

# Semantic Web and ISO Standards to Create Multilingual Terminology e-Dictionary for Cultural Heritage: A Mini-Review

Tong Wei<sup>1, 2, 3\*</sup>, Christophe Roche<sup>2, 3</sup>, Maria Papadopoulou<sup>2, 3</sup> and Yangli Jia<sup>3</sup>

<sup>1</sup>Department of Information Management, Peking University, Beijing, China

<sup>2</sup>Condillac Research Group of LISTIC Lab, University Savoie Mont-Blanc, Rue Marcoz, France

<sup>3</sup>The School of Computer Science, Liaocheng University, Liaocheng, China

## Abstract

This paper presents the basic theory and methodology underlying the building of multilingual terminological e-dictionary in the domain of cultural heritage. A multilingual terminological e-dictionary is broadly defined as a reference resource that gathers, structures and describes in a systematic way linguistic data in a specific domain, by means of concepts (extralinguistic information) denoted by terms (linguistic information). The aim is to define cultural heritage terminology based on ontology. The results are: 1) a multilingual terminology e-dictionary in cultural heritage; 2) a new tool-assisted methodology for humanities scholars to build a multilingual terminology e-dictionary. In the context of digital humanities, it is believed that such an approach also contributes to the drafting of well-structured and more logically consistent definitions in natural language for specialized communication purposes.

## Keywords

Ontology • Terminology • e-Dictionary • Multilingual • Digital humanities  
• Cultural heritage

## Introduction

A multilingual terminology e-dictionary is a knowledge-based terminological product, which contributes to the communication and dissemination of domain knowledge. The multilingual terminology e-dictionary is broadly defined as a reference resource in a specific domain that gathers, structures, and describes linguistic data in a systematic way in one, two or more languages, in order to define concepts (extralinguistic information) denoted by terms (linguistic information) [1]. The theory and methodology underlying the construction of multilingual terminology e-dictionary come from the Semantic Web, Linked Data and ISO standards (ISO 1087 and ISO 704), which focus on systematizing both term systems that are about Terminology and concept systems that are about Ontology in the sense of Knowledge Engineering.

Cultural heritage is the legacy of physical artefacts and intangible attributes of a group or society that is inherited from past generations [2]. The lack of definitions of terms between different languages is a key obstacle to communication of cultural heritage. Libraries need to rely on a standard terminology system to establish consistent communication methods. Few terminology systems in heritage institutions are accepted because heritage institutions focus on building knowledge about past facts and times rather than heritage classification and behaviour. The collaborative work

between Terminology and Ontology provides new perspectives that create multilingual terminological resources to solve this obstacle in the digital age. In order to agree on terms and their meanings, terminology science postulates that the meaning of a term is the concept designated by that term. Perceived as extra linguistic, the concept is language-independent and can thus be shared. Terminology has, therefore, a double dimension: A linguistic dimension, through the terms which comprise it and a conceptual dimension, through the concepts designated by the abovementioned terms.

In this work, we want to present the theoretical and methodological aspects of building knowledge-based multilingual terminology e-dictionary and some suggestions in building multilingual terminological digital resources. It is believed that such an approach also contributes to the drafting of well-structured and more logically consistent definitions in natural language for specialized communication purposes. In addition, the multilingual terminology e-dictionary will be a new way of cultural communication on the Web. The organization of this paper is as following: Section 2 aimed to present the theory and methodology of building multilingual terminological e-dictionary; section 3 presented the literature review; section 4 introduced the multilingual terminological e-dictionary; section 5 is about the discussion of subjective; section 6 is about the conclusion of this paper.

## Literature Review

The multilingual terminological e-dictionary aims to meet the requirement of communication tool between different languages in the digital age. There are many cases in the medical fields [3-6]. EndoTerm project is a knowledge-based terminological resource, which aims to bridge that gap by contributing, on the one hand, to a more effective way of organizing and sharing the current knowledge regarding this particular medical condition and, on the other hand, to enhance its rather marginal presence in existing biomedical terminological resources [1]. The literature presents the French-language COVID-19 terminology international or localized, which de-terminologization is being accelerated by the popular media and leading to linguistic differences is the language situation [6].

There are some cases in cultural heritage [7-9]. OntoAndalus is an ontology of pottery artefacts of al-Andalus, which aims to serve as the conceptual backbone of a future terminological knowledge base [9]. Within

\*Address for Correspondence: Tong Wei, Department of Information Management, Peking University, Beijing, China; E-mail: weitong@pku.edu.cn

Copyright: © 2021 Wei T, et al. This is an open-access article distributed under the terms of the creative commons attribution license which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

Received date: December 02, 2021; Accepted date: December 16, 2021; Published date: December 23, 2021

the context of a FAIR terminology management, TriMED focuses on the structural standardization of an existing multilingual database, which is based on the Terminological Markup Framework (TMF) meta-model [10]. The European Commission (DG III) published the Multilingual Glossary of Popular and Technical Medical Terms [11]. The YourTerm MED22 project consists in the realization of a structured linguistic database based on the cognitive structure of the Frame-based terminology [12].

## Theory and methodology

**Ontology theory:** In the sense of Knowledge Engineering, ontologies are defined as specifications of a conceptualization [13]. A concept is used to structure and organize objects into sets and classes. Ontology is a formal expression of concepts for computer readable languages. In the sense of Knowledge Engineering, Ontology is used to structure knowledge domain. In the sense of the Semantic Web, ontologies aim to structure and publish data in machine-actionable formats for data integration and linked data. Ontology can be used as knowledge base in multilingual terminological e-dictionaries, which provide structure for domain knowledge for publishing data onto the Semantic Web.

**Terminology theory ISO 1087 and ISO 704:** Terminology is a tool for the dissemination and communication of cultural heritage. The main goal of defining any terminology (i.e., a list of terms in a domain of knowledge) reduces human communication ambiguity. The work of Terminology includes two dimensions: linguistic dimension and conceptual dimension. The linguistic dimension focuses on identifying terms, when conceptual dimension aims to identify concepts denoted by terms. In the multilingual terminological e-dictionary, the definition of a concept is based on the ISO 1087 standard on Terminology [14,15]. A concept is defined as a “unit of knowledge created by a unique combination of characteristics”, when a characteristic is defined as an “abstraction of a property of an object.” The ISO standard on Terminology put the focus on the notions of essential characteristic, a “characteristic which is indispensable to understanding a concept,” and of delimiting characteristic, an “essential characteristic used for distinguishing a concept from related concepts” [14]. The formal analysis of concepts establishes the relationship between terminology and ontology, which also is the methodology for building ontology in this article (Figure 1).

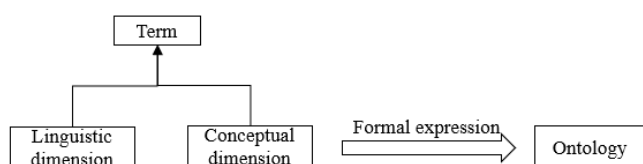


Figure 1. The relationship between terminology and ontology.

**Methodological:** The key for building multilingual terminological e-dictionary is to build a domain ontology. Domain ontology is to organize the domain knowledge and publish the open data on the Semantic Web. However, it is difficult to develop a domain ontology for humanities scholars who do not know formal languages and description logics. In a “term-and-characteristic” guided method is proposed, in order to create domain ontology of tangible cultural heritage [2] (Figure 2). The theory and methodological of this method is oriented from the Semantic Web, Linked Data, and ISO standards (ISO 1087-1 and ISO 704), which focuses on systematize both term systems that are about Terminology and concept systems that are about Ontologies-in the sense of Knowledge Engineering. The “term-and-characteristic” guided method reduces dependency on formal languages and description logics. It extends the building methods of domain ontology. Although the humanities scholars and ontology engineers could cooperate to develop ontology, this type of cooperation is difficult to implement in practice. In the future, the main work will focus on collaboration between humanities scholars and ontology engineers [16].

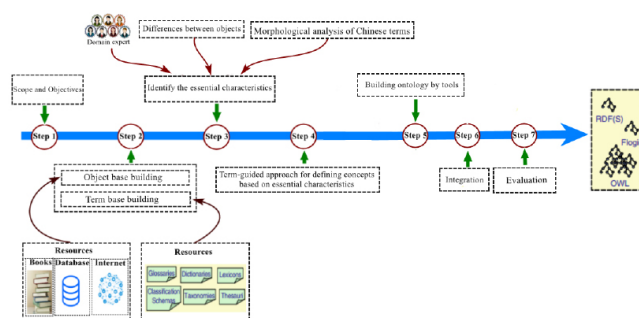


Figure 2. The workflow of “term-and-characteristic” guided method [2].

## Multilingual Terminological E-Dictionary

The terminology is a communication tool in different domain, which needs a standard terminology system to establish consistent communication methods. In the digital age, traditional terminological resources could not meet the requirement of humanity scholars. In the international environment, multilingual terminological tools have become a major demand. Therefore, the multilingual terminological e-dictionary will play an important role in communication of international. Especially, in the domain of cultural heritage, this kind of multilingual terminological e-dictionary will become a new terminology tool in the digital age.

The building of multilingual terminological e-dictionary is based on the Semantic Web and Ontology technology. The building of domain ontology depends on the domain knowledge, which is provided by the domain experts. So, in this context, the multilingual terminological e-dictionary as a new perspective needs to be collaboration between humanities scholars and IT, which is consistent with digital humanities. From this point of view, the building multilingual terminological e-dictionary would be an important work in the context of digital humanities.

## Discussion

The advent of the Semantic Web, Digital Humanities, and Linked Open Data initiative opens new perspectives for terminology work and IT applications. Let us quote content management systems, cross-language information retrieval, and multilingual terminological e-dictionaries. Especially, in the context of digital humanities, ontology and terminology should play an important role in solving the problem of semantic interoperability in cultural heritage data.

From a methodological perspective, it offers new perspectives for humanities scholars to build domain ontology.

From the evaluation perspective, there are many methods to evaluate the domain ontology. However, there is no standard evaluation method of the terminological resource production based on ontology.

## Conclusion

In this paper, we present the review of building multilingual terminological e-dictionary by using Semantic Web and ISO standards. We detail the theoretical background, which is the ISO Standard on Terminology (ISO 1087 and ISO 704). We describe the methodology for building domain ontologies for humanities scholars. Also, we introduce the state of the art for multilingual terminological e-dictionary in different domains. Last but not least, we discuss the open issues about multilingual terminological e-dictionary and propose several suggestions.

## References

1. Roche, Christophe, Rute Costa, Sara Carvalho and Bruno Almeida. "Knowledge-Based Terminological E-Dictionaries: The EndoTerm and al-Andalus Pottery Projects." *Terminol Int J Theor Appl Issues Spec Commun* 25 (2019): 259-290.
2. Wei, Tong, Christophe Roche, Maria Papadopoulou and Yangli Jia. "Using ISO and Semantic Web Standard for Building a Multilingual Terminology E-Dictionary: A Use Case of Chinese Ceramic Vases." *J Inf Sci* 1 (2021): 1-22.
3. Jucks, Regina, Rainer Bromme. "Choice of Words in Doctor-Patient Communication: An Analysis of Health-Related Internet Sites." *Health Commun* 21 (2007): 267-277.
4. Wilkinson, Mark D, Michel Dumontier, IJsbrand Jan Aalbersberg and Gabrielle Appleton, et al. "The Fair Guiding Principles for Scientific Data Management and Stewardship." *Sci Data* 3 (2016): 1-9.
5. Tercedor, Maribel, Juan Antonio Prieto Velasco. "Las Barreras En La Comunicacion Medico-Paciente: El Proyecto Varimed" 1 (2014): 1-22.
6. Bowker, Lynne. "French-Language Covid-19 Terminology: International or Localized?" *J Int Localization* 7 (2020): 1-27.
7. Wei, Tong, Christophe Roche, Maria Papadopoulou and Yangli Jia. "An Ontology of Chinese Ceramic Vases." *Int Conf Knowl Eng Ontol Dev* 1 (2020): 53-63.
8. Papadopoulou, Maria and Christophe Roche. "Building Ontology-Based Dictionaries for Greek Material Culture Terms." *ODOCH@ CAiSE* 1 (2019): 61-71.
9. Almeida, Bruno, Christophe Roche and Rute Costa. *Archaeological Classification and Ontoterminology: The Case of Islamic Archaeology of the Al-Andalus*. Chambéry: Université Savoie Mont Blanc, France, (2017).
10. ISO-16642, *Computer applications in terminology—Terminological markup framework*, CH: International Organization for Standardization, Geneva 2017.
11. *Multilingual Glossary of Technical and Popular Medical Terms in Nine European Languages*.
12. Vezzani, Federica and Giorgio Maria Di Nunzio. "On the Formal Standardization of Terminology Resources: The Case Study of TriMED." *TriMED* 1 (2020):1-12.
13. Gruber, Thomas R. "A Translation Approach to Portable Ontology Specifications." *Knowl Acquis* 5 (1993): 199-220.
14. ISO 1087-1, *Standard ISO 1087: Terminology-Vocabulary*, International Standards Organization, Geneva 2000.
15. ISO 704, *Terminology Work – Principles and Methods*, International Standards Organisation, Geneva 2009.
16. Westerinen, Andrea and Rebecca Tauber. "Ontology Development by Domain Experts (Without Using the "O" Word)." *Appl Ontol* 12 (2017): 299-311.

**How to cite this article:** Wei, Tong, Christophe Roche, Maria Papadopoulou, and Yangli Jia. "Semantic Web and ISO Standards to Create Multilingual Terminology e-Dictionary for Cultural Heritage: A Mini-Review." *J Sens Netw Data Commun* S5 (2021): 114.