Query Answering from Traditional Dance Videos: Case Study of Zapin Dances

Sihem Belabbes[1]   Chi Wee Tan[2]   Tri-Thuc Vo[3]   Yacine Izza[1]   Karim Tabia[1]   Sylvain Lagrue[4]   Salem Benferhat[1]

[1]*CRIL & CNRS, Université d'Artois*, Lens, France
{belabbes,izza,tabia,benferhat}@cril.fr

[2]*School of Computing, Universiti Teknologi Malaysia*, Johor, Malaysia
seantan8989@gmail.com

[3]*CICT, Can Tho University*, Can Tho, Vietnam
vtthuc@ctu.edu.vn

[4]*Heudiasyc & CNRS, Université de Technologie de Compiègne*, Compiègne, France
sylvain.lagrue@hds.utc.fr

Abstract—The aim of this paper is to highlight two important issues related to the annotation and querying of Intangible Cultural Heritage video datasets. First, we focus on ontology completion by annotating dance videos. In order to build video training sets and to enrich the proposed ontology, manual video annotation is performed based on background knowledge formalized in an ontology, representing a semantics of a traditional dance. The paper provides a case study on Malaysian Zapin dances. Second, we address the question of how can end-users efficiently query the datasets of annotated videos that are built.

Index Terms—Ontologies, Video Annotation, Querying

I. Introduction

Southeast Asia is a rich and ethnically diverse region. The United Nations Educational, Scientific and Cultural Organization (UNESCO) lists many of the living traditional art forms in this region as intangible cultural heritage (ICH) needing preservation [1]. Examples include traditional dances as well as local operas and theatres.

This work falls within the framework of the "AniAge" research project [2] which aims at tackling challenging problems such as archiving and reproducing style-preserved ICH contents. The overall aim of AniAge is to develop novel techniques and tools to reduce the production costs as well as to improve the level of automation without compromising the control of the artists.

A formalism that has been widely used in various studies on ICH preservation is that of an ontology [3], [4], [11], [13]. In this work, we propose an OWL ontology (a.k.a. terminological knowledge base or TBox) to digitally conceptualize the cultural aspects of Southeast traditional dances, illustrated here by Malaysian Zapin dances. An ontology is one of the two components of a knowledge base (KB). The other component is a base of ground facts (a.k.a. assertional base or ABox).

The first contribution of this paper consists in an ontology-based modelling of Zapin traditional dances. We specify the knowledge base in DL-Lite [2], a family of lightweight fragments of Description Logics (DLs) which are particularly interesting since they provide a good trade-off between expressive power and computational complexity. Indeed, a big advantage of DL-Lite is that query answering can be performed efficiently using query rewriting [6].

The second contribution of this paper consists in using knowledge representation approaches to "enrich" raw data (mainly videos of traditional dances) to allow for better exploitation and dissemination of these data. Dance videos can be enriched by adding annotations describing their content (e.g. description and semantics of some expressions, postures, movements, costumes and props related to a dance), the structure (e.g. the different parts of a dance) and any other relevant information.

To this aim, a completion of the ontology is done by domain experts by annotating dance videos using an annotation tool that we have developed. The annotations of any given video can then be translated into an ABox which can in turn be queried.

We present the contributions of this paper as follows. In Section II, we describe our ontology for Malaysian Zapin dances. In Section III, we present our tool for annotating videos. In Section IV, we discuss query answering based on annotated videos. We then conclude the paper in Section V.

II. An Ontology for Zapin Dances

A. Overview of Malay Traditional Dances

The Malay traditional dance in Malaysia, called Zapin, is a mixture compound of Malay ethnic lifestyle resulting from the settlement of Zapin Arabs [12]. Johor Malay Zapin has developed into Malay art in the early days of settlements in coastal areas and the main river banks in the State of Johor. As modern types of amusement increasingly take place in social life [10], traditional elements can no longer sustain their position in society. To the best of our knowledge, there are no studies dealing with an ontology-based modelling of Malay Zapin dances. The exposure of the original form of Zapin Johor in a structural way can only be seen in the film

---

[1]https://ich.unesco.org/en/lists
[2]http://www.cril.univ-artois.fr/aniage/

"Zaleha Ayam Patah" by Shahrul Azli (Uniputra Production, Universiti Putra Malaysia, 2002) and a documentary video named "Let's Learn Zapin Johor" by Yayasan Warisan Johor. Some studies [5], [7] have focused on Zapin dances which have become one of the most popular folk dances spread over Southeast Asia. In the sequel, we build an ontology to conceptualize Zapin dances.

B. The Proposed Zapin Ontology

We have built the Zapin Ontology, depicted in Figure 1, by selecting the most significant features of Zapin dances [5], [7]. Indeed, this type of dances consists of four main parts, namely: Dancer, Movement, Music and Plot.

For Dancer, costume is divided into three main groups: Accessory, Body Wear and Head Part. The costumes worn by male dancers are a series of traditional Malay outfits. Meanwhile, women dancers use some Malay special clothes with Accessory such as Golden Necklace, and Hair Bun with a Decorative Flower.

Another important feature is the Plot of a particular dance. Each dance must start with salutation then proceed with the main dance and end with some movement related to the Movement class. For lack of space, we cannot explicitly provide all concepts and roles of our ontology. The next section provides some details on two main features of Zapin dances, namely "Movement" and "Plot".

C. Main Concepts of Zapin Traditional Dances

The name Zapin comes from the Arabic word "Zaffan" which means dancers and "Al-Zapin" which means leg movement. This explains why there are many types of movements regarding leg and foot in Zapin dances. Movements are composed of those of an individual dancer and those of interactions between dancers. They are discussed in the next two subsections.

1) Movements of an Individual Dancer: Generally, Zapin dances consist of a collection of movements interspersed with Malay daily life elements and culture. Nature-inspired dance motifs are linked and performed together in Zapin dance sequences throughout the dance. Ten types of movements of an individual dancer included in a Zapin choreography are listed in Figure 2. These are:

- *Taksim (Prelude)*: Performed in the beginning phrase of the dance as a prelude and salutation to the audience.
- *Asas (Basic)*: The fundamental of Zapin dances which consists of two parts [9]. In the first part: all movements on hold, left foot step forward, right foot step forward and left foot step. For the second part: hold position of left leg and right foot swings behind left foot as body turns 180 degrees, right foot steps while left foot turns 180 degrees to the right, left foot steps back while raising the body and right foot in a leg gesture and ending with right foot returns to the floor, left foot remains in place.
- *Titi Batang (Footlog)*: A type of movement depicting the uncertainties of body movement when a person is stepping on a slippery wooden footlog pathway.

- *Ayam Patah (Chicken Fracture)*: Dancer performs a movement imitating a chicken with a broken leg. According to [8], chicken movements symbolize events in nature that teach the mind about difficulties faced in life.
- *Kopak (Broken)*: A series of skipping and squatting position, used to connect one part of the Zapin movement to another movement.
- *Lompat Injut*: Another type of skipping movement mixed with Basic movement.
- *Pusar Pelanak Besar (Big Twirling)*: Inspired from the movement of twirling and skipping of a mudskipper (amphibious fish) in mud at low tide, to represent dexterity and agility.
- *Pusar Pelanak Kecil (Small Twirling)*: Same with Pusar Pelanak Besar but with smaller body movement.
- *Tapak (Foot)*: A type of movement moving forward and backward which emphasises on dancer's foot.
- *Wainab (Embraced)*: The word "wainab" is derived from Arabic *wainaq* which means embrace or hug. This ends the dance to signify the climatic yearning of wanting to be embraced by God. It also implies the meaning of appreciation for the audience and the end of the dance.

2) Movements of Interactions: Symbolically, movements in Zapin dances showcase the shared interests of dancers rather than an individual dancer's effects, in addition to dispel between dancers in groups. In this perspective, the interaction between male and female dancers is crucial. Figure 3 shows Interaction, one of the sub-classes of Movement.

There are three types of interactions involved, namely: Challenging, Enjoying (the music) and Playing, which showcase the uniqueness and distinctiveness of Zapin dances.
We now discuss the Plot concept.

3) Plot of Zapin Dances: A plot is a causal sequence of the dance that leads the audience into the performance. Essentially, it depicts the events that take place in a dance. Three elements are involved in a plot:

- *Opening*: Zapin makes use of Pantun (Malay poem). There is an improvisatory opening segment known as Taksim which is used to signify the dancer's performance and to divide one part of a song from another. Taksim is generally instrumental and is played either by the gambus instrument or the violin.
- *Main*: The main segment of the dance consists of Ayam Patah, Kopak, Lompat Injut, Pusar Pelanak Besar, Pusar Pelanak Kecil, Tapak, Titi Batang and Basic.
- *Closing*: As there is an opening, there is also a closing segment for Zapin dances, Wainab (already described as a subclass of Individual under the Movement concept).

D. Properties of the ontology

In this section, we interpret the concepts described above into an OWL document representing the ontology. Various ontology languages are available, such as N-Triples, RDF-XML and OWL abstract syntax. We choose Ontology Web Language (OWL) to develop our ontology. OWL is a language for building ontologies that provides high-level descriptions of Web content. It is built on a description logic foundation (as we shall see in Section IV, DL-Lite languages will be used) and
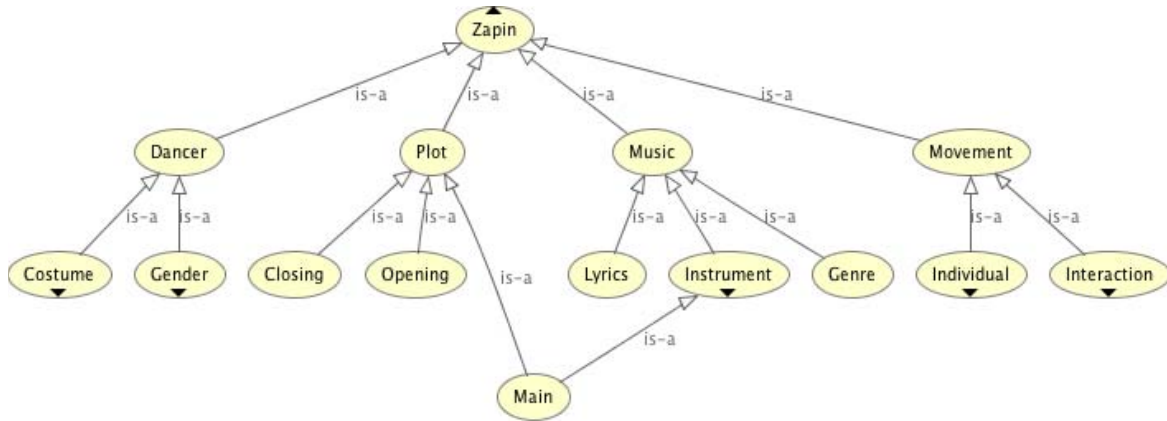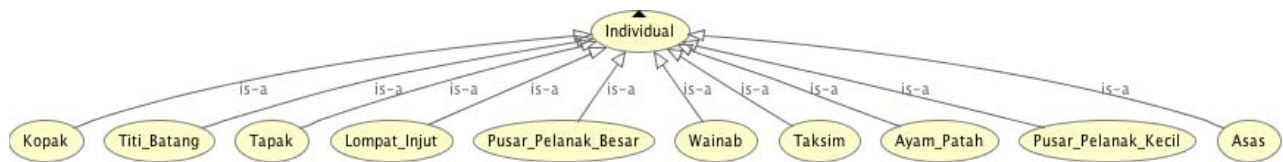
Fig. 1: Overview of the proposed Zapin Ontology

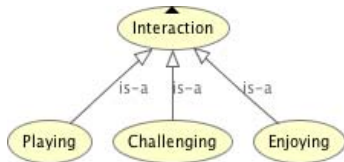Fig. 2: Individual: sub-class of Movement

Fig. 3: Interaction: sub-class of Movement

provides strong formal consistency guarantees when checking ontologies for inconsistencies and when deducing new pieces of knowledge.

We used the free and open source tool Protégé [3] to build the Zapin Ontology. The resulting Zapin dance ontology is an OWL document that comprises: 190 axioms, 55 classes, 7 object properties, 2 data properties and 12 individuals.

The classes express the concepts described in sections II-B and II-C and the properties express the relationships between concepts. All classes are stored in a hierarchical structure. Then the major class is *Zapin* which has the sub-classes *Dancer*, *Movement*, *Music* and *Plot*. The OWL document of the Zapin Ontology is available at http://www.cril.univ-artois.fr/aniage/ICTAI19 together with an overview of all entities of the ontology.

We have evaluated the completeness and consistency of the ontology. Consistency here is understood in the sense that the TBox admits a model and there is no empty concepts or roles. Namely, if A is a concept name, then $A \sqsubseteq \bot$ cannot be

derived from the ontology. We have used HermiT reasoner [4] to first check the consistency and then query the ontology using SPARQL.

III. FROM VIDEOS TO ABOXES

The Zapin Ontology conceptualizes the Southeast Asian traditional dances called Zapin. In order to build a dataset that enriches the ontology, we consider examples of Zapin dance videos. However, we have very little to no information at all on the semantic content of those videos. On the other hand, there are domain experts who are knowledgeable about Zapin dances and who can assist us in annotating Zapin dance videos. Hence, there is a need to develop a user-friendly tool that experts can use to annotate dance videos on the basis of the taxonomy defined in the ontology (i.e., the TBox). The idea is then to translate those annotations into a set of assertional facts called ABox, which can in turn be used in answering queries on the videos.

To this aim, we have developed a tool for producing ABoxes from a TBox. This tool is an annotation tool that enables the use of concepts and relations of the ontology in a consistent way by different users. This contributes to minimizing discrepancies in annotations. Moreover, users access the set of concepts and knowledge domains related to a specific dance. Therefore, they can easily select annotated content for each video frame or segment.

The diagram of the annotation tool is presented in Figure 4. The inputs of the annotation tool are on the one side dance videos and on the other side the OWL document of the
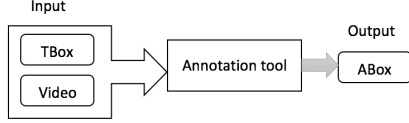
Fig. 4: Annotation tool diagram

ontology corresponding to dances. The output is the annotation data corresponding to an input video, and it is rendered as an ABox in OWL format. The tool enables users (experts) to upload and display videos and then to annotate them. Figure 5 shows a screenshot of one Zapin Dance video selected as an example to be annotated using the tool. In order to annotate a video, concepts and roles of the ontology are uploaded to the tool and they are displayed into a drop-down menu (right part of Figure 5) with options that can be selected by the users during the annotation process. Users mark video segments where each segment corresponds to a set of video frames. After marking video segments, users select concepts and roles from the menu options and choose the value corresponding to a concept or role for each segment. For instance, a value v associated with a concept A will represent an assertion $A(v)$ of the ABox (namely, the fact $A(v)$ is present in the video). The annotated data serves as an ABox associated with the TBox (terminological base). Once all frames are annotated, the Web-VTT document can be downloaded, and the annotations can be translated into assertions of the ABox. The translation process consists in parsing each frame of the Web-VTT, called *cue*, and generate the assertions of the current frame. The generated assertions are written in DL-Lite language and saved in an OWL document representing the ABox. The Web-VTT document and the OWL ABox associated with the video shown as an example in Figure 5 are available at http://www.cril.univ-artois.fr/aniage/ICTAI19.
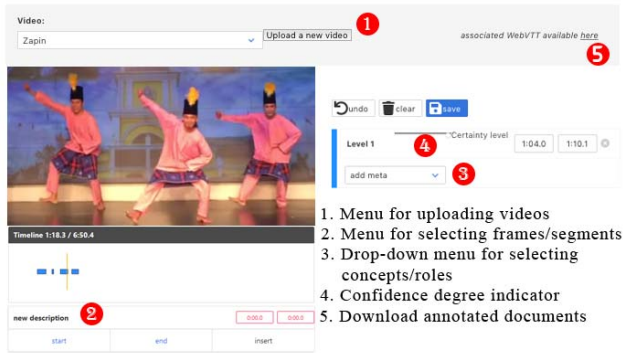

Fig. 5: Screenshot of Zapin Dance video annotation

IV. QUERY ANSWERING ANNOTATED DATA

Ontology-Mediated Query Answering (OMQA) is an active research area that takes advantage of the semantic knowledge specified in an ontology to answer queries about data. Typically, querying is performed on a knowledge base (KB)

which is composed of an ontology (a TBox) and a dataset of assertional facts (an ABox). We use DL-Lite, a family of lightweight languages, to encode our ontology (for lack of space, we do not provide the full transcription). We provide a brief reminder on DL-Lite, then discuss query answering.

A. Brief Reminder on DL-Lite

We recall the basics of DL-Lite [2] (specifically, the DL-Lite$_R$ variant) used for formalising lightweight ontologies.

A *knowledge base* (KB) is built upon a finite set of *concept names* C, a finite set of *role names* R and a finite set of *individual names* I, such that C, R and I are mutually disjoint.

The DL-Lite$_R$ language is defined according to the following rules: $R \rightarrow P \mid P^-$; $E \rightarrow R \mid \neg R$; $B \rightarrow A \mid \exists R$; $C \rightarrow B \mid \neg B$; where $A \in$ C, $P \in$ R, and P^- is the *inverse* of P. With R (resp. E) we denote a *basic* (resp. *complex*) *role*, while B (resp. C) denotes a *basic* (resp. *complex*) *concept*.

An *inclusion axiom* on concepts (resp. on roles) is a statement of the form $B \sqsubseteq C$ (resp. $R \sqsubseteq E$). Inclusions with $\neg$ in the right-hand side are called *negative inclusions*, otherwise they are called *positive inclusions*. A *TBox* $\mathcal{T}$ is a finite set of both positive and negative inclusion axioms. An *assertion* is a statement of the form $A(a)$ or $P(a, b)$, with $a, b \in$ I. An *ABox* $\mathcal{A}$ is a finite set of assertions. For a given $\mathcal{T}$ and $\mathcal{A}$, we denote a KB with $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$. A KB is *consistent* if it admits at least one model, it is *inconsistent* otherwise. A TBox $\mathcal{T}$ is *incoherent* if there is $A \in$ C such that A is empty in every model of $\mathcal{T}$, it is *coherent* otherwise. Henceforth, we refer to DL-Lite$_R$ simply as DL-Lite.

B. Querying ABoxes

In this section, we propose an OMQA approach for querying KBs representing Zapin dances. We take the reasonable assumption stating that the TBox is stable, coherent and reliable. For the purpose of illustration, we use the following example.

Example 1. *Assume a video v of a Zapin dance that has been annotated by two experts E_1 and E_2. Let the considered video be the one of Figure 5. For the sake of simplicity, let us focus only on the annotations associated to one frame, denoted f, in video v. Let us also assume that the annotations given by experts E_1 and E_2 for f are listed in tables I and II.*

TABLE I: Annotations of expert E_1 on frame f

Expert E_1
Annotation (assertion)
Male(f)
Taksim(f)
Enjoying(f)

TABLE II: Annotations of expert E_2 on frame f

Expert E_2
Annotation (assertion)
Male(f)
Tapak(f)

The annotations of expert E_1 (Table I) state that frame f depicts a man (f is an instance of the Zapin Ontology's

concept Male*) performing* Taksim *(f is an instance of the concept* Taksim*) and the interaction is of enjoying type (f is an instance of the concept* Enjoying*). Similarly, Table II contains the annotations of expert* E_2 *for frame f. Recall that annotations correspond to assertions, i.e., instantiated concepts and roles of the ontology.*

We denote the TBox by $\mathcal{T}$ *and the ABox by* $\mathcal{A}$*. Space considerations prevent us from providing the full set of axioms of the TBox, but we mention some axioms when needed.* □

We assume that all experts agree in their annotations of any given video. This means that the ABox is consistent w.r.t. the TBox. In this case, query answering (QA) simply amounts to using a standard DL-Lite QA tool. The clear advantage of DL-Lite is that QA is very efficient, namely its data complexity is in AC_0 [1]. Thus the input of the QA tool consists of a set of annotated videos (an ABox), an ontology (a TBox) and a conjunctive query $q(\vec{x})$. The output is a set of answers X. Note that when $\vec{x}$ is empty, then $q(.)$ is a boolean query and its answer is either yes or no.

Example 1. (continued) *Consider the annotations of expert* E_1*. Thus* $\mathcal{A} = \{Male(f), Taksim(f), Enjoying(f)\}$*. Consider the boolean query* $q(.) \leftarrow \exists x, Interaction(x)$*. Since* $\mathcal{A}$ *contains* $Enjoying(f)$ *and* $\mathcal{T}$ *contains the axiom* $Enjoying \sqsubseteq Interaction$ *(i.e.,* Enjoying *is a type of* Interaction*), then the query answer is yes.* □

One may also want to rank-order videos instead of answers. This comes down to comparing sets of answers associated with videos. There are different strategies to rank-order these sets of answers. Due to the fact that answers are fully reliable, we simply use the cardinality criterion. This can be achieved by first collecting the answers associated with each video.

Definition 1. *Let* v_1, v_2 *be two videos. Then* v_1 *is presented to the user before* v_2*, denoted* $v_1 > v_2$*, iff* $|X(v_1)| > |X(v_2)|$*, where* $|X(v_i)|, i = 1, 2$*, is the size of the set of answers obtained from* v_i *and the ontology.*

Example 1. (continued) *Assume we have another ABox* $\mathcal{A}'$ *issued from another video* v' *and another expert, where* $\mathcal{A}' = \{Male(f'), Female(f'), Taksim(f')\}$ *and* f' *is a frame of video* v'*. Consider now a query* $q(x) \leftarrow \forall x, Interaction(x)$ *asking for all individuals that have* Interaction*. Querying* $\mathcal{A}$ *returns* $X(v) = \{f\}$ *while querying* $\mathcal{A}'$ *returns* $X(v') = \emptyset$*. Since* $|X(v)| > |X(v')|$*, then video* v *will be shown first to the user because its component better fits the query.* □

V. CONCLUSION

In this paper, we first proposed an ontology, represented in DL languages, to capture the cultural knowledge conveyed by a type of Malaysian traditional dances known as Zapin. We then proposed to enrich the ontology by annotating dance videos. A tool for manually annotating videos, based on ontologies, has been developed. Lastly, we proposed a strategy for query answering from annotated dance videos.

Space limitations prevent us from considering situations where several experts may annotate the same dance video, but they may disagree about some elements of the dance. Hence for a given video, these differing annotations may potentially cause conflicts between assertions of the ABox w.r.t. the TBox. In such case, the KB is said to be inconsistent. Moreover, experts can express their confidence in their annotations. Our annotation tool already boasts a functionality for assigning confidence degrees to annotations (Figure 5, item 4). The translation of an annotation coupled with a confidence degree into an assertion can be captured by applying a priority relation over the assertions of the ABox in order to express the fact that some assertions are more reliable than others. In future work, we plan to develop a query answering approach that handles conflicts in inconsistent KBs when the ABox is prioritized by way of a total preorder.

An avenue for further investigation consists in developing a method for revising ontologies in the presence of conflicting assertions issued from different experts. Another future work is to use our annotation tool to build a training set from a set of annotated videos, then to use machine learning techniques to train machine learning models to annotate videos automatically.

REFERENCES

[1] Alessandro Artale, Diego Calvanese, Roman Kontchakov, and Michael Zakharyaschev. The DL-Lite family and relations. *Journal of Artificial Intelligence Research*, 36:1–69, 2009.

[2] Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, and Riccardo Rosati. Tractable reasoning and efficient query answering in description logics: The DL-Lite family. *Journal of Automated Reasoning*, 39(3):385–429, 2007.

[3] Giannis Chantas, Sotiris Karavarsamis, Spiros Nikolopoulos, and Ioannis Kompatsiaris. A probabilistic, ontological framework for safeguarding the intangible cultural heritage. *Journal on Computing and Cultural Heritage*, 11(3):12:1–12:29, 2018.

[4] Annabel Clarance. A proposal for the creation of a dance ontology. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, 145:86–99, 2015.

[5] Wan Muhammad Afiq Wan Muhamad Fauzan and Sharifah Kartini Said Husain. Geometri dalam tarian zapin (Geometry in Zapin Dance). *Asian Journal of Environment, History and Heritage*, 2:331–342, 2018.

[6] Roman Kontchakov, Carsten Lutz, David Toman, Frank Wolter, and Michael Zakharyaschev. The combined approach to query answering in DL-Lite. In *KR, Toronto, Canada*, pages 247–257, 2010.

[7] Mohd Anis Md Nor. *Zapin: folk dance of the Malay world*. Oxford University Press, 2010.

[8] Mohd Anis Md Nor and Kendra Stepputat. *Sounding the dance, moving the music : Choreomusicological Perspectives on Maritime Southeast Asian Performing Arts*. Routledge, 2017.

[9] Mohd Anis Md Nor. Blurring Images, Glowing Likenesses: Old and New Styles in Traditional Dances of Malaysia. *Yearbook for Traditional Music*, 33:65, 2001.

[10] Norazit Selat, Hashim Awang, and Nor Hisham B. Osman. *Meniti zaman : masyarakat Melayu antara tradisi dan moden*. Akademi Pengajian Melayu, Universiti Malaya, Kuala Lumpur, 1997.

[11] Cokorda Pramartha, Joseph Davis, and Kevin Kuan. Digital preservation of cultural heritage: An ontology- based approach. In *The 28th Australasian Conference on Information Systems*, 2017.

[12] Ellya Roza and Nur Aisyah Zulkifli. The contribution of zapin as one of malay traditional arts in curriculum 2013. *Al-Ta'lim Journal*, 24(2):142–156, 2017.

[13] Abdelmoutia Telli, Ma Thi Chau, Mustapha Bourahla, Karim Tabia, and Salem Benferhat. An ontology for classifying vietnamese dance movements. In *International Conference on Pattern Recognition and Artificial Intelligence*, pages 23–29, 2018.