Knowledge Organization of Historical Events in Thailand

Ekkachai Sae-Jueng and Kulthida Tuamsuk

Information and Communication Department, Faculty of Humanities and Social Sciences, Khon Kaen University, Thailand

Keywords: Knowledge Organization, Historical Events, Thailand, Semantic Web, Linked Data, Getty Vocabularies.

Abstract: The purpose of this research is to analyse and organize the knowledge of historical events in Thailand in order to construct the knowledge structure and its relationship for further use in a linked data system. A qualitative research based on content analysis and development of knowledge organization approach is used. Related information resources included research publications from the e-Library of Thailand Research Fund, Getty Vocabularies, and Geographic Information System of Cultural Heritage Site of Thailand have been studied and used for the knowledge organization. The research results provide the knowledge structure of historical events in Thailand that contain five classes and relationships of the terms in each class and between classes. This can be further used for development of linked data and semantic web in the future.

1 INTRODUCTION

History learning can support to understand the changes and to recognize the civilization in which we live. Historical study is the method of systematically learning from past events to give an account of what has happened in the past (Carr, 2018). Historical studies effort to provide information and understanding of past historical, legal and policy events and this is certainly one of the reasons all modern nations encourage its teaching in some learning model. Because chronological event is a continuous, systematic description of past events as relating to the country, people, period, person, and phenomena that reflects the socio-cultural and economic aspects (Matthews, 1937), the knowledge of historical events of the nations can always been the valuable learning resources. In contexts of information science, knowledge organization research has been successfully used the content analysis approach integrating with technological techniques, such text mining, data analytics, ontology modelling, etc. to collect and manage the knowledge domain of valuable topics (Ibekwe-Sanjuan and Sanjuan, 2011). In this research, therefore knowledge organization of the knowledge on historical events of Thailand has been conducted so that the domain knowledge on the selected events can be scoped and structured, and then used for further studies in the field to ontology, linked data and semantic web by following the digital humanities research (Tuamsuk,

2015).

At present, many organizations have been developing the knowledge based system in historical contents and providing for open learning. For example, the development of "Museum Finland" or "Culture Sampo" by using Finnish Ontology Library Service, metadata schemas, and semantic technology tool such as ONKI – is a Finnish ontology library service, SAHA – is a browser based semantic annotation tool, POKA – is a framework for automatic annotation, and VERA - is a validation and quality assistant for Semantic Web data. The "Museum Finland" system can integrate the Finnish cultural heritage data from related information sources in the nation including museums, national library, and national archives and use a single search technique to retrieve and present the historical information by time dimension, location, map, and event chronicle (Hyvönen et al., 2009).

The history of Thailand has been recorded traced back about 800 years ago, divided into four kingdoms: Sukhothai, Ayutthaya, Thonburi, and Rattanakosin, Kingdoms. The learning resources for Thai history available in libraries and media are in the narrative and documentary formats. This research attempts to develop the tools for organizing and searching the knowledge on Thai history so that the data will be standardized and enabling for semantic search and linked with the well-known data sources such as Getty Vocabularies (Harpring and Baca,

2013) and Schema.org Vocabularies (Schema.org Community Group,).

This paper presents the research results of historical events knowledge structure in Thailand which was studied by using a content analysis and knowledge organization technique in order to identify the scope and related information resources. The knowledge structure of Thai historical events will be use for the development of semantic web for Thai historical events system (THES) in next research phase.

2 OBJECTIVES

This research was aimed at analyzing and organizing the knowledge of historical events in Thailand in order to identify the scope and develop the structure of the knowledge for further use as a resource for ontology and semantic web.

3 BACKGROUND AND RELATED RESEARCH

The history of Thailand comprised five periods by kingdom including the Sukhothai kingdom (1238–1538 A.D.), the Ayutthaya kingdom (1350–1767 A.D.), the Thon Buri kingdom (1767–1782 A.D.), the Rattanakosin kingdom (1782 A.D. – present), and Democratic period (1932 A.D. – present). In each period has important events that Thai people known. For instance, the Thon Buri kingdom under the reign of King Taksin the Great occurred between 1767 A.D. and 1782 A.D. It is 15 years there were 4 major events: 1) the recovery of Siam independence, 2) the kingdom establishment, 3) the kingdom enlargement, and 4) Thon Buri kingdom ending.

Nowadays information and communication technology or ICT was used for increasing chronological learning. There are numerous ICT-based learning platforms, for example, A WebQuest is an inquiry-oriented lesson format. It provides tools and medias for self-learning, multidisciplinary learning, external classroom, sharing learning experience, and virtual learning on its platform (Ott and Pozzi, 2011). Literature reviews found that there were many research projects based-on ICT for cultural heritage information management. For instance, the study "Culture Sampo: A National Publication System of Cultural Heritage on the Semantic Web 2.0" has developed

the Museum Finland on web platform which can associated cultural heritage information from museums, archives, and libraries in Finland. So it can presented the information in the features of maps, times, and history in sequential order (Hyvönen, 2009). The cultural heritage information presentation or "CHIP project" of Rijks Museum in Netherlands which is the artwork recommender based on semantic web technology. It can provide semantic browsing, searching, semantic recommendations (Wang et al., 2010).

This research uses the simple knowledge organization systems (SKOS) for knowledge classification and collection. It is a field of study to support the use of knowledge organization systems (KOS). There are many tools or categorize technics for KOS including thesauri, subject heading systems, Web Ontology language (OWL), classification schemes, taxonomies, Resource Description Framework (RDF) within the framework of the Semantic Web (Zeng and Mayr, 2018).

The Semantic Web or "Web of data" is useful to enhance web content by machine that can process in semantic way. The way to approach the Semantic Web is based on semantic annotations. It uses to describe the meaning of certain parts of web information. The major purpose of the Semantic Web is to enable machines(computers) to do more advantage work and to develop systems that can provide trusted collaborations over the computer network (Antoniou and Plexousakis, 2016). There are many technologies such SKOS, SPARQL, OWL, and RDF is a subset in Semantic Web technologies. It can support people to build data stores on the Web, create new vocabularies, and create rules for handling data.

In Thailand, the Princess Maha Chakri Sirindhorn Anthropology Centre developed the key archaeological sites system. This project used Dublin core metadata schema for data description and presented the information of archaeological sites on the common name, bibliography, geological context, topography, artifacts of the site, waterways, location, archaeological summary, route to site, tourism potential, responsible organization, etc. (SAC, 2019)

A number of projects have used the Getty vocabularies which comprise controlled terminology for arts, architecture, geographic names, artist names and bibliographic materials for linked data development. The Getty vocabularies are assembled to let their use in the linked data and can be published into the linked open data or LOD. The Getty vocabularies collect terminology in several subjects including Union List of Artist Names (ULAN), Thesaurus of Geographic Names (TGN) and Art and

Architecture Thesaurus (AAT) which are existing as LOD Getty Research Institute (2019).

AAT, TGN and ULAN are controlled vocabularies. For examples of each controlled vocabularies, TGN comprises place name, place descriptions, ancient cities, kingdoms, archaeological sites, physical features and etc., AAT contains terms, descriptions, and other attributes related to art, architecture, decorative arts, archival materials, conservation and etc., and ULAN contains person name, biographies, related people, and other attributes related to artists, architects, organizations (Getty Research Institute, 2019).

4 RESEARCH METHODOLOGY

A qualitative research based on content analysis and development of knowledge organization approach is used for this research which comprises of the following steps.

4.1 Identification of Information Resources

4.1.1 The e-Library of Thailand Research Fund

The research publications that contained the term "history" in the subject headings and keywords in its records were retrieved from the e-Library of Thailand Research Fund. It was found that there were 428 items. An analysis of data of each research items was conducted to identify whether there were research that have already focused on historical knowledge structure development. It was found that there were six research documents providing the knowledge structure on the historical contents relating to the specific topics of each study. The titles of these six documents were:

Doc.1- The study on relationship of the ancient through present culture for the development of cultural and civilization database for GMS and Malay Peninsula Regions Phase 1.

Doc.2- The study on relationship of the ancient through present culture for the development of cultural and civilization database for GMS and Malay Peninsula Regions Phase 2.

Doc.3- Exploration and sustainable heritage management in Pai-Pang Mapha- Khun Yuam Districts, Mae Hong Son Province

Doc.4- Living Angkor Road Project Phase I

Doc.5- Living Angkor Road Project Phase II

Doc.6- Research and development project on community-based museums, phase 2: Digital archives on community-based museums in Thailand.

The examples of knowledge structure from Doc.1 and Doc.2 are in Figure 1 and Figure 2.

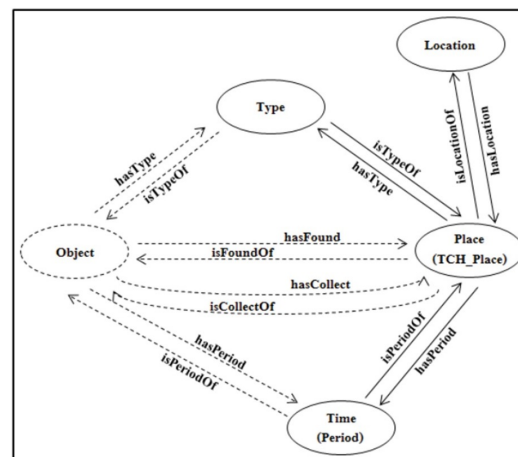Figure 1: Example of a knowledge structure from research literature (Doc.1)

Figure 2: Example of a knowledge structure from research literature (Doc.2)

4.1.2 The Study of Getty Vocabularies

The AAT, CONA, TGN, and ULAN of the Getty Vocabulary Program have been created, compiled, and disseminated in the Getty Research Institute, with the purpose of improving access to various information about art, architecture, decorative arts, archival materials, conservation, visual surrogates, and bibliographic materials. There are compliant with international standards and provide authoritative information for catalogers, librarian, researchers, software developer, and data providers (Getty Research Institute, 2019).

This research explored the Getty vocabularies structures including AAT (the Art and Architecture Thesaurus), TGN (Thesaurus of Geographic Names), ULAN (Union List of Artist Names) to identify the terms and concepts that can be used for historical events topic.

The relationship between controlled vocabularies can be shown in human-readable formats and machine-readable formats. For example, a human-readable format was used for human can read related contents about controlled vocabularies in a webpage format or hierarchy are format in Figure 3, and a machine-readable format was used for a computer programming, application programming interface(API) that its five most common formats: JSON, JSONLD, RDF/XML, N3/Turtle and N-Triples.

Figure 3: Associative relationships of ULAN (Image from http://vocab.getty.edu/ulan/500344436)

The Getty Vocabularies can serve in many format include XML/RDF, JSON, N-Turtles, Relational Tables, Linked Open Data (LoD), and through application programming interface (APIs). The figure

4 shows an example application of knowledge about "Person" in Thai historical events, Phraya Phichai (thes_person:1004) that linked with TGN and AAT.
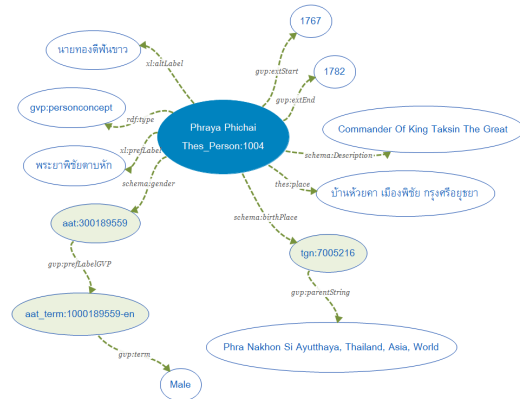

Figure 4: Example of person's linked data

4.1.3 Information Resources for Geographical Names

The Geographic Information System (GIS) of Cultural Heritage Site from the Fine Art Department of Thailand was used as the information resources for geographical names in Thailand. In this database, the data on cultural heritage site includes location, address, age, period, and important note as shown in Figure 5.

4.2 Knowledge Organization

The data, elements, attributes, and knowledge from all information resources were collected and analyzed by using SKOS principle. Similarity terms were grouped and categorized into the subject domains or classes as shown in Table 1.

The controlled vocabularies by the Getty vocabulary is matched with these knowledge structure in the context of historical events. The terminologies on "place" were derived from TGN, "object" were derived from CONA, and "person" were derived from ULAN.

5 RESULTS

The results of these analysis and organization of knowledge on the historical events in Thailand reveals that the knowledge can be classified into 5 classes. The details of each class are shown in Figure 6.

(1) **Period** comprises of name of the historical period and timespan of the period.

- Archaeological site(โบราณสถาน)
    - official name(ชื่อโบราณสถาน)
    - Other name(ชื่ออื่น(ถ้ามี))
    - Is World Heritage Sites(แหล่งมรดกโลก)
- Archaeological site(ลักษณะโบราณสถาน)
    - Religion and Beliefs(ศาสนาและความเชื่อ)
    - Type(ประเภทของแหล่ง)
    - Commercial site(แหล่งการพาณิชย์)
    - Government site(แหล่งสถานที่ราชการ)
    - Monument(อนุสาวรีย์หรืออนุสรณ์สถาน)
    - Cave/cliff (แหล่งถ้ำ/เพิงผา)
    - Old industry site(แหล่งอุตสาหกรรมโบราณ)
    - Ancient City(เมืองโบราณ)
    - Construction materials(วัสดุในการก่อสร้าง)
- Address(ที่ตั้ง)
    - Province(จังหวัด) / District(ตำบล/แขวง) / Postcode(รหัสไปรษณีย์)
- Age(ยุคสมัย)
    - Prehistory(ยุคก่อนประวัติศาสตร์)
        - Paleolithic(สมัยหินเก่า)
        - Metal Age(ยุคโลหะ)
    - Thailand history(ยุคประวัติศาสตร์)
        - Sukhothai period(สมัยสุโขทัย)
        - Ayutthaya period(สมัยอยุธยา)
        - Thonburi period(สมัยธนบุรี)
        - Rattanakosin period(สมัยรัตนโกสินทร
- Important(ความสำคัญ)
    - History note(ประวัติโบราณสถานหรือข้อมูลแหล่ง)
    - Architecture(ลักษณะทางสถาปัตยกรรม ศิลปกรรม หรือสภาพแหล่ง)
    - Current condition(สภาพปัจจุบัน และลักษณะการใช้งานปัจจุบัน)

Figure 5: Example of data in the GIS Cultural Heritage Site

Table 1: Analysis of knowledge structure from identified information resources.

Knowledge structure		Research project and Related software							Similar count	Require items
		Doc.1	Doc.2	Doc.3	Doc.4	Doc.5	Doc.6	GIS System		
Period	Period name	✓	✓			✓		✓	4	✓
	Timespan	✓	✓	✓		✓		✓	5	✓
Historical site / Place	Historical site name	✓	✓	✓		✓	✓	✓	6	✓
	Other name	✓	✓	✓				✓	4	✓
	Address	✓		✓	✓	✓	✓	✓	6	✓
	Province	✓			✓		✓	✓	4	✓
	Place type	✓	✓	✓		✓	✓	✓	6	✓
	is Registered	✓						✓	2	✓
	Geographic coordinates	✓		✓		✓			3	✓
	Image	✓		✓		✓		✓	4	✓
	Timespan	✓ Period	✓ Timespan	✓ Timespan		✓ Period			4	✓
	Aged							✓	1	✓
	Architecture type			✓		✓ Area type		✓	3	✓
	Current condition			✓		✓		✓	3	✓
	History note							✓	1	✓
	Religion and Beliefs	✓	✓		✓			✓	4	✓
Antique / Object		✓	✓			✓			3	✓

(2) **Event** comprises of name of the historical event and timespan of the event.

(3) **Object** comprises of name of object or antique in the historical event and its story (history note) that can be link to other controlled vocabularies.

(4) **Person** comprises of name of the person in the historical event and its story (history note) that can be linked to other controlled vocabularies.

(5) **Place** comprises of name of place or historical

1) Period
- Period name
- Timespan

2) Event
- Event name
- Timespan

3) Object/Antique
- Object name
- History note
- CONA id (http://vocab.getty.edu)

4) Person
- Person name
- History note
- ULAN id (http://vocab.getty.edu)

5) Place/Historical site
- Place name
- Other name
- Address
- Province
- Place type
- Is registered
- Geographic coordinator
- Image
- Timespan
- Age
- Architecture type
- Current condition
- History note
- Religion and beliefs
- TGN id (http://vocab.getty.edu)

Figure 6: Knowledge structure of Historical Events in Thailand.

site in the historical event, address, province, place type, historical registered status in Thailand, geographic coordinator, image, timespan, age, architecture type, current condition, religion and beliefs and its story (history note) that can be linked to other controlled vocabularies.

The sub-event class is a sub class of event class can be linked with every classes as the relationship between classes shown in Figure 7.
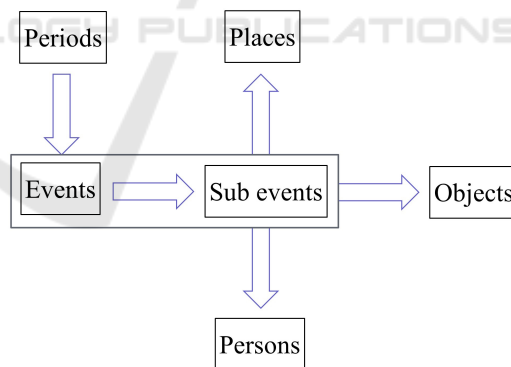
Figure 7: Relationship between classes.

6 CONCLUSION AND FUTURE WORK

The knowledge structure of historical events in Thailand resulted from this research that was collected and analyzed from many information sources. The results were based on existing standard dataset (The Getty vocabularies), therefore it can be

connected to other systems via linked data concept. Knowledge classification of historical events covers all components of the contents that can describe the historical events of Thailand. For example, the historical event on the recovery of Siam independence in the Thon Buri period comprises of the knowledge on "Person" such as King Taksin the Great and Nai (Mr.) Thong In. The knowledge on "Place" include the places such as the Pho Sam Ton Camp and the Chao Phraya River near Thonburi city. These "Person" and "Place" can be occurred in other historical events in the Thonburi period, and may be other Thai historical periods. These can be linked by using the relationship model between classes and sub-classes. When these data are published within linked data concept that it can be linked to other controlled vocabularies, for example, Thonburi city is in the TGN data (Subject_ID:7005221) in the Getty vocabulary; Thonburi is the preferred term for the concepts in similar terms Thon Buri and Dhonburi. It is currently located under the facet World, Asia continent, Thailand nation and Bangkok province.

Future development of this research will be the development of semantic web of historical events in Thailand, of which the knowledge structure in this study will be used for linked data (RDF) implementation.

ACKNOWLEDGMENTS

REFERENCES

Antoniou, G. and Plexousakis, D. (2016). *Semantic Web*, pages 1–5. Springer New York, New York, NY.

Carr, E. H. (2018). *What is history?* Penguin, UK.

Getty Research Institute, . (2019). *Getty Vocabularies*. http://www.getty.edu/research/tools/vocabularies/ (accessed January 20, 2019).

Harpring, P. and Baca, M. (2013). *Introduction to controlled vocabularies: Terminology for art*. architecture, and other cultural works.

Hyvönen et al. *2009*. The semantic web: Research and applications, Lecture Notes in Computer Science.

Ibekwe-Sanjuan, F. and Sanjuan, E. (2011). Knowledge organization research in the last two decades: 1988-2008. *arXiv preprint arXiv:1110.5419*.

Matthews, D. W. R. (2019). What is an historical event? *Proceedings of the Aristotelian Society*, 38:207–216.

Ott, M. and Pozzi, F. (2011). Towards a new era for cultural heritage education: Discussing the role of ict. *Computers in Human Behavior*, 27(4):1365–1371.

SAC (2019). *The development of database of key archaeological sites in Thailand*. Princess Maha Chakri Sirindhorn Anthropology Centre.

Schema.org Community Group, . *Schema.org vocabularies*. https://schema.org/docs/ (accessed January 20, 2019).

Tuamsuk, K. (2015). Digital humanities research at Khon Kaen University, Thailand. Vol. 9469:334-335.

Wang, Y., Wang, S., Stash, N., Aroyo, L., and Schreiber, G. (2010). Enhancing content-based recommendation with the task model of classification. In *International Conference on Knowledge Engineering and Knowledge Management*, pages 431–440. Springer.

Zeng, M. L. and Mayr, P. (2018). Knowledge Organization Systems (KOS) in the Semantic Web: a multi-dimensional review. *International Journal on Digital Libraries*, pages 1–22.