The Traditional Malay Textile (TMT) Knowledge Model: Transformation towards Automated Mapping

Syerina Azlin Md Nasir, Nor Laila Md Noor, Suriyati Razali

Abstract—The growing interest on national heritage preservation has led to intensive efforts on digital documentation of cultural heritage knowledge. Encapsulated within this effort is the focus on ontology development that will help facilitate the organization and retrieval of the knowledge. Ontologies surrounding cultural heritage domain are related to archives, museum and library information such as archaeology, artifacts, paintings, etc. The growth in number and size of ontologies indicates the well acceptance of its semantic enrichment in many emerging applications. Nowadays, there are many heritage information systems available for access. Among others is community-based e-museum designed to support the digital cultural heritage preservation. This work extends previous effort of developing the Traditional Malay Textile (TMT) Knowledge Model where the model is designed with the intention of auxiliary effort of developing the Traditional Malay Textile (TMT) Knowledge Model and CIDOC CRM as exemplars in refining the facets in the model particularly involving TMT-Artifact class. The result is an extensible model which could lead to a common view for automated mapping with CIDOC CRM. Hence, it promotes integration and exchange of textile information especially batik-related between communities in e-museum applications.

Keywords—automated mapping, cultural heritage, knowledge model, textile practice

I. INTRODUCTION

Preserving national heritage has become a key agenda in many countries in the world. In Malaysia, one of its preserved values is Traditional Malay Textiles (TMT) such as Songket, Batik, Pua Kumbu, Cindai and etc. [18]. At present, most of TMT collections have been exhibited in Malaysian museums using traditional curation methods with constraints of limited descriptions and articulations to portray the distinctiveness of TMT artifacts. There are also those individuals who posses and own priceless TMT artifacts and knowledge and those who produce these textiles and preserve the techniques of making them. Apart from that, attempts have been made by several institutions to initiate the digitizing of documents and artifacts. However, most of the works are isolate, unknown and for private usage. Therefore, work has been done to collect all these Malay textiles to strengthen the conservation for national cultural heritage preservation in museums and documented as TMT Knowledge Model [2]. To fully grasp the advances of web technology, the construction of this model will be extended by mapping the model with the standard ontology in cultural heritage domain, CIDOC CRM.

The growth in number and size of ontologies indicates the well acceptance of its semantic enrichment in many emerging applications. Each ontology being designed has been represented in different ways with different purposes in mind. The current scenario requires these available ontologies to be mapped together [9], [11] which raises an issue to find similar entities and semantic correspondences that exist between them. In this case, CIDOC CRM provides both a global and extensible model into which data originating from distinct sources can be mapped and integrated, and base concepts that future metadata initiatives could build on when developing domain specific vocabularies [29]. Mapping TMT Knowledge Model and CIDOC CRM raises several challenges that need further investigation. The first and primary challenge is the diversity of the metadata terms used to describe and structure the collections [3]. The second challenge is the rigid structure of TMT Knowledge Model which makes it unable to deliver very rich meaningful artifacts [20]. Finally, the result of the mapping will hopefully allow room for ontology growth in e-museum application and create a wealth of opportunities for interactions with collections of experts’ knowledge from the community.

Numerous studies and a number of solutions have been proposed to solve the ontology mapping problem. There are reports on the success and difficulties of ontology building using either manual approaches or with tool supports. Each claims one is better than the other in terms of its performance and scalability. [24] states that a lot of research has been monopolized by heuristics and machine learning approach (automation). Various mapping techniques have been
explored and categorized [15] to further understand the mapping process. At the same time, a lot of work has been reported on the development of heuristics and machine learning tools (e.g. AnchorPrompt [26]; COMA++ [10]; GLUE [4]; S-Match [16]). Most of these tools, methods or techniques are available. However, in reality, they are not freely available or still in development stage or in the form of algorithms that need clarifications before it can be further exploited. Therefore, this brings us back to the basic of mapping by reviewing the manual approach. Some of the works claim that manual mapping is laborious, time consuming, error prone, difficult to maintain and update [10], [19] and yet, [7] suggest the creation of benchmark ontologies by manually defining mapping between concepts. [25] elaborates on the concept of ‘shared ontology’ which is agreed upon by developers from various applications for future adoption and extension to their specific application. Thus, this study aims to transform the TMT Knowledge Model accordingly in which would enable the representation of TMT by modeling in an adequately consistent way the conceptualization of the reality behind textile practice.

The remainder of this paper is organized as follows. Section II summarizes the previous work concerning standard ontology, CIDOC CRM. Section III then describes the approach used in this work. The actual results are presented in Section IV. The paper concludes with Section V that discusses on the anticipated future work directions.

II. RELATED WORK

Of late, the number of ontologies created for various domains has been increasing. Each specific domain expects a well-developed ontology, where most of the vocabularies that are needed for meaningful interaction and communication are available between people and application systems, hence, is able to handle vast amounts of distributed and heterogeneous computer-based information. There are several domain ontologies which gain popularity among researchers like in the area of medicine [28], [6], food industry, tourism [22], [8], bioinformatics [2], [1] and cultural heritage [13]. Studies on creating a common ontology also rise in different domains. Among the earliest ontologies being built are Suggested Upper Merged Ontology (SUMO) and Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE). Over the years, several other reference ontologies have been constructed for specific domains of discourse.

The existence of upper level ontology in specific domains helps many ontology developers in creating ontology which conform to the standard. For example, in cultural heritage domain, ontology for cultural heritage information like CIDOC CRM is referred. Among works that adopt CIDOC CRM are the implementation of cultural heritage web information system on the Semantic Web such as SCULPTUER, AMA, CRM-EH and Digital Museums. SCULPTUER project handles museum multimedia collections by mapping the museum’s partner legacy system with CIDOC-CRM for cross collection searching [30]. The Archive Mapper for Archaeology (AMA) project aims to create tools for semi-

<table>
<thead>
<tr>
<th>TABLE I SUMMARY OF CIDOC CRM HARMONIZATION PROJECTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Domain</strong></td>
</tr>
<tr>
<td><strong>area</strong></td>
</tr>
<tr>
<td><strong>Purpose</strong></td>
</tr>
<tr>
<td><strong>Source of Data</strong></td>
</tr>
<tr>
<td><strong>Methods for ontology mapping</strong></td>
</tr>
<tr>
<td><strong>Tool or system associated</strong></td>
</tr>
</tbody>
</table>
developers to include semantics and structures as a way of transforming this unstructured information into a format that machines could understand. The harmonization projects are summarized as shown in Table 1. All these projects aim to create a single ontology that represents the conceptualization of reality in the domain area. The works are done manually which involve experts and individuals from related institutions before the common ontology is agreed upon. In relation to this study, the investigation on CIDOC CRM is made to provide an insight on the importance of an upper ontology in mapping between two ontologies, especially when related to TMT.

III. METHODOLOGY

A. Data Description

This study will employ Traditional Malay Textile (TMT) Knowledge Model which is designed by Suriyati et al. [32]. As defined by Siti Zainon, I [31], TMT means the handicrafts which relates to the process of weaving invented by human beings for daily clothing in Malay Archipelago with the understanding of historical factors, symbolic and aesthetic values. The research is solely focused on the historical factors and limited to the description of textile from the Malay Peninsula. Due to the current nature of the Malay Textile Knowledge and Artifacts that are scattered among domain experts and novices, this model was designed. The taxonomy aims to represent Malay Textile content in organized structures for retrieval by user preference and to allow association between system development and user interaction platform (Malay Textile Community and Museum Community that have different backgrounds).

B. Motivating Examples

Example 1. ABC ontology is a model for the exchange and integration of digital library information. It was designed to model physical, digital and analogue objects of all media types; abstract concepts such as intellectual content and temporal entities; and describe other entities that occur across many domains. It encompasses 14 classes and 25 properties that present the conceptual basis for generating domain, role or community specific vocabularies across different domains as guidance to interested communities. The primary concerns are on volatile digital objects and intellectual property rights [21]. The mapping was done in 2001 by using the OntoClean approach previously used in integrating WordNet with OCT ontology [27].

In this example, the study described a detailed specification of all ABC classes and properties and represented it in graphical form. This provides an example on how entities are modeled and relationships are built between them. As such, it guides on how to build descriptive ontology [21]. It also provides examples of the model implementation which allow for construction of metadata repository of RDF description and sophisticated queries through the search interface. It pays more attention on the conceptualization of object transformation over time in which this case, on TEMPORALITY category.

Example 2. The Moving Pictures Expert Group (MPEG) released the “Multimedia Content Description Interface” standard for MPEG-7 for describing multimedia content. The combining effort between MPEG-7 and CIDOC CRM metadata models resulted in the creation of a standardized model for describing and managing museum multimedia content. The work merged both ontologies and further extended MPEG-7 with additional sub-classes and sub-properties to cater for multimedia concepts and descriptions. The vast diversity and variety of multimedia resources especially in the museum domain makes it a necessity to record these elements in the most eloquent manner for preservation and dissemination purposes [28]. Through this example, it is learned that certain categories covered by CIDOC CRM requires for extension. This gives an alternative on how to represent the detailed description on the making process of textiles which is not represented in the model. The study shows how extension is carried out to further describe the digital multimedia contents by extending the INFORMATION OBJECT category.

C. Purpose

The TMT Knowledge Model was originally designed to preserve the Malay textiles through semantic relationship between artifact and its precedent knowledge by mapping the model with CIDOC CRM. This study aims to create Batik Heritage Ontology as the result of the mapping between both models to embody the whole process of batik making. However, the current state of the TMT Knowledge Model hinders the mapping process and requires for transformation to take place beforehand. Hence, this study is motivated to reach a common view for automated mapping with CIDOC CRM and to capture and represent the underlying semantic of Malay textile information for ease of integration and exchange between communities in e-museum applications.

D. Transformation Process

Initially, TMT Knowledge Model is developed using Protégé software. It consists of six concepts and thirteen subconcepts. Based on the examples of the harmonization works, this study will adapt both works. Consequently, TMT-Knowledge Model (see Fig. 2) will undergo a refinement process to capture its concepts and properties. In this work, all facets in the model are re-created by transforming concepts and subconcepts into RDFS classes and properties.
Due to the current state of CIDOC CRM classes which is limited in describing TMT-Artifact category, therefore, this category will be further improved to capture the details of the textile making. As a result, the model is extended by classifying the artifact into three categories which are Aesthetics Value, Materialization and Production. The Aesthetics Value category defines any value that causes an object to be ‘a work of art’. Materialization category is further divided into sub-categories describing Material (e.g. brush, block, chanting, cloth, etc), Motif (e.g. traditionally-based on flora & fauna, dongson etc) and Color (e.g. chemical dye, natural dye, combination, etc). Production category depicts the actual textile practice concerning Design, Technique and Making-Process. Design sub-category involves normal practices and skills of producing the artifacts. Lastly, Making-Process details the actual making which depends on the design and technique used.

IV. RESULTS

This section presents the new version of TMT Knowledge Model where the primitive category is an entity. Fig. 4 shows the facets of the model which encompasses five main classes: idea, existence, activity, time, location and four subclasses: handwork, community, artifact and subject which are described below. Classes are shown in rectangles and sub-class relationships are indicated by solid lines. The modified TMT-Artifact categories as shown in Fig. 3 are also included in the model as an extension. All the classes in the model were given a name and an identifier consists of the word TMT constructed according to the conventions used in the CIDOC CRM model.

A. Idea Class

Subclass of: Entity
Superclass of: Handwork
Description:
The Idea category is created to enable the expression of concepts or ideas. It is used to express the notion of Handwork which exists through sensible way such as when it is told, demonstrated and shown in some way as a means to bind several Materializations. Idea is inspired by an instance of Aesthetics Value in which included in Production, which may be an instance of Design, Technique and Making-Process.

B. Existence Class

Subclass of: Entity
Superclass of: Community, Artifact, Subject
Description:
The Existence is similar to entity Persistent Item (E77) which stands in contrast to the Idea category. It expresses tangible cultural heritage which refers to something that remains intact to see, hold and is movable and normally preserved in its original form. TMT expresses this notion through the inPlace that associates Existence with Location.

C. Activity Class

Subclass of: Entity
Description:
The Activity is equivalent to entity Activity (E7) which involves Community in the context of Production. TMT expresses this notion through hasInvolve that associates Activity with Community.

D. Time Class

Subclass of: Entity
Description:
The Time is the same as entity Time-Span (E52) which represents a time span or point in time. The influence property binds the Time and the Activity in which activities are being influenced by some period of time.
E. Location Class

**Subclass of:** Entity

**Description:**

The *Location* is similar to entity Place (E53) which represents the origin or place where the artifacts can be found. The *hasHistory* property associates *Location* with *Time* to describe the actual existence of artifacts.

F. Handwork Subclass

**Subclass of:** Idea

**Description:**

The *Handwork* is a conceptual notion which only disclosed when it has been actualized through *hasRealization* property in some *Materialization*.

G. Community Subclass

**Subclass of:** Existence

**Description:**

The *Community* participates during an *Activity* which may be experts, semi-experts, novices, group from cultural institutions, etc.

H. Artifact Subclass

**Subclass of:** Existence

**Superclass of:** Aesthetics Value, Materialization, Production

**Description:**

The *Artifact* expresses logical entities that are tangible realization of concepts and can be manifested in many ways such as through Aesthetics Value, Materialization and Production.

I. Subject Subclass

**Subclass of:** Existence

**Description:**

The *Subject* refers to actual things that exist and symbolized by the *Motif* in the *Design* through the *represents* property.

V. FUTURE WORK AND CONCLUSIONS

With the transformation of TMT Knowledge Model, the next step is to test, evaluate and refine it by using the real collections of cultural artifacts within museums. Digitizing the heterogeneity of cultural heritage collections is a major challenge for ontology developers in order to support the demand of online applications. For example, the nature and
type of collections may consist of text, written on different materials, paintings, photographs, 3D objects, sound recording or even maps [7]. Similarly, the scope of artifacts covered in TMT may vary widely and the chance to discuss or mention the same objects is relatively small and sometimes limited. Therefore, mapping TMT with the upper ontology, CIDOC-CRM could resolve the problem of disagreements and misinterpretation of these artifacts. Hence, a future goal is to construct Batik Heritage Ontology through automated mapping between TMT Knowledge Model and CIDOC CRM in order to preserve batik-related information.

To conclude, this research intends to provide a platform for the museum community to enrich knowledge on cultural heritage artifacts and deploy an e-museum design pattern through community based e-museum. For that reason it could be implemented in any museum in Malaysia. Once the knowledge structure is constructed, the community could promptly access cultural heritage knowledge as well as sustain ongoing discussions on particular topics of interest. In addition, community based e-museum may also create opportunities to enhance user interest, new exciting experiences and to experiment with new strategies for user engagement that will help to build the community and attract new web audiences.

ACKNOWLEDGEMENT

Special thanks to Martin Doer for his suggestions and recommendations in this study. He is a senior researcher in the Institute of Computer Science, Foundation for Research and Technology – Hellas (ICS/FORTH), Greece. He has been involved in most of the harmonization projects involving CIDOC CRM.

REFERENCES