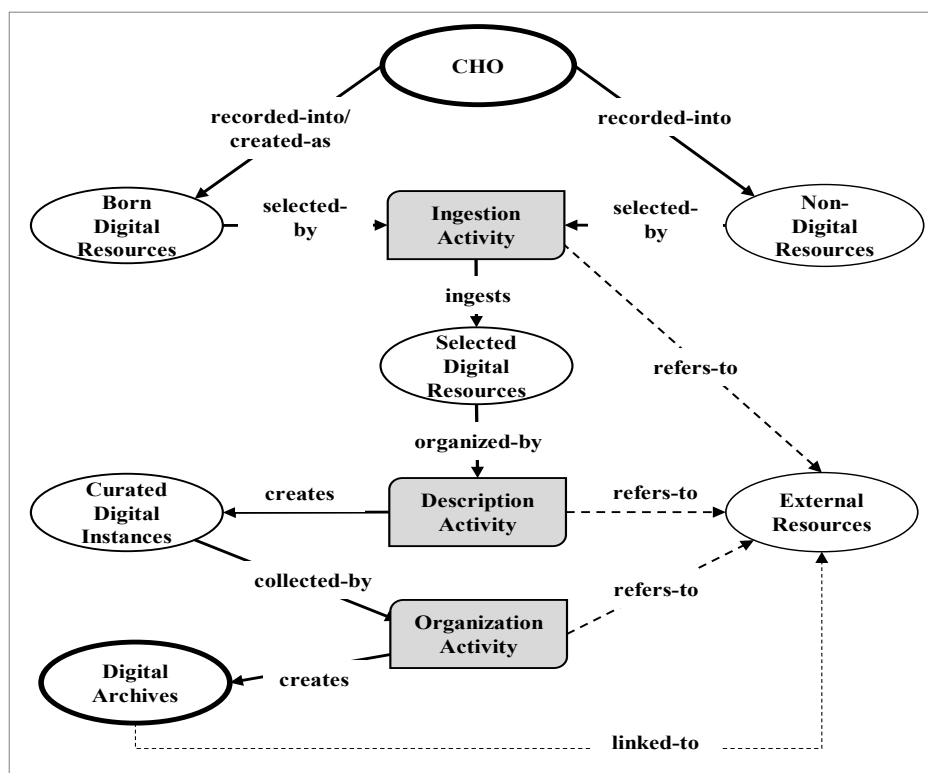


Figure 2. CHDE curation process

### CHDE Curation Process

Digital curation has emerged as a new inter-disciplinary practice that seeks to set guidelines for disciplined management of information. For example, the Digital Curation Center (DCC) is an internationally recognized institution which is involved in R&D activities related to the digital curation domain. According to the DCC, “digital curation involves maintaining, preserving and adding value to digital research data throughout its lifecycle”.<sup>11</sup> They have introduced a model called DCC curation lifecycle model which facilitates a graphical high-level overview of phases need for successful curation and preservation of digital object from its start to the end (Higgins, 2008). The same model was further extended by the Digital Curation Unit at the Athena Research Centre in Greece (Constantopoulos, Dallas, & Androutsopoulos, 2009). The model shown in Figure 2 is designed based on these general digital curation process models.

This paper focuses on a few key phases in the process shown in Figure 2. The oval labeled as CHO denotes the intangible and tangible heritage entities in the physical world. CHOs can be recorded into a carrier which may or may not be digital, where tangible cultural heritage can be recorded directly, and intangible cultural heritage can be recorded as their *Instantiations*.

Figure 2 consists of three main *Activities* (gray rectangles) and *Resources* (ovals). In the *Ingestion Activity*, the *Non-Digital Resources* of physical carriers are digitized, and it also imports digital resources such as *Born-Digital* from various places (e.g., digital portals).

<sup>11</sup> <http://www.dcc.ac.uk/digital-curation/what-digital-curation>



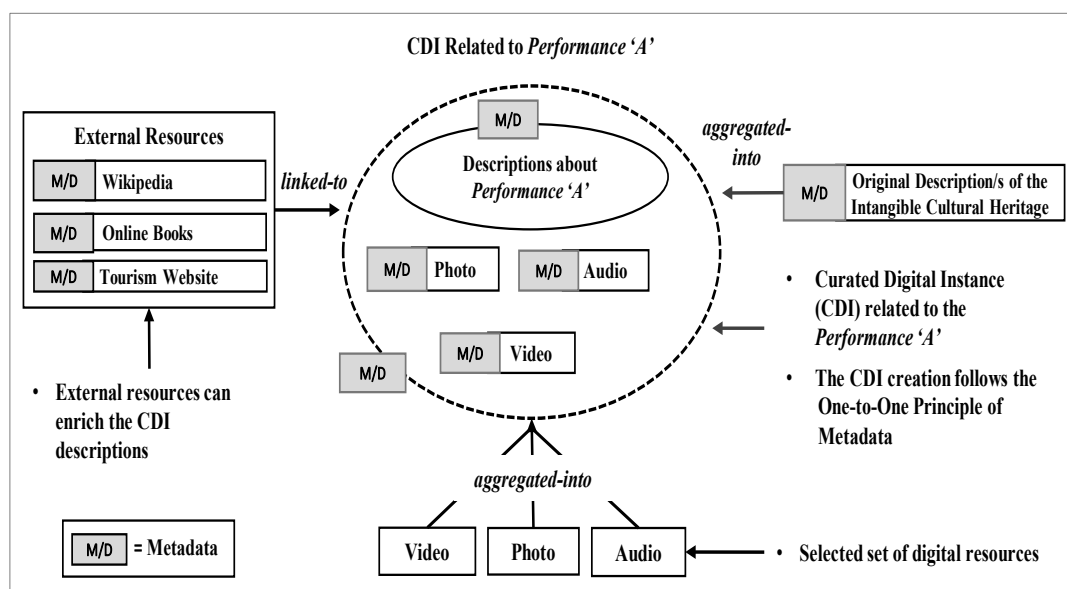


Figure 3. Curated Digital Instance (CDI)

After the *Ingestion Activity*, the resources are selected according to the institutional and user requirements. The digital resources are selected depending on various factors such as the usefulness of the content, data capacity, reliability, cost, etc. Then, the digital resources are given contextual information as a part of its metadata in the *Description Activity*.

CDI is an essential output of the *Description Activity*. A single CDI has to contain a non-empty set of digital resources. The CDI acts as a container for those context-rich digital resources related to a CHO in the physical world. This idea will be further described later.

In the *Organization Activity*, CDIs are stored in a *Digital Archive* based on the organization policy of the archive. During all three *Activity* phases (*Ingestion*, *Description* and *Organization*) *External Resources*, which might be online or offline, may be used as an additional information resource to understand the CHO and to enrich the contextual information of the digital resources. A *Digital Archive* is created as the final output of the whole process. As a whole through the entire process, the digital resources (i.e. CHIs) are created, contextualized and organized for dissemination.

### Curated Digital Instance (CDI)

The CDI in the CHDE Model acts as an aggregated instance in the *Digital Space*, and is created from CHOs in the *Physical Space*. As discussed above, a *Digital Archive* in the CHDE process is a collection of these *Curated CDIs* which is created according to the One-to-One Principle of metadata concept (Figure 3). This paper defines a CDI as “a collection of digital resources and their descriptions representing a single CHO.” A CDI is not a single CHI instance but a collection of CHI instances.

As a whole, Figure 3 depicts a CDI instance representing an intangible cultural heritage entity labeled *Performance ‘A’*. The CDI aggregates various digital objects/resources (e.g., a video, a photograph, and an audio record), each of which should be given a metadata record. Similarly, the instantiation of intangible cultural heritage, which is denoted as *Performance ‘A’* has its own metadata (i.e., description about the performance). *External Resources* are vital when identifying and enriching the CDI via non-institutional resources. Moreover, *Original Descriptions* are given to the intangible cultural heritage in the physical environment

also can be aggregated into the CDI. Each of these resource components has its own set of metadata, and a CDI itself has a separate metadata record. Identification of individual metadata descriptions related to individual resources is based on the One-to-One Principle of metadata. This differentiates individual objects separately without any confusion with other resources/objects.

## USE CASES

This section describes a few use cases of the CHDE model applied to real-world examples involving tangible and intangible cultural heritage. The use cases are illustrated in Figure 4 and 5 respectively.

Figure 4 shows a tangible CHO known as the *Statue of Tara*, which is a bronze artifact of the 8<sup>th</sup> Century AD originated in Sri Lanka, and now housed at the British Museum.<sup>12</sup> This statue can be recorded directly into physical media such as a printed photograph, a video recording or a printed article which consists of some background information of the object. Further, these *Offline Resources* can be converted into *Digital Resources* in the *Digital Space*. For instance, VHS recording can be converted into a YouTube video which can be in MOV, MP4 or WMV format. In addition, there can be resources which are already in digital formats which exist as *Offline Resources* and can be directly collected in the *Digital Space*. Finally, all these collected resources are aggregated as a CDI which represents the *Statue of Tara* as a single comprehensive unit. Hence, resources related to the tangible cultural heritage artifact can be identified as separated CHIs. As a result, it supports the possibility of identifying metadata descriptions individually. Since it collects and aggregates institutional and item-centric metadata from the British Museum with non-institutional metadata such as YouTube video or BBC article, it makes the CDI more context-rich as well.

Figure 5 shows an intangible cultural heritage entity known as the *Kandy Esala Perahera* festival in Sri Lanka. This can be identified as a major intangible heritage related to Sri Lanka which occurs annually, and it may be associated with many *Agents* such as dancers, performers, singers, etc. An intangible cultural heritage entity is frequently accompanied by various tangible CHOs such as props and instruments, which are excluded in Figure 5 to keep the figure simple.

Generally, these tangible CHOs are utilized during some intangible cultural heritage activities (e.g., props and masks used during Thai *Khon* performances), and some tangible CHO is produced during an intangible cultural heritage activity such as traditional craftsmanship (e.g., hand-dyed cotton and silk garments known as *Indonesian Batik*).

*Kandy Esala Perahera* festival is a tradition which is a conceptual entity. The local community carries out the festival every year, which is modeled as an *Instantiation* in the CHDE model. For example, the *Performance in 2016* is a single intangible cultural heritage *Instantiation*. This specific instantiation is the physical existence of the *Kandy Esala Perahera* which can be recorded into physical and/or digital resources. Thus, *Kandy Esala Perahera* can be instantiated at a particular time and place, and by/with particular agents. The *Performance in 2016 Instantiation* of *Kandy Esala Perahera* can be captured by different physical mediums such as a printed photo or a VHS tape. Also, there can be a printed performance schedule which describes the *Performance in 2016*. All these resources are not connected to the networked environment and exist as *Offline Resources* in the *Physical Space*.

---

<sup>12</sup> [http://www.britishmuseum.org/research/collection\\_online/collection\\_object\\_details.aspx?objectId=251954&partId=1](http://www.britishmuseum.org/research/collection_online/collection_object_details.aspx?objectId=251954&partId=1)

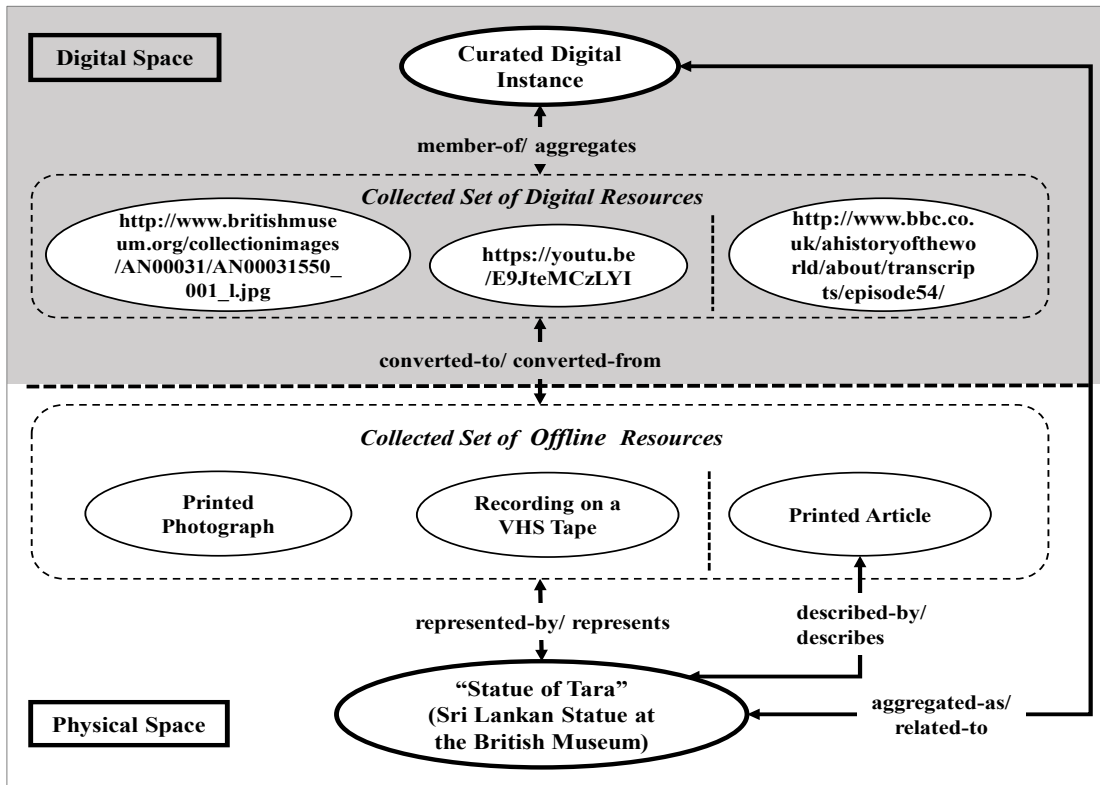


Figure 4. CHDE model: Tangible cultural heritage aggregation

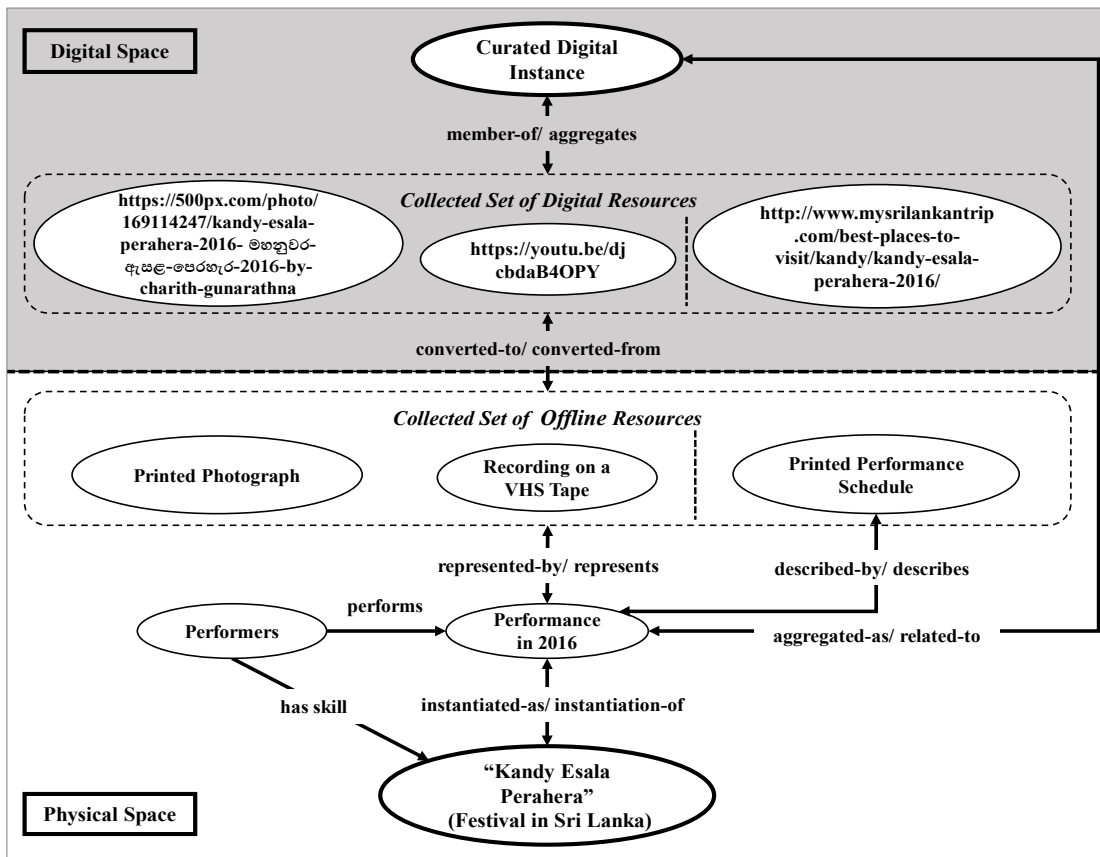


Figure 5. CHDE model: Intangible cultural heritage aggregation

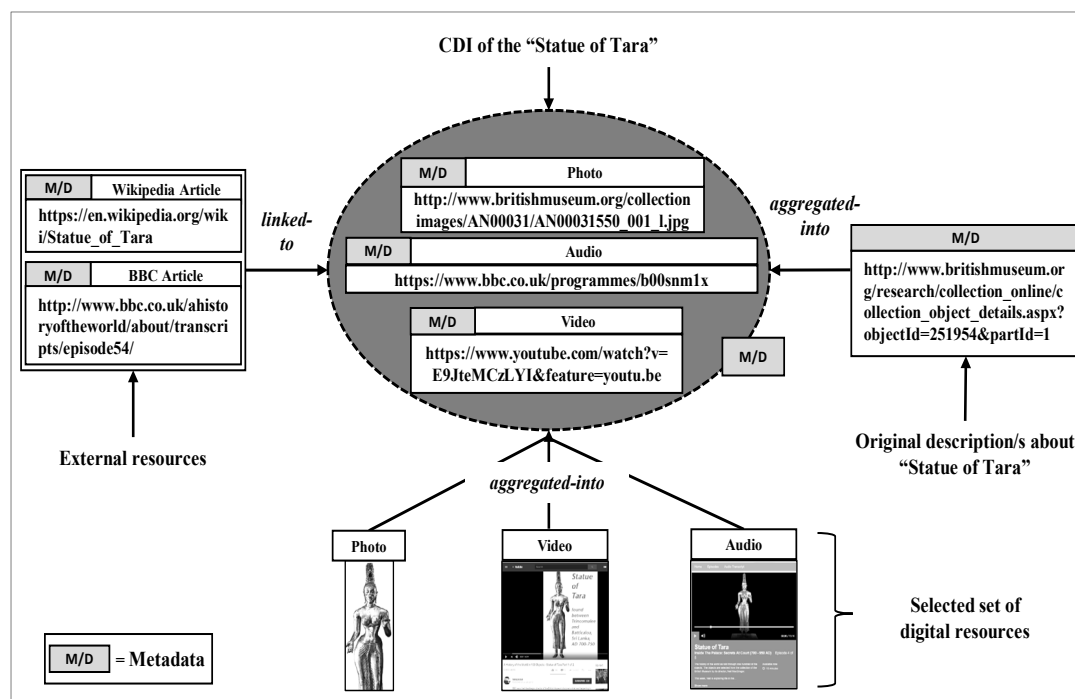


Figure 6. Example depicting the curated digital instance (CDI)

Later, these *Offline Resources* are converted into *Digital Resources*, such as a JPEG image and MPEG file on a website. Similarly, Printed Performance Schedule can be converted into an HTML webpage. Finally, all these collected *Digital Resources* are aggregated under the CDI which represents the *Performance in 2016*.

The CDI entity in the model acts as a container of aggregated resources and corresponds to a CHO in the *Physical Space*, that is, a tangible CHO or an *Instantiation* of an intangible cultural heritage entity. Figure 6 shows an example of the aggregation in the *Digital Space*.

Starting from the bottom, the photo, video and audio instances are digital resources related to the artifact *Statue of Tara* which is a tangible cultural heritage object in the *Physical Space*. These resources are collected and aggregated into the CDI. According to the One-to-One Principle of Metadata, each of these digital resources has its own metadata

External resources such as the Wikipedia article and the BBC article have some useful information related to the *Statue of Tara* which can be linked to enrich the CDI. The descriptions about the *Statue of Tara* on the right side is another critical information which should be aggregated to the CDI. The CDI instance created through the CHDE curation process model consists of digital resources, descriptions of the original heritage object, and some external resources. All these resources might be individually described by their metadata which is used for aggregation.

In the Linked Open Data (LOD) environment, RDF (Resource Description Framework) plays a prominent role and converting the CHDE into such a formal model is very helpful to realize it in the LOD environment. Figure 7 shows an RDF-based representation of a few entities of the CHDE model applied to *Kandy Esala Perahera*. Here, a few CHI instances are shown using simple RDF triples.

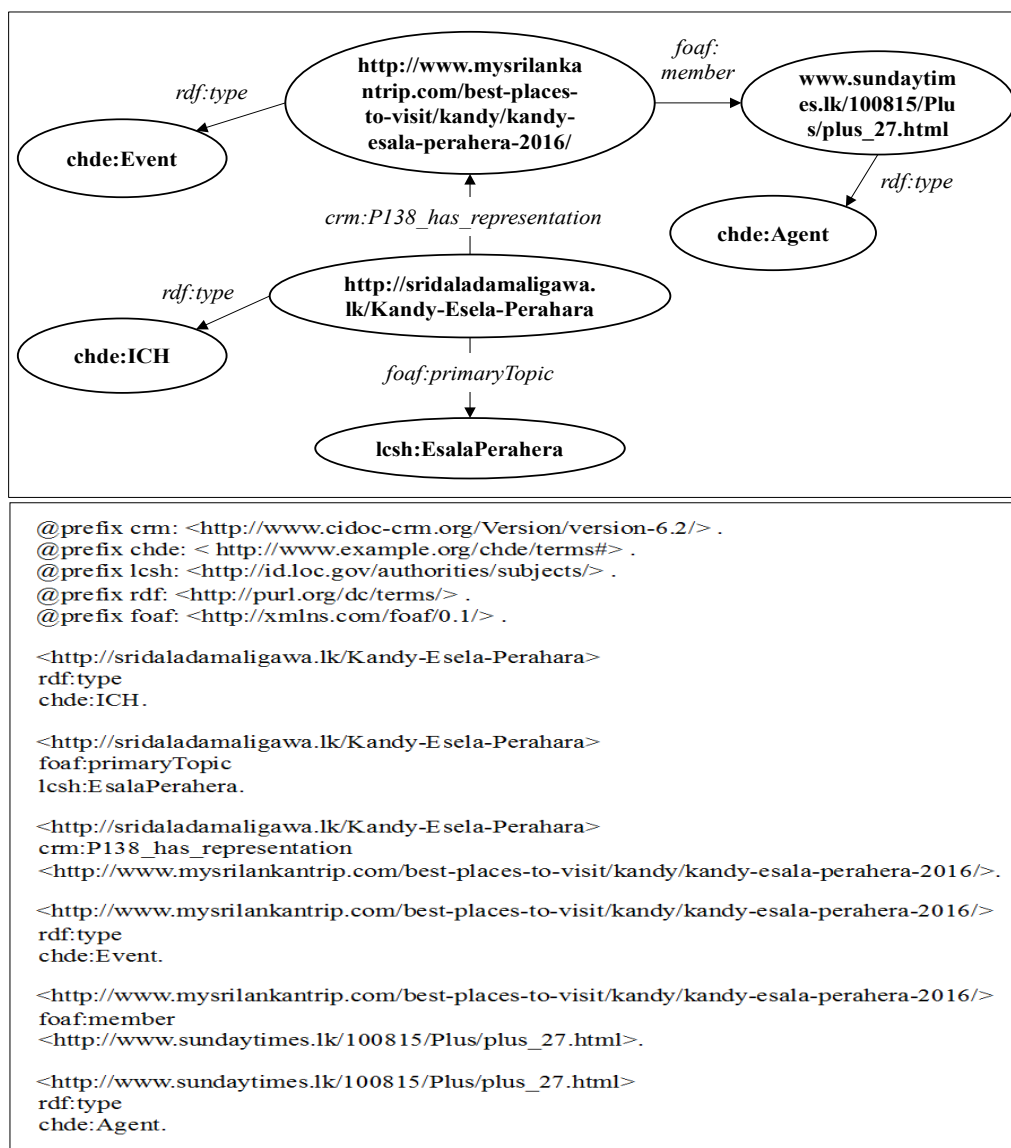


Figure 7. Fragment of the CHDE model in RDF

## CONCLUSION AND DISCUSSION

The paper has proposed an abstract level metadata model to organize and aggregate heritage related information in the networked information environment. This section discussed how the two research questions have been addressed, issues encountered and limitations of the study.

### Instantiation as Physical Resource for Building Digital Archives of Intangible Cultural Heritage

The first research question asked what are the intangible cultural heritage digital resources that digital archives represent? A digital archive of cultural heritage is primarily a collection of digital surrogates of CHOs. Those digital surrogates are mostly realized as visual and/or audio data. A fundamental requirement for creating a digital surrogate is that the original cultural heritage has to be presented in a form recordable into audio-visual media. However, intangible cultural heritage is primarily composed of knowledge and skills inherited from our



ancestors and can be presented as a performance by people or a group of people who owns the knowledge and skills. A digital surrogate of intangible cultural heritage in a digital archive is not a surrogate of the intangible cultural heritage but a surrogate of a single performance. This distinction seems trivial but is important to properly organize digital archives of intangible cultural heritage. For instance, a memory institution might recognize a record taken at *Kandy Esala Perahera* in 2016 as an intangible cultural heritage entity of Sri Lanka. There is no way to digitally record *Kandy Esala Perahera* as an intangible cultural heritage entity, as *Kandy Esala Perahera* is an event which takes place every year. In principle, intangible cultural heritage cannot be digitized but only its *Instantiations* can. Aggregation of digital surrogates of *Instantiations* may be a quasi-surrogate for intangible cultural heritage: it has the significant advantage of showing historical and contextual information which is a crucial aspect for digital archives of intangible cultural heritage.

*Instantiation* may be used to model artifacts which physically exist only in a particular time period and place, that is, dynamic artifacts such as fireworks, installation, theater play, etc. These artifacts may have one or more instantiations. Here, a clear distinction between the dynamic artifacts and their instantiations is useful to build a digital archive of dynamic artifacts.

Digital archives have been created to record events such as natural and man-made disasters and activities such as sports and game plays. Unlike the dynamic artifacts referred to above, *Instantiation* may not apply to events and activities even if we use recording media to archive the events and activities, because they are physical entities that existed at some point in a given time and location. CIDOC-CRM has *E5 Event* and *E7 Activity* which is a sub-class of *E5 Event*. As shown below (Table 2), *Instantiation* is defined as a subclass of *E7 Activity* because *Instantiation* can be defined as an “*Activity*” specialized to present a physical entity from an artifact of intangible cultural heritage.

### **Metadata Model for Aggregating Institutional and Non-Institutional Metadata**

The second research question was, how can we model metadata for digital object to be created by aggregation of fragments extracted from existing digital archives and other Web resources? This question is connected to the requirement of providing users with more contextual information by linking institutional and non-institutional CHI. User needs are diverse, and they might need multiple information resources related to a single tangible or intangible heritage object. The authors found that conventional institutional heritage metadata are more item-centric and non-institutional metadata are more comprehensive and context rich. Therefore, CHDE tries to connect these item-centric and comprehensive information through an entity called CDI (Figure 3 and 6). According to Figures 3 and 6, CDI aggregates individual items, their original metadata descriptions plus their related external resources such as Wikipedia. CHDEs’ CDI entity aggregates all these resources while providing users more enriched and linked form of CHI by linking institutional and non-institutional metadata descriptions. Therefore, CDI is a set of digital resources connected to a single tangible or intangible cultural heritage instance which enabling item-to-collection level data aggregation.

Since CHDE collects and aggregate diverse information from institutions and non-institutions, it may have various metadata schemas and mixed form of metadata. During this process, the authors used the One-to-One Principle of metadata as the key concept to build up the CHDE model and each layer of the CHDE adheres to the above principle.

**Table 2. Semantic relationship between the CHDE and existing schema classes**

CHDE Classes		Relationship with Existing Schema Classes	
Class Label	Class Name	Relationship	Schema Class
1	Curated Digital Instance	chde:CDI	rdfs:subClassOf ore:Aggregation
2	Collected Set of Digital Resources	chde:DigitalSet	rdfs:subClassOf dcmiType:Collection
			rdfs:seeAlso aat:Collection
3	Digital Resource	chde:Digital	rdfs:subClassOf crm:E73_Information_Object
			rdfs:seeAlso edm:WebResource
			rdfs:seeAlso aat:Recording
			rdfs:seeAlso aat:Digital
4	Collected Set of Offline Resources	chde:OfflineSet	rdfs:subClassOf dcmiType:Collection
			rdfs:seeAlso aat:Collection
5	Offline Resource	chde:Offline	rdfs:subClassOf FRBR <sub>oo</sub> :F26_Recording
			rdfs:subClassOf dcmi:PhysicalResource
			rdfs:seeAlso aat:Recording
6	Instantiation	chde:Instantiation	dcterms:hasPart edm:Event
			rdfs:subClassOf crm:E7_Activity
			rdfs:seeAlso aat:Event
			rdfs:seeAlso aat:TimeBasedWorks
			rdfs:seeAlso aat:PerformanceArt
7	Agent	chde:Agent	rdfs:subClassOf crm:E39_Actor
			rdfs:subClassOf dcmi:Agent
			rdfs:seeAlso aat:Agent
8	Intangible Cultural Heritage	chde:ICH	rdfs:subClassOf crm:E28_Conceptual_Object
			rdfs:seeAlso aat:IntangibleCulturalHeritage
9	Tangible CHO	chde:TCH	rdfs:subClassOf crm:E18_Physical_Thing
			rdfs:seeAlso aat:TangibleCulturalHeritage

Since one object can be associated with many digital surrogates, they used One-to-One Principle as a core rule to maintain the intentions of the digital surrogates (resources) and their original resource descriptions. Although metadata aggregation from different sources is a key issue in this research, this paper does not mention about the technology for the metadata aggregation. As in many cases, metadata of cultural heritage digital archive is hybrid, we need to identify objects described by source metadata which should be aggregated into a CDI. In other words, we have to identify an object corresponding to a description component in the source metadata. Figure 3 and 6 conceptually show this metadata aggregation scheme. The authors are considering that the DCMI Application Profiles model, in particular, its Description Set Profiles model would work well to build a model of metadata aggregation for the CDI.

## Semantic Relationship Between the CHDE and Existing Schema Classes

This study made a semantic mapping between CHDE classes and existing related schema classes (Table 2). As CHDE model is generalized and it uses RDF like instances it can be easily converted into the RDF format as well. Representing CHDE in RDF enables the presentation of the CHDE entities and their relationships in a more formal and exclusive way. This was discussed previously (Figure 7) and the authors tried to present a fragment of the CHDE through RDF. This kind of RDF realization is very important as we need to link the CHI resources and realize them in the LOD environment.

## Limitations

An *Instantiation* of an intangible cultural heritage entity is the key unit as an answer to the first research question, that is, we need to use physically existing entities but not abstract entities as an instance to be curated into a digital archive. CIDOC-CRM has class *Activity*, which is used as a super-class of *Instantiation* in the mapping table shown in Table 2. However, the authors realized that existing schema classes cannot fully represent the idea of *Instantiation* in the CHDE which is intended to define activities which can be recorded into physical media whose contents can be collected into a digital archive. As there are many *Instantiations* of a single intangible cultural heritage entity, a single intangible cultural heritage may be presented as a collection of *Instantiations*. According to the CHDE, *Instantiation* is a physical entity and it can be recorded in *Offline Resources*. Curators select these recorded *Instantiations* based on their institutional policies. Nevertheless, the first research question “How can we describe intangible cultural heritage for digital archives?” would include contextual and historical descriptions about an intangible cultural heritage entity, which means that we would need to collect not only digital surrogates of *Instantiations* but also those resources that explain the intangible cultural heritage such as Wikipedia articles and websites. There may be cases where the identification of a single intangible cultural heritage entity or a single *Instantiation* is not very clear because of the nature of intangible cultural heritage. However, such an identification issue should be solved by domain specialists and is out of the scope of this study.

Regarding the second research question of modeling the metadata for a digital object created by aggregation of fragments extracted from existing digital archives and other Web resources, this study answers mainly in the abstract entity level, but not in the levels of metadata schema and implementation. The CHDE model has the entities named *Collected Set of Digital Resources* and *Collected Set of Offline Resources*. These entities explicitly show that a CDI is an aggregation of existing archived resources in the entity level. The DCAP model applies to the schema and implementation levels since it defines a metadata instance as a set of descriptions (i.e. Description Set) where each description is a metadata component describing a resource. However, a detailed discussion of metadata aggregation in the metadata schema and implementation levels are left to a future paper.

We have limited the scope of the discussion to the abstract entity level so that this paper does not address requirements in the metadata schema and implementation levels. Aggregation on the metadata schema level requires matchings across metadata schemas. Metadata schema matching obviously requires *property level* matching, which can only be done manually unless metadata schemas are built on shared metadata vocabularies. However, the schema matching of the *property level* is complicated.

In conclusion, the paper discussed an approach to integrate diverse and dispersed CHI on the Web. The model is intended for South and Southeast Asian heritage yet, this can be

applied to any region. As the CHDE model used more broad and general classes, it can be considered as a generalized framework which can straightforwardly adaptable for both tangible and intangible cultural heritage entities which realized as digital objects in the networked information environment.

## ACKNOWLEDGMENT

This paper is based on a research partially supported by JSPS KAKENHI Grant (in aid of scientific research) number: (A) #16H01754. The authors wish to thank Professor Atsuyuki Morishima and all the members at the Metadata Lab for their support and guidance provided.

## REFERENCE

- Amin, R., Baker, O.F., Deraman, A., & Yatim, N.F.M. (2012). Transforming model to meta model for knowledge repository of Malay intangible culture heritage of Malaysia. *International Journal of Electrical and Computer Engineering*, 2(2), 231–238. doi:10.11591/ijece.v2i2.205
- Carboni, N., & de Luca, L. (2016). Towards a conceptual foundation for documenting tangible and intangible elements of a cultural object. *Digital Applications in Archaeology and Cultural Heritage*, 3(4), 108–116. doi:10.1016/j.daach.2016.11.001
- Chryssoula, B., Doerr, M., Le Bœuf, P., & Riva, P. (Eds.). (2015). *Definition of FRBRoo: A conceptual model for bibliographic information in object-oriented formalism* (Ver. 2.4). The Hague: IFLA International Working Group on FRBR and CIDOC CRM Harmonisation.
- Cononstantopoulos, P., Dallas, C., & Androutsopoulos, I. (2009). DCC & U: An extended digital curation lifecycle model. *International Journal of Digital Curation*, 4(1).
- Coyle, K., & Baker, T. (2009). *Guidelines for Dublin core application profiles*. Dublin Core Metadata Initiative. Retrieved from <http://dublincore.org/documents/profile-guidelines/>
- Gilliland, A. J. (2008). *Setting the stage*. In M. Baca (Ed.), *Introduction to metadata version 2.0* (online version). Los Angeles: Getty Publications. Retrieved from [http://www.getty.edu/research/publications/electronic\\_publications/intrometadata/setting.html](http://www.getty.edu/research/publications/electronic_publications/intrometadata/setting.html)
- Higgins, S. (2008). The DCC curation lifecycle model. *International Journal of Digital Curation*, 3(1).
- Hyvönen, E. (2012). Publishing and using cultural heritage linked data on the semantic Web. *Synthesis Lectures on the Semantic Web: Theory and Technology*, 2(1), 1-159
- Isaac, A. (Ed.). (2013). *Europeana data model primer*. Retrieved from [https://pro.europeana.eu/files/Europeana\\_Professional/Share\\_your\\_data/Technical\\_requirements/EDM\\_Documentation/EDM\\_Primer\\_130714.pdf](https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Primer_130714.pdf)
- Le Boeuf, P., Doerr, M., Ore, C.E., & Stead, S. (Eds.). (2015). *Definition of the CIDOC conceptual reference model* (version 6.2). ICOM/CIDOC Documentation Standards Group and CIDOC CRM Special Interest Group. Retrieved from [http://www.cidoc-crm.org/sites/default/files2/cidoc\\_crm\\_version\\_6.2.1.pdf](http://www.cidoc-crm.org/sites/default/files2/cidoc_crm_version_6.2.1.pdf)
- Lor, P. J., & Britz, J. J. (2012). An ethical perspective on political: Economic issues in the long-term preservation of digital heritage. *Journal of the American Society for Information Science and Technology*, 63(11), 2153-2164. doi: 10.1002/asi.22725
- Miller, S. J. (2010). The one-to-one principle: Challenges in current practice. In *Proceedings on International Conference on Dublin Core and Metadata Applications 20-22 October 2010, Pittsburgh, USA* (pp. 150-164). Dublin Core Metadata Initiative.

- Retrieved from <http://dcpapers.dublincore.org/pubs/article/view/1043/992>
- Orgel, T., Höffernig, M., Bailer, W., & Russegger, S. (2015). A metadata model and mapping approach for facilitating access to heterogeneous cultural heritage assets. *International Journal on Digital Libraries*, 15(2-4), 189-207.
- Peroni, S., Tomasi, F., & Vitali, F. (2012). Reflecting on the Europeana Data Model. In M. Agosti, F. Esposito, S. Ferilli, & N. Ferro (Eds), *Digital libraries and archives, IRCDL 2012* (pp. 228-240). Berlin: Springer.
- Smiraglia, R. P. (2005). Content metadata: An analysis of Etruscan artifacts in a museum of archeology. *Cataloging & Classification Quarterly*, 40(3-4), 135-151.
- Sugimoto, S. (2014). Digital archives and metadata as critical infrastructure to keep community memory safe for the future lessons from Japanese activities. *Archives and Manuscripts*, 42(1), 61-72. doi:10.1080/01576895.2014.893833
- Wickett, K. M., Isaac, A., Doerr, M., Fenlon, K., Meghini, C., & Palmer, C. L. (2014). Representing cultural collections in digital aggregation and exchange environments. *D-Lib Magazine*, 20 (5/6).
- Wijesundara, C., Monika W., & Sugimoto, S. (2017). A metadata model to organize cultural heritage resources in heterogeneous information environments. In *Proceedings of 19th International Conference on Asia-Pacific Digital Libraries, ICADL 2017, Bangkok, Thailand, 13-15 Nov 2017* (Lecture Notes in Computer Science 10647, pp. 81-94). Heidelberg: Springer. doi:10.1007/978-3-319-70232-2\_7
- Wijesundara, C., Sugimoto, S., Narayan, B., & Tuamsuk, K. (2016). Bringing cultural heritage information from developing regions to the global information space as linked open data: an exploratory metadata aggregation model for Sri Lankan heritage and its extension. In *Proceedings of 7th Asia-Pacific Conference on Library and Information Education and Practice (A-LIEP 2016), 3-4 Nov 2016, Nanjing, China* (pp. 117-132). Nanjing: School of Information Management, Nanjing University.
- Woodley, M. S. (2016). Metadata matters: Connecting people and information. In M. Baca (Ed.), *Introduction to metadata* (3rd ed.). Los Angeles: Getty Publications. Retrieved from <https://www.getty.edu/publications/intrometadata/metadata-matters/>
- Zeng, M. L., & Qin, J. (2016). *Metadata* (2nd ed.). Chicago: ALA Neal-Schuman Publishers.