HOME: Hybrid Ontology Mapping Evaluation Tool for Computer Science Curricula

Chayan Nuntawong¹, Chakkrit Snae Namahoot¹ and Michael Brückner²
¹Department of Computer Science and Information Technology, Faculty of Science, Naresuan University, Thailand.
²Department of Educational Technology and Communication, Faculty of Education, Naresuan University, Thailand.
chakkrits@nu.ac.th

Abstract—This paper presents a hybrid ontology mapping tool for evaluating the standard of computer science subjects against the Thailand Qualification Framework for Higher Education (HQF: HEEd). This can improve the standard of curriculum of universities in Thailand with higher accuracy and enable the decrease of processing time. Three ontologies have been designed: course, TQF: HEEd and the standard curriculum of computer science. They were used for comparing course contents by applying a combination of ontology mapping techniques (semantic-based using extended Wu & Palmer’s algorithm and structure-based using SKOS features). Test with the sample data show that the tool based on a hybrid ontology mapping worked sufficiently well and can inform the efforts for curriculum improvement.

Index Terms—Ontology; Ontology Mapping; Semantic Similarity; Qualification Framework; Web Application; String Algorithm.

I. INTRODUCTION

One of the problems in the management of higher education in Thailand is the variety of standards for curriculum design and teaching quality at Thai universities. Most students try to enter famous universities that can provide the best quality of education for them. However, each curriculum amongst the well-known universities can receive only limited number of students. Students who miss the cut off admission have to look for other universities, which may not match the teaching quality laid out in the Thailand Qualification Framework (TQF). If this is the case, students have lesser prospects of good careers after their graduation. This is one of the vital problems relating to the quality of standard education. Therefore, the Ministry of Education, by the Office of the Higher Education Commission (OHEC) [1] has developed a Thailand Qualification Framework for Higher Education (TQF: HEEd) to be used by all universities in Thailand as a framework to improve the quality of teaching and curricula [2]. Given that Thailand is entering the ASEAN Economic Community (AEC) [3] in 2015, collaboration in educational sector is one of core areas of policy for increasing relationships and fostering career exchanges in the community, so that students can transfer credits to every university in ASEAN [4]. The standard of education is one of the important issues that must be addressed.

However, in practice, there are problems evaluating the quality of the curricula of Thai universities because they differ from one to another. The complicated process of curriculum improvement is still one of the problems, which consume a lot of resources, effort and time. Even though every university in Thailand tries to improve their curriculum, some universities still have issues with TQF: HEEd or there are different patterns from other universities relating to the same curriculum. This is due to the reason that some universities use different word variations in the course syllabus, course descriptions and teaching guidance: In fact, those variations represent the same meaning.

Ontology is a key to represent education information and to find the accordance of contexts in each ontology using various techniques in Ontology Mapping, such as structure-based, linguistic-based, property relation-based, instance-based and constraint-based [5]. This work is similar to the work that shows different forms of data in e-Learning domain that can be shoveled by Wongkalasin and Archint [6], but they lack of curriculum standard. Additionally, there has been a work on the structure-based Ontology Mapping tool for Computer Science curriculum and TQF: HEEd comparison presented by Nuntawong and Snae [7] still cannot guarantee the accuracy in mapping process. The technique seem to produce insufficient results due to a small sample used in practice. Further, Nuntawong et al. [8] presented the semantic similarity technique using Wu & Palmer’s algorithm for mapping process. Although the testing results were an extended work, it only covered semantic similarity, hence it cannot guarantee the output as a whole, which still needs a structure based on the Ontology Mapping technique to improve the accuracy of mapping results and performance.

The semantic web language mostly used in Knowledge Organization System (KOS) for knowledge representation of thesaurus or classification schemes is the SKOS (Simple Knowledge Organization System) [9] since the structure and property of language are appropriate for knowledge representation in defining terminology. Defined as the standard by W3C, the SKOS is popular in vocabulary work or thesaurus, such as the AGROVOC [10] the food and agriculture organization of the United Nations (FAO), and the ACM Computing Classification System (CCS) [11].

The SKOS have been used in ontology research, such as Shi [12] used SKOS to design a water resource ontology. The water resource and water usage data were designed to ease a search in terms of thesaurus and categories. Zoghlami et al. [13] have developed a web application tool that can create and exchange knowledge resources using SKOS for a specified vocabulary dataset. The dataset contains properties of terminology such as word, derivation, description, origin and example of word used in the form of sentences. In addition, Salama [14] has studied a tool to create a semantic web for Nubian language, and then designed a terminology ontology using SKOS features. The system can be used to search Nubian character using SPARQL query and present the translation of Nubian words in English and Arabic language. It can be seen that the research mentioned above.
used the SKOS features to store the property of words to represent knowledge, and the property of SKOS was designed in the form of RDF (Resource Development Framework) standard.

From the problems and literature mentioned above, we developed a tool that can determine the consistency between a given computer science curriculum and TQF: HEd using a combination of ontology mapping techniques. The extended Wu & Palmer’s algorithm in semantic-based ontology mapping and the SKOS features in structure-based ontology mapping were used for evaluating the courses. We have designed standard of curriculum data based on SKOS and used them in the comparison process.

II. ONTOLOGY DESIGN

In this paper, we used an ontology of TQF: HEd in computer science (OTQF) which has been designed previously [7]. The OTQF (Figure 1) contains classes of all courses in the curriculum and each class contains the “Knowledge Area” class, e.g. Intelligent Systems, Programming Languages, Human-Computer Interaction, etc. Each knowledge area class also contains the “Body of Knowledge” class which can be defined as follows: “Fundamental Issue”, “Basic Search Strategies”, “Problem Solving” and “Knowledge-based Reasoning”.

Figure 1: An example of TQF: HEd ontology in Intelligent Systems knowledge area

For curriculum ontology (OCC), the developed tool can import the curriculum data in English language and automatically convert them to an ontology format.

In this paper, we used the knowledge based of standard curriculum in computer science (SKOSCC) as the standard data in both semantic-based and structured-based ontology mapping processes. The SKOSCC is extracted from the Computer Science Curricula 2013 (CS2013) [15] written by ACM and IEE, which represents a standard guideline for curriculum design. In this research, we converted the SKOSCC into an ontology format using Protégé [16] and the example of SKOSCC ontology is shown in Figure 2.

According to Figure 2, to use the features of SKOSCC, we need to import SKOSCC into a format of the SKOS language, which is similar to the ontology format. Figure 3 shows components of SKOSCC in a format of the SKOS language. Each topic has a Concept ID (unique) stored in <NamedIndividual> tag. A topic name (Problem Solving by Search) is stored in <skos:PrefLabel> tag, which has synonyms such as “Investigative by Search”, “Diagnosis by Search”, and “Problem Solving by Investigation”, stored in the <skos:AltLabel> tag. Additionally, it also has a parent topic as a BROADER ID stored in <skos:broader> tag.

Figure 2: Example of SKOSCC in a format of ontology

III. SYSTEM ARCHITECTURE

Figure 4 illustrates the system architecture, which can be described as follows:

1. Import the computer science curriculum that contains the course description, course syllabus or teaching guidance of each subject in English language and convert it to the course ontology (OCC).
2. Match OCC and OTQF using semantic-based ontology mapping process with extended Wu & Palmer’s algorithm and WordNet [17] (see subsection 3.1). SPARQL query is used in this process in order to extract information from both ontologies.
3. Match OCC and OTQF using structure-based ontology mapping process with SKOS features (see subsection 3.2).
4. Get the result of the subject mapping from 2 and 3 with details, weight values of course description, and percentage of overall correspondence.
A. Semantic-based Ontology Mapping

This technique takes the subjects with course descriptions in OCC to find synonyms related to the topics in the course descriptions from SKOSCC. Then, the synonyms are used to compare with the OTQF for standard course description evaluation. If no synonyms is found, the system then looks for topics in WordNet by segmenting words and retrieves the list of synonyms of each word. The group of synonyms is used to calculate the semantic similarity values using the extended Wu & Palmer’s algorithm [8], which is represented by (1)

$$Sim_{matrix} = \sum_{i=1}^{m} \sum_{j=1}^{n} \{ Sim_{wp}(Word_{wordNet}, Topic_{O}) \} \times 100 \quad (1)$$

where Word_{wordNet} is the word from WordNet, Topic_{O} is the word from the class name in OCC or OTQF.

The algorithm creates a matrix with size m x n, where m is the number of word in group of words from WordNet and n is the word in group of words from OCC or OTQF. The results are calculated based on the similarity values from each column, and then the semantic similarity values of the group of words are summarized.

We used only a group of words that has a maximum average of semantic similarity values (values that have the most similarity with the original group of words from the class name) to compare the OCC and OTQF again in the structure-based Ontology Mapping process.

The algorithm of the Semantic-based Ontology Mapping is shown in Figure 5 as pseudo code.

B. Structure-based Ontology Mapping

In the process described in the previous subsection, synonyms were used to find the level of correspondence. This is because in some cases, the topic of the subject corresponds with the body of knowledge in its semantic meaning, but it is not in the correct position in the TQF: HED structure.

First, the tool will check the position in each class of both the OCC and OTQF by making comparison with the SKOSCC (as in the process described in previous subsection). Each class of OCC and OTQF has two types of levels: node level and parent node level. Then we proceeded in the Structure-based Ontology Mapping technique as follows (see Figure 6).

```python
FUNCTION semanticMappingOnto(OCC, OTQF, SKOSCC)
    Topics <- Word of node_OCC and node_OTQF
    Word_occ <- Synonym word from WordNet
    FOR ALL Topics, DO
        IF Topics = word of node_SKOSCC THEN
            nodeLevel of node_SKOSCC = nodeLevel of Topics
            parentNodeLevel of node_SKOSCC = parentNodeLevel of Topics
        END IF
        RETURN Topics.MATCH WITH SEMANTIC STANDARD
    END FOR
    END FUNCTION
```

Figure 4: The system architecture

![Figure 4: The system architecture](image)

**Figure 4**: The system architecture

**Figure 5**: A Semantic-based Ontology Mapping algorithm

```python
FUNCTION structureMappingOnto(OCC, OTQF)
    FOR ALL node_OCC DO
        IF nodeLevel of node_OCC = nodeLevel of node_OTQF THEN
            IF parentNodeLevel of node_OCC = parentNodeLevel of node_OTQF THEN
                RETURN node_OCC.MATCH WITH node_OTQF correspondNode_OCC = 1
            END IF
        ELSE
            RETURN node_OCC.DOES NOT MATCH WITH node_OTQF
        END IF
    END FOR
    END FUNCTION
```

**Figure 6**: A Structure-based Ontology Mapping Algorithm

The tool compares all classes in OCC to find the correspondence as the structure-based until it gets the overall correspondence value in a course.

Figure 7 shows an example of the Structure-based Ontology Mapping process in “Problem Solving” topic of the “Artificial Intelligence” course from OCC.

First, we compared this class from OCC with SKOSCC and found that it corresponds with the “Problem Solving by Search” topic from SKOSCC that has the “Basic Search Strategies” as a parent node. Next, we compared the OTQF with SKOSCC and found no class match with “Problem Solving by Search” in SKOSCC; but the closest match is “Basic Search Strategies” topic, which corresponds with the “Basic Search Strategies” topic from SKOSCC and has “Intelligent Systems” as a parent node.

When we compared the OCC and OTQF, we found no class in OTQF that matches with “Problem Solving” topic from OCC. So, we considered the parent node level and found that this corresponds with “Basic Search Strategies” in OTQF as a child node.
Finally, we derived from the examples that the “Problem Solving” class from OCC has a correspondence as a child node with “Basic Search Strategies” class from OTQF.

IV. TESTING AND RESULTS

Figure 8, 9, and 10 show the Interface of the hybrid ontology mapping evaluation tool (HOME) for evaluating the standard of computer science subjects against the Thailand Qualification Framework for Higher Education (TQF: HEEd). For testing the tool, we used some course syllabuses from various subjects in undergraduate computer science curriculum from universities involved in the course evaluation according to the TQF: HEEd. Figure 8 shows the subject evaluation using the course syllabus of software engineering and the result shows all correct mapping in both semantic-based and structure-based ontology mappings with 100% of overall correspondence. This means that the subject of software engineering matches the standard of the TQF: HEEd evaluation, since all topics in the course description followed the semantic writing and course structure standards. The weight values indicate how much the course of software engineering has stressed the importance of content in teaching in each body of knowledge, e.g. software design is a very important piece of knowledge and should be emphasized in teaching due to the highest weight value.

However, Figure 9 shows a case where the content is not compliant with the standards using the semantic based technique. The subject descriptions of artificial intelligence were tested and the result shows that the content was not within the standard of TQF: HEEd, e.g. missing topics in the category of knowledge-based reasoning. The tool returns the introduced topics, which should be included in this subject to meet the standard of TQF: HEEd. Users can add these topics to the course description by clicking the “add” button. It is also found that by using the structure based technique, there is one topic that did not comply with knowledge. In other words, the topic of “AI programming with LISP” was not relevant to the knowledge-based reasoning and was supposed to be taught in another course instead. The tool allowed users to edit or delete this topic from the course syllabus and suggested to move the topic to the body of knowledge of Programming language or Functional Programming instead.

Figure 10 shows the content of an Artificial Intelligence subject that has been modified in accordance with the standard. The user has edited the content according to the suggestion shown in Figure 9 and clicked the “Re-Calculate” button for reevaluating and this time, the result shows 100% of overall correspondence with all correct mappings applying both ontology mapping techniques.
The results of weight values can be used for the description data from the computer science curriculum. The test results the accuracy improvement of the mapping process. SKOS features to determine the correctness of body of the standard. The structure may be written with synonyms of terms but correspond with algor based ontology mapping use the computer science subjects and TQF: HEd. The semantic techniques for standards determining and correspondence of evaluation tool, which combines two ontology mapping algorithms. In this paper, we present a hybrid ontology mapping evaluation tool, which combines two ontology mapping techniques for standards determining and correspondence of the computer science subjects and TQF: HEd. The semantic-based ontology mapping used the extended Wu & Palmer’s algorithm to find the semantic similarity of course details that may be written with synonyms of terms but correspond with the standard. The structure-based ontology mapping used SKOS features to determine the correctness of body of knowledge categories. The standard curriculum data from the TQF: HEd and Computer Science Curricula written by ACM and IEEE were imported into the ontology format to suggest the accuracy improvement of the mapping process. The test results were encouraging; however, we plan to extend the features of the tool to test and analyze all course description data from the computer science curriculum. Moreover, the results of weight values can be used for the evaluation of teaching objectives and comparison of academic trend amongst universities in Thailand.

**REFERENCES**


Figure 10: The Interface testing with an Artificial Intelligence subject which have already modified

**V. CONCLUSION**

In this paper, we presented a hybrid ontology mapping evaluation tool, which combines two ontology mapping techniques for standards determining and correspondence of the computer science subjects and TQF: HEd. The semantic-based ontology mapping used the extended Wu & Palmer’s algorithm to find the semantic similarity of course details that may be written with synonyms of terms but correspond with the standard. The structure-based ontology mapping used SKOS features to determine the correctness of body of knowledge categories. The standard curriculum data from the TQF: HEd and Computer Science Curricula written by ACM and IEEE were imported into the ontology format to suggest the accuracy improvement of the mapping process. The test results were encouraging; however, we plan to extend the features of the tool to test and analyze all course description data from the computer science curriculum. Moreover, the results of weight values can be used for the evaluation of teaching objectives and comparison of academic trend amongst universities in Thailand.